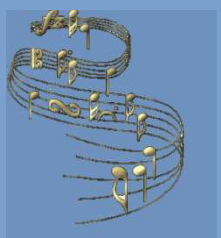


UNIVERSITEIT GENT

Faculteit Geneeskunde en Gezondheidswetenschappen

Development and validation of the ‘Voice Handicap Index aangepast aan de zangstem’

Scriptie voorgedragen tot het behalen van de graad van master in de logopedische en audiologische wetenschappen

Student: Anneleen D’haeseleer

Promotor: Prof. dr. Marc De Bodt
Copromotor: Prof. dr. Dominique Morsomme

Academiejaar: 2010 - 2011

Table of contents

Abstract: English

Purpose: The purpose of this study was to develop a reliable and valid version of the 'VHI adapté aux chanteurs' (Morsomme et al., 2007), an instrument to evaluate the impact of voice disorders for the specific population of classical singers, in Dutch: 'VHI aangepast aan de zangstem'.

Introduction: The Voice Handicap Index (VHI) (Jacobson et al.,1997) is an instrument for self-assessment, designed to evaluate the impact of voice disorders on patient's quality of life and the consequent level of handicap. However, this questionnaire is not sensitive enough for the population of singers. Therefore, Morsomme et al., (2007) created the 'VHI adapté aux chanteurs' in French and Cohen et al., (2007) created the Singing Voice Handicap Index (SVHI) in English. The adaptations of the VHI to the population of singers were performed in different languages (for example: Swedish, Spanish,…), however, until present, not in Dutch.

Methods: Participants of this study were divided into two groups: 85 classical singers without voice disorders (test group) and 5 classical singers with voice disorders (control group). After translating the 'VHI adapté aux chanteurs' (Morsomme et al.,2007) into the Dutch version 'VHI aangepast aan de zangstem', several statistical analysis were performed. To standardize the test we investigated the test-rest reliability by performing a Wilcoxon Signed Ranks Test, the Intraclass Correlation Coefficient (ICC) and the Spearman's Rho coefficient. The Cronbach's alpha was calculated to investigate the internal consistency of the questionnaire. The Wilcoxon Signed Ranks Test was performed to investigate the validation of the test. Finally the effects of the variables: 'gender', 'age', 'smoking', 'singing style' and 'following or having followed singing classes' were examined by the Mann-Whitney U-Test and the variables: 'grade of profession', 'type of profession' and 'classification of the singing voice' were examined by the one-way ANOVA (after performing the Levene's Test).

Results: A good test-retest reliability was shown by the ICC values (total score: ICC=0.744 ; subscores range: 0.738 – 0.651 ; $p<0.001$) and Spearman's Rho correlation coefficients (total score: $r_s=0.794$; subscales range: $r_s=0.747$ – $r_s=0.782$ ($p<0.001$)). The Wilcoxon Signed Ranks Test showed that all scores on the retest were significantly lower than those on the test (total score: 12; subscores range: 6 to 9). We obtained good Cronbach's alpha values for the total score (r=0.875) and subscores (range: r=0.687 - r=0.775), which means that the questionnaire had a good internal consistency. The Mann-Whitney U-Test showed that the questionnaire is sensitive and specific, as significant mean differences ($p<0.001$ – $p=0.006$) between all scores

of both groups were obtained, where singers with voice disorders had systematically higher scores than singers without voice disorders. The same test showed significant mean differences between test and retest scores of both groups for total scores (p=0.016) and emotional (p=0.022) and physical (p=0.039) subscores. Only the variables 'retesting' (p<0.001) and 'grade of profession' (F-scale: p=0.039 ; P-scale: p=0.027) had significant influences on the results of the questionnaire.

Conclusion: The Dutch questionnaire 'VHI aangepast aan de zangstem' is a reliable and valid clinical tool to evaluate the impact of voice disorders, which can be used for the population of classical singers.

Abstract: Nederlands

Doelstelling: De hoofddoelstelling van deze thesis was om een betrouwbare en valide versie te ontwikkelen van de 'VHI adapté aux chanteurs' (Morsomme et al., 2007), een instrument om de impact van stemstoornissen te evalueren bij de specifieke populatie van klassieke zangers, in het Nederlands: 'VHI aangepast aan de zangstem'.

Inleiding: De Voice Handicap Index (Jacobson et al., 1997) is een self-assessment instrument dat ontwikkeld is om de impact van stemstoornissen op de kwaliteit van leven en het niveau van gepercipieerde handicap van de patiënt na te gaan. Deze vragenlijst is echter onvoldoende sensitief voor de specifieke populatie van zangers. Daarom ontwikkelde Morsomme et al., (2007) de Franstalige vragenlijst 'VHI adapté aux chanteurs', aangepast aan de specifieke populatie van klassieke zangers. Ook Cohen et al., (2007) ontwikkelde een Engelstalige vragenlijst: 'Singing Voice Handicap Index (SVHI) voor de volledige populatie zangers. In navolging van Morsomme et al., (2007) en Cohen et al., (2007) werden beide vragenlijsten vertaald naar verschillende andere talen (Vb.: Zweeds, Spaans,…). Tot op heden werd echter geen Nederlandstalige versie ontwikkeld.

Methods: De proefpersonen in deze studie werden onderverdeeld in twee groepen: 85 klassieke zangers zonder stemstoornissen (test groep) en 5 klassieke zangers met stemstoornissen (controle groep). Na vertaling van de 'VHI adapté aux chanteurs' (Morsomme et al., 2007), werden verschillende statistische analyses uitgevoerd. Om de test te standaardiseren werd de test-hertest betrouwbaarheid nagegaan door berekening van de Wilcoxon Signed Ranks Test, de Intraclass Correlation Coefficient (ICC) en de Spearman's Rho correlatie coëfficiënt. Ook de Cronbach's alpha werd berekend om de interne consistentie van de vragenlijst na te gaan. De validiteit werd nagegaan door de Wilcoxon Signed Ranks

Test. Tenslotte werd de invloed van de variabelen: ‘geslacht’, ‘leeftijd’, ‘roken’, ‘zangstijl’ en ‘zangles volgen of gevolgd hebben’ berekend door de Mann-Whitney U-Test. De resterende variabelen: ‘professionele graad’, ‘type van professie’ en ‘classificatie van de zangstem’ werden berekend door de one-way ANOVA (na berekening van de Levene’s Test).

Resultaten: De ICC waarden (totale score: ICC=0,744; subscores: range: 0,738 – 0,651) en Spearman’s Rho correlatie coëfficiënten (totale score: r_s=0,794; subscores: range: 0,747 – 0,782 (p<0,001)) toonden een goede test-hertest betrouwbaarheid aan. De Wilcoxon Signed Ranks Test toonde aan dat alle scores van de hertest significant lager waren dan de scores van de test (totale score:12; subscores: range: 6 – 9). Er werden goede Cronbach’s alpha waarden aangetoond (totale score: r=0,875; subscores: range: r=0,687 – r=0,775), wat wijst op een goede interne consistentie van de vragenlijst. Sensitiviteit en specificiteit werden aangetoond door significante verschillen tussen alle scores van beide groepen (range: p<0.001 – p=0.006), waarbij zangers met stemstoornissen systematisch hogere scores behaalden dan zangers zonder stemstoornissen. Er werden ook significante gemiddelde verschillen aangetoond tussen test en hertest van beide groepen, en dit voor de totale scores (p=0.016), de emotionele (p=0.022) en de fysieke subscores (p=0.039) (Mann-Whitney U-Test). De enige variabelen die een significante invloed hadden op de vragenlijst, waren de variabele hertesting (p<0,001) en de variabele graad van professie (F-schaal: p=0,039; P-schaal: p=0,027).

Conclusie: De Nederlandstalige vragenlijst: ‘VHI aangepast aan de zangstem’ is een betrouwbaar en valied klinisch instrument om de impact van stemstoornissen te evalueren. De vragenlijst is bruikbaar voor de populatie van klassieke zangers.

Introduction

Measuring the severity of a voice disorder is difficult. Methods have ranged from subjective measures of voice disorder severity including perceptual judgements (e.g., GRBAS(I) (Hirano et al.,1981) to objective measures of voice characteristics (e.g., acoustic and aerodynamic measurements, Dysphonia Severity Index (DSI) (Wuyts et al., 2000). Traditional clinical examination of the larynx or objective acoustic and aerodynamic measures of voice characteristics have been used to assess medical, surgical or behavioral interventions. (Woodson et al., 1992). The perceptual assessment is the eldest and most common method to examine voice disorders. Most experts (Dejonckere et al., 1993; Hirano, 1981; Baken, 1987; Sataloff, 1996; Kearns and Simmons, 1988; Isshiki and Takeuchi, 1970) agree that perceptual assessment is irreplaceable, but they also underline the limitations about this way of assessment (e.g., they are poorly standardized (Sataloff, 1996) and there is a lack of general accepted concepts for a descriptive terminology (Baken, 1987)). However, perceptual assessment remains the primary instrument for differential diagnosis and clinical management (Kearns and Simmons, 1988). Auditory perceptual assessments are often the Golden Standard to assess voice disorders in clinical decisions and are the standard where the objective measurements are being compared to (De Bodt et al., 2009). Although these methods can yield valuable data, they do not provide insight into why patients with similar voice disorders experience different levels of handicap and disability.

Therefore, various health status instruments have been designed to measure the impact of voice problems on patient quality of life and the consequent level of handicap. The World Health Organization (1980) defines a handicap as “a social, economic, or environmental disadvantage resulting from an impairment or disability” and a disability as “a restriction or lack of ability manifested in the performance of daily tasks”. Outcome research places the focus of the impact of disease and subsequent treatment on the patient’s subjective perspective as opposed to the results of a clinical examination (subjective assessment) or an objective test which can measure pre- and post treatment changes (Benninger et al., 1997).

Different instruments of self-assessment of voice problems exist, for example the Voice Handicap Index (VHI) (Jacobson et al., 1997), the Voice Activity and Participation Profile (VAPP) (Ma and Yiu, 2001),…

The voice handicap index is a self-administered questionnaire that consists of 30 statements. The subject has to respond according to the appropriateness of each item (0=never, 1=almost never, 2=sometimes, 3=almost always, 4= always). This questionnaire was constructed with a three subscale structure: a functional scale, a physiological scale, and an emotional scale. Each subscale contains 10 items and is worth 40 points. The VHI score varies between 0 and 120 with the latter representing the maximum perceived disability due to voice difficulties. In 2002, the Agency of Health Care Research and Quality officially recognized the VHI to be a 'reliable and valid diagnostic tool'.

Franic et al. (2005) made a comparison between different clinical tools that assess voice problems. They concluded that the VHI (Jacobson et al., 1997) was more reliable and more easy to use, compared to eight other health status instruments (the Voice Activity and Participation Profile (VaPP) (Ma and Yiu, 2001), the Voice-Related Quality of Life Measure (V-RQOL) (Hogikan and Sethuraman, 1999), Voice Clinic Quality of life Questionnaire (VC-QOLQ) (in Morsomme et al. (2007), Voice Disability Index (VDI) (in Morsomme et al., (2007), Voice Outcome Survey (VOS) (Gliklich, 1999), Pediatric Voice Outcome Survey (Pediatric VOS) (Hartnick, 2002), University of Michigan Head and Neck Quality of Life Instrument (HNQOL) (in Morsomme et al. (2007), University of Michigan Quality of Life Questionnaire Revised(UW-QOL-R) (in Morsomme et al., (2007). The VHI was developed to measure the patient's perception of disability due to any type of voice disorder. Its internal consistency, test-retest reliability and construct validity has been proved (Jacobson et al., 1997). Hsiung et al., (2002) showed that an instrument as the VHI is crucial as a completion to objective measurements to make a complete voice report, targeting different aspects of a voice disorder.

Several new versions of the VHI were derived from the original VHI of Jacobson et al. (1997). These are displayed in **Table 1**.

Table 1: VHI and new versions derived from the VHI

Table 1							
Name	**VHI**	**VHI-10**	**VHI adapté aux chanteurs**	**SVHI**	**SVHI-10**	**The Swedish version of the Voice Handicap Index adapted for singers: RHI-S**	**S-VHI**
Author	Jacobson et al.	Rosen et al.	Morsomme et al.	Cohen et al.	Cohen et al.	Lamarche et al.	Garcia-López et al.
Publication Year	1997	2004	2007	2007	2009	2009	2010
Journal	American Journal of Speech-Language Pathology	The Laryngoscope	Rev Laryngol Otol Rhinol	Annals of Otology, Rhinology & Laryngology	The Laryngoscope	Logopedics Phoniatrics Vocology	Acta Otorrino-laringológica Española
Language	English	English	French	English	English	Swedish	Spanish
Sample: study group	65 consecutive patients from a Voice Clinic	3 groups of dysphonic patients: group 1: 100 group 2: 59 group 3: 819	192 singers: 54 dysphonic 138 nondysphonic	112 dyphonic singers	group 1: 297 dysphonic singers group 2: 91 dysphonic singers	96 nondyphonic singers	29 dysphonic singers
Sample: control group	No control group	159 nondysphonic individuals	27 nondysphonic non-singers	129 nondysphonic singers	99 nondysphonic singers	30 dysphonic singers	81 nondysponic singers
Population (standerdized)	Persons with or without voice disorders	Persons with or without voice disorders	Classical singers	Variety of singers	Variety of singers	Variety of singers	Classical singers, light or pop
Adapted to the singing voice	No	No	Yes	Yes	Yes	Yes	Yes
Scale	5-point Likert scale	5-point Likert scale	5-point Likert scale	5-point Likert scale	5-point Likert scale	5-point Likert scale	5-point Likert scale
Subscales	3	3	3	0	0	3	3
Number of items in each subscale	10	F-scale: 5 E-scale: 2 P-scale: 3	10	/	/	10	/
Total number of items	30	10	30	36	10	30	36
Total score	0-120	0-40	0-120	0- 144	0-40	0-120	0-144
Test-retest reliability: Pearson	r= 0.92	/	/	/	/	r=0.91	r=0.63
Test-retest reliability: Spearman	/	/	r_s=0.89 (p<0.001)	r_s=0.92 (p<0.001)	r_s=0.86 (p<0.001)	/	/
Total cronbach's alpha	r= 0.95	group 1: r=0.97 Control group: r=0.96	/	r=0.97	r=0.94	/	r=0.96

However, Rosen and Murry (2000) showed that the VHI is not sensitive enough for singers with vocal complaints compared to non-singers with vocal complaints. They indicated that singers represent a unique group of individuals with voice problems, of which some present

with hoarseness in their speaking voice, but many present with problems specifically related to their singing voice. Singers with voice problems had significant lower VHI scores compared to the non-singers with voice problems. Rosen and Murry suggested that the low VHI scores might be related to the questions in the VHI focusing on voice problems related to talking and not specifically to singing (only 5 to 30 items were established sensitive for singers' problems). According to Rosen and Murry (2000) another explanation for the low VHI scores is that singers are more sensitive than non-singers to early symptoms of voice problems and thus present earlier in the development of their voice problem. Conversely, non-singers may present at a later time when their speaking voice reaches a communication handicapping level, i.e. discomfort, vocal fatigue or voice loss. Rosen and Murry (2000) warned the clinicians of the seriousness of a relatively low VHI score in singing subjects, because it might represent a significant handicap that should not be ignored when considering the severity of a singer's voice problem.

Following Rosen and Murry (2000), several adaptations have been made to the original VHI (Jacobson et al., 1997) to meet the needs of the singing voice.

In 2003 Schmitt and Balteau attempted the first adaptation for singers of the VHI (Jacobson et al.,1997). They selected 46 items. Five of these items were preserved from the original VHI, and were found sensitive by Rosen and Murry (2000). The other items of the VHI of Jacobson et al. (1997) were adapted to the singing voice. Finally Schmitt and Balteau also added new items, formulated by E.N.T. specialists and speech and language therapists specialized in voice, based on their experience with singer opinions. The relevance of these 46 items was evaluated by an experimental group of professional voice users and singers and a control group of non-singers. They preserved 31 items (10 items functional, 10 items emotional and 11 items physical). Balteau (2003) investigated these items in 53 singers with different singing styles (e.g., classical, gospel, karaoke, Arabic classical, rock, jazz,…). All these singers had a voice disorder . The scale was found reliable, but the number of preserved items was restricted to 18 (3 items functional, 8 items emotional and 7 items physical). As a result of the heterogeneity of the population the reliability of the scale probably was diminished. Morsomme et al. (2005) modified the items which were irrelevant according to the study of Balteau et al. (2003) and reconstructed the scale consisting now of 28 items. This scale was validated by a population of 132 classical singers (with and without voice disorders, professionals and amateurs, soloists and chorus members).

Morsomme et al. (2005) obtained excellent test-retest stability and a good internal consistency, but at the end of their study, they suggested to modify 9 items to achieve a better internal consistency of the scale. Lesage et al. (2005) elaborated this suggestion of Morsomme at al. (2005). They reformulated some items to adapt them to the goal population and added two items that were previously eliminated. Finally they obtained a scale of 30 items which they validated in a sample of 52 divers singers with or without voice disorders. Again, excellent test-retest stability and a very good internal consistency was obtained. Lesage et al. (2005) suggested some modifications, for the same reason Morsomme et al. (2005) did.

In 2007 Morsomme et al. were the first authors who adapted the Voice Handicap Index to the singing voice. They developed the self rating scale: 'Voice Handicap Index adapté aux chanteurs', which measures classical singers' perception of voice handicap. This version was the result of five former studies, that obtained statistical validity, reliability and internal consistency. This scale, consisting of 30 items, has been developed for use in the daily clinic. It permits to evaluate the impact of a voice disorder on the quality of life of singers. The reliability of the questionnaire and the internal coherence of the three subscales (functional (F), emotional (E) and physical (P)), obtained in a population of 192 classical singers and 27 non-singers, was good. Morsomme et al. (2007) also examined nine variables: vocal complaint, smoker, soloist or chorus member, age, professional or amateur, gender, singing lessons and classification of voice type. The presence of a vocal complaint was significant, which confirmed the hypothesis that the 'VHI adapté aux chanteurs' is sensible for a vocal complaint of singers and that the items are relevant to measure the impact of the voice disorder on the quality of life. The variables 'amateur' and 'soloist' were increased compared to respectively the variables 'professional performers' and 'chorus members'. The variables 'age' and 'gender' had an influence on the scores of the questionnaire. The scores were slightly higher in the female singers compared to the male singers. Older subjects showed slightly higher scores. Other variables had no influence on the results of the questionnaire. Morsomme et al. (2007) concluded that the 'Voice Handicap Index adapté aux chanteurs' is a reliable and valid instrument that can evaluate voice disorders in the population of classical singers. The questionnaire is sensible for vocal complaints.

The purpose of this study was to develop a reliable and valid version of the 'VHI adapté aux chanteurs' (Morsomme et al., 2007), an instrument to evaluate the impact of voice disorders for the specific population of classical singers, in Dutch: 'VHI aangepast aan de zangstem'.

Methods

Participants

A specific population of classical singers was selected because this is a homogenous population compared to a group of singers with other singing styles. A sample of 90 subjects, all classical singers, were included in this study and divided into two groups: the test group, consisting of singers without voice disorders (n=85) and the control group, consisting of singers with voice disorders (n=5). All participants sang at least one of the following repertoires: opera, oratorio or classical songs.
All subjects were informed about the purpose of this study and participated voluntary. The subjects were assured an anonymous participation.

Test group

Ninety-two singers without voice disorders completed questionnaire. Seven of these singers were excluded, because they had vocal complaints (for example: hoarseness) or health complaints (for example: flu) either on test or retest. This brought us to a test group of 85 singers without voice disorders, with a mean age of 37 years (range: 18 – 66 years, SD: 11 years). The test group consisted of 29 men (34%) and 56 women (66%). Four singers (one man and three women) who did not complete the retest questionnaire were included in the test group and were only excluded from statistical analysis which involved the retests of the subjects.

To be included in the test group, the participants had to be classical singers without voice disorders or vocal complaints, who sang regularly (minimum 2 hours a week). Their native language had to be Dutch and they had to be at least 16 years old (to focus on an adult population with a mature larynx). The last inclusion criterion was that the participants may not follow speech therapy during one month before completing the test and during the interval between the test and retest, because treatment might influence the results of the questionnaire (Cohen et al.,2008).
If, by a person, one of these inclusion criteria on test or retest was not satisfied, this person was excluded.

Control group

Classical singers with voice disorders were included in the control group. These subjects were very hard to find. The main reason is that there was a lot of resistance in participation within these subjects, although an anonymous participation was assured. Sataloff (1997) showed that

openly disclosing a voice disorder seems to remain taboo in the singing world. Another reason was that they might consult more than one E.N.T. specialist or speech-language therapist for advice. Another reason might be that the prevalence of singers with voice disorders could be little. Rosen & Murry, (2000) and Murry et al., (2007), suggested that classical singers are typically very sensitive to minor nuances and variations in their singing voice, identifying voice abnormalities during singing that are not perceived alike during speaking. This might indicate a less severe voice problem. Moreover, classical singers have a bigger rigor in following the rules of voice hygiene (Ormezzano, 2000). This could be indicating that there are less voice disorders in the population of classical singers. Several hospitals, University clinics, E.N.T. specialists and speech therapists have been contacted. Only 6 singers with voice disorders could be contacted. Five of them (1 man and 4 women), with a mean age of 37 years (range: 21 – 50 years, SD: 14 years), completed the test and the retest. They form the control group.

Inclusion criteria were Dutch speaking classical singers with a (self-perceived) voice disorder, who sang regularly (minimum 2 hours a week) and who were at least 16 years old (to focus on an adult population with a mature larynx).

Procedure

Translation

The first goal of this thesis was to translate the 'VHI adapté aux chanteurs', accomplished in French (Morsomme et al., 2007), into Dutch. The patient's history was also translated into Dutch. Two independent translations of the questionnaire, one by a native Dutch speaking speech-language therapist student A.D., and one by a linguist P.S. (member of the department ISLV Editing and Translation Services of the University of Luik, Smith Phyllis) (Appendix 1 and Appendix 2), were made. Thereafter, A.D. compared these two translations and modified them into the final translation (Appendix 3). Modifications who were made were mostly subtle content differences (e.g., item F22: 'Ik heb er moeite mee om mijn emoties al zingend te uiten' was changed into: 'Ik heb er moeite mee om mijn emoties zingend over te brengen'). Item F6 had a double meaning and had been changed into a clear meaning: 'Ik heb het gevoel dat ik weerhouden word van 'projecten' omwille van mijn stem' was changed into: 'Ik heb het gevoel dat ik niet bij 'projecten' word betrokken omwille van mijn stem'. Other differences were for example: differences in style or formation as for example the selection of prepositions. For example item F3: 'Ik vermijd zingen onder begeleiding (piano, orkest, instrumentale groep)' was changed into: 'Ik vermijd zingen met begeleiding (piano, orkest,

instrumentale groep)'. Differences in sentence structure were performed, for example: item E27: 'Het idee te moeten zingen maakt me bang' was changed into 'Ik word bang bij het idee te moeten zingen'.
All 30 items and the subscale structure from the original 'VHI adapté aux chanteurs', were preserved. At least subtle differences were made in order to take the specific jargon of the classical singers into account. For example: item E15: 'Ik hou niet van de kleur van mijn stem (timbre, schelheid, korrel,…)' was changed into 'Ik hou niet van de klankkleur van mijn stem (timbre, schelheid, ruwheid,…)'.

Testing

All participants had to complete the 'VHI aangepast aan de zangstem' and a participant's history.

'VHI aangepast aan de zangstem'

The 'Voice Handicap Index aangepast aan de zangstem' is a self-administered questionnaire that consists of 30 statements. It is a translation of the 'VHI adapté aux chanteurs' of Morsomme et al. (2007) into Dutch, which is adapted to the specific population of classical singers.
The subject has to respond according to the appropriateness of each item (0=never, 1=almost never, 2=sometimes, 3=almost always, 4= always). This questionnaire is constructed with a three subscale structure: a functional scale, a physiological scale and an emotional scale. Each subscale contains 10 items and is worth 40 points. The total score varies between 0 and 120 with the latter representing the maximum perceived disability due to voice disorders or vocal complaints.

The following instructions were given orally (to singers who were tested personally) or written (to singers who were tested by email), at the beginning of the testing of the 'VHI aangepast aan de zangstem':

> *" Hieronder vindt u statements die door de meerderheid van de zangers gebruikt worden om hun (zang)stem en de weerslag daarvan op het dagelijkse leven te beschrijven. Duid aan in welke mate de statements bij u van toepassing waren in de afgelopen maand. « Altijd» komt overeen met «ja, dit is altijd een probleem » ; « nooit » komt overeen met «neen, geen enkel probleem». "*

The subjects had to complete the questionnaire twice, with an interval of 14 days between test and retest ((De Bodt et al., (1997), Hogikyan and Sethuraman, (1999), Wuyts et al., (1999)). To receive and complete the test and the retest, the participants had the choice to receive and complete the questionnaire by email or personally. Personal appointments took place in the private accommodation of the singers themselves in a non-professional context (Aalst, Lede, Nieuwerkerken, Leuven and Gent), during a repetition of the chorus (Brussel) or during singing classes (Leuven, Aalst, Gent).
Singers, who decided to complete the retest by email, received the questionnaire exactly 10 days after the day they completed the test. When they did not return the retest in time, they received a phone call or a new email to remind them. Others, who decided to complete the retest personally, received a new appointment exactly 14 days after the day they completed the test. The mean interval was 14.5 days (range: 10 – 33 days, SD: 4 days) for the test group and 18.6 days (range: 9 – 26 days, SD: 6 days) for the control group.

Questions were answered orally (to singers who were tested in person) and written (to singers who were tested by email).

Participant's history:

The participants history (Appendix 4) had to be completed once, at the first testing. This questionnaire investigated several variables (see further: 'variables'). For example: being a smoker or not, type of singing voice, grade of profession, exc.
The questions were all closed questions with several given options. There was always an option: 'others', to give the singers the opportunity to provide another answer, when none of the possibilities fitted. One exception was made for the variable: 'vocal complaint'. This was an open question in order not to introduce a bias by giving suggestive examples of vocal complaints. If subjects from the test group had a vocal complaint on a test moment (either test or retest), they were excluded (see study group). The subjects from the control group had to define the duration of their voice disorder and whether they were following or did follow speech therapy for this problem or not, because treatment could influence the results on the questionnaire (Cohen et al., 2008).

Statistical analysis

The statistical analysis was performed by a statistical program: “PASW Statistics 18”.

Standardization of the ‘VHI aangepast aan de zangstem’

Reliability of the ‘VHI aangepast aan de zangstem’

To calculate the test-retest reliability of the ‘VHI aangepast aan de zangstem’ three methods of statistical analysis were performed: A Wilcoxon Signed Ranks Test, the Intraclass Correlation Coefficient (ICC) and the Spearman’s Rho correlation coefficient. Comparisons were made for the total VHI score and the three sub-scores (functional (F), emotional (E) and physical scale (P)).

The non-parametric Wilcoxon matched-pairs Signed Ranks Test was used, to compare the mean difference of the scores from the test group on the test and retest. Significance level was set at $\alpha=0.05$.

The Intraclass Correlation Coefficient (ICC) calculates which part of the total variability is due to the subjects themselves and which part is due to a bias. The total variability represents the sum of the variability within the subjects and the variability between the subjects.

The correlation coefficient: Spearman’s Rho was performed in order to calculate the correlations between the test and the retest scores. The correlations between the sub-scores: F-E, F-P and E-P, were also examined. Significance level was set at $\alpha=0.05$.

Internal consistency

The internal consistency of each sub-scale was examined by calculating the Cronbach’s alpha. Improved values of the Cronbach’s alpha were considered by deleting items (one by one) who were less coherent with other items.

Validation of the ‘VHI aangepast aan de zangstem’

The results between singers without (test group) and with (control group) vocal complaints were compared, using the non-parametric Wilcoxon Signed Ranks Test. Significance level was set at $\alpha=0.05$.

At First, the mean differences between the scores of the test group and those of the control group were compared. Subsequently, the mean differences of the test and retest scores of both groups were compared.

Variables

The effect of several variables on the total score and sub-scores was calculated.

The non-parametric Mann-Whitney U-Test was used for the variables: ‘gender’, ‘age’, ‘smoker’, ‘singing style’ and ‘following or having followed singing classes’. Comparisons were made for the total VHI score and the three sub-scores (functional, emotional and physical scale). Significance level was set at $\alpha=0.05$.
For the remaining variables: ‘grade of profession’, ‘type of profession’ and ‘classification of the singing voice’, a One-Way Analysis of Variance (ANOVA) was used. The condition of equality of variances, required for ANOVA, was set at $\alpha>0.01$. Therefore, the homogeneity test of variance (Levene’s tests) was performed.

Results

The results of the scores on the questionnaire (test and retest) of the test group and control group are shown in **Table 2.**

Table 2: Results of the scores on test and retest of the test group and control group

		Group																
		Test group									Control group							
		N						Percentiles			n						Percentiles	
		valid	missing	mean	SD	Min	Max	25	50	75	Valid	Missing	Mean	SD	Min	Max	25	50	75
Total	**test**	85	0	12.0	8.94	0	31	3	12	20	5	0	34.0	12.04	21	53	25	30	45
	retest	81	4	9.6	9.02	0	46	3	7	16	5	0	36.2	11.45	27	55	28	31	47
F-scale	**test**	85	0	3.6	3.06	0	13	1	3	5	5	0	10.8	4.02	6	15	6	13	14
	retest	81	4	3.0	2.82	0	11	0	2	5	5	0	10.8	3.90	5	15	7	12	14
E-scale	**test**	85	0	3.8	3.88	0	15	0	2	6	5	0	13.2	7.33	8	26	8	11	19
	retest	81	4	3.0	3.54	0	15	0	2	4	5	0	14.2	7.95	8	28	9	12	20
P-scale	**test**	85	0	4.5	3.63	0	13	1	4	8	5	0	10.0	3.39	5	13	6	12	12
	retest	81	4	3.7	4.13	0	29	1	3	5	5	0	11.2	3.11	8	15	8	12	14

Part 1: Standardization of the 'VHI aangepast aan de zangstem'

Reliability of the 'VHI aangepast aan de zangstem'

To calculate the test-retest reliability three methods of statistical analysis were performed on the test group: the Wilcoxon Signed Ranks Test, the Intraclass Correlation Coefficient (ICC) and the Spearman's Rho coefficient.

Wilcoxon Signed Ranks Test

A comparison between the scores (total scores and subscores: functional (F), emotional (E) and physical (P)) on test and retest from the test group (N=81) was made. Therefore, we calculated the mean differences of the test scores and the retest scores. To prove whether the mean differences were significant or not, the non-parametric Wilcoxon matched-pairs Signed Ranks Test was performed. The results are shown in **Table 3**.

Table 3: Mean differences between test and retest of the total scores and the subscores

	total score	**scores on the F-scale**	**scores on the E-scale**	**scores on the P- scale**
Mean difference between test and retest	12	8	6	9
Z-value	-3.185	-2.575	-2.524	-2.667
p-value	0.001	0.010	0.012	0.008

Table 3 shows that the mean differences of the total score and subscores (F, E and P) between test and retest are significant (resp. p=0.001 ; p=0.010 ; p=0.012 ; p= 0.008). This means that all scores of the test are significantly higher compared to the retest.

Intraclass Correlation Coefficient (ICC)

The Intraclass Correlation Coefficient calculates which part of the total variance (variance between the subjects and the variance between test and retest), from the scores of the test group (N=81), is attributed to the real score (the variance of the subjects themselves) and which part is attributed to a bias. The results of the ICC value of total scores and subscores are shown in **table 4**.

Table 4: ICC value of the total score and subscores

	ICC	95% Confidence Interval		p-value
		Lower Bound	Upper Bound	
Total scores	0.744	0.628	0.827	<0.001
F-scale	0.738	0.621	0.823	<0.001
E-scale	0.709	0.581	0.802	<0.001
P-scale	0.651	0.506	0.761	<0.001

As shown in **Table 4** the ICC values vary from 0.651 to 0.744, which is good (Fleiss, 1986). All values were strongly significant (p<0.001).

Spearman's Rho correlation coefficient (r_s)

The test-retest reliability was also investigated by calculating the Spearman's Rho correlation coefficient (r_s) of the test and the retest scores of the total score and the subscores from the test group (N=81). **Table 5** shows the results.

Table 5: Spearman's Rho correlation coefficient (r_s) between test and retest scores

	Spearman's Rho (r_s)	p-value
Total score	0.794	<0.001
F-scale	0.773	<0.001
E-scale	0.782	<0.001
P-scale	0.747	<0.001

Table 5 shows that the test and retest scores are positively correlated (range: 0.747 – 0.794), which is further illustrated in the scatterplots (**Figure 1 – 4**). All correlations are significant ($p<0.001$).

Figure 1: Correlation of the total scores

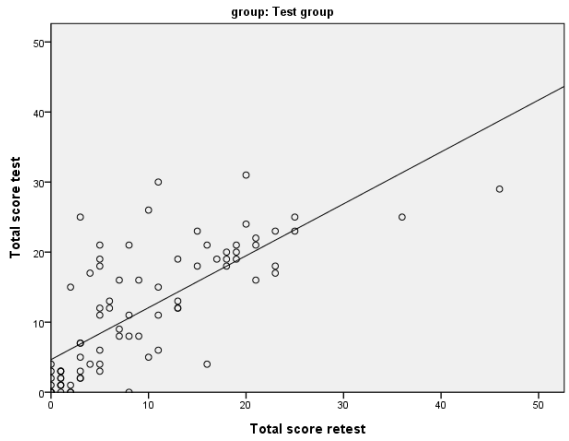


Figure 2: Correlation of the F-scores

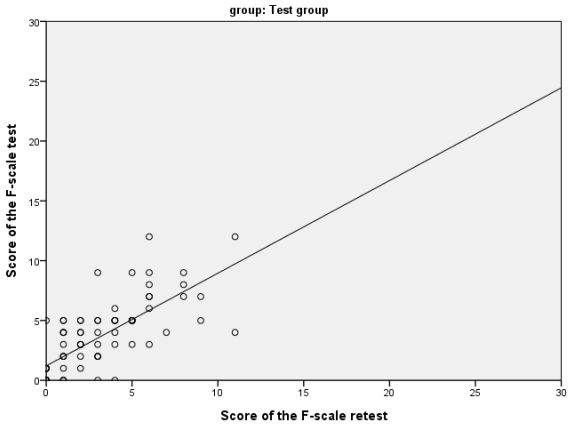


Figure 3: Correlation of the E-scores

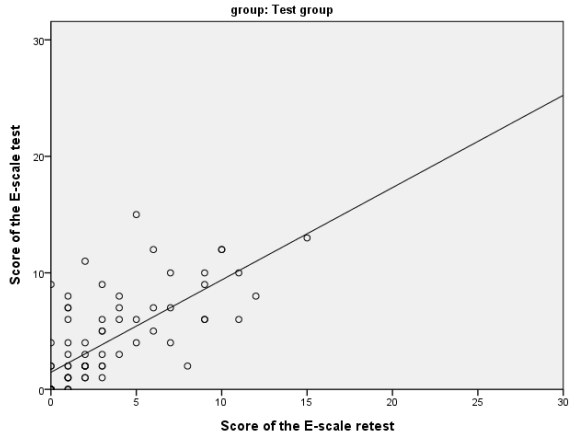
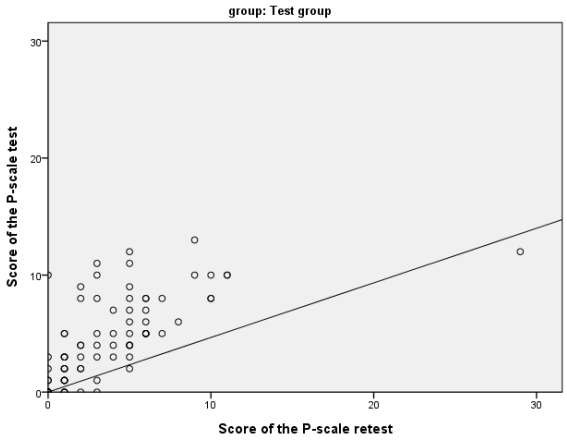


Figure 4: Correlation of the P-scores



We also compared the different subscales (F-E, F-P and E-P) of the questionnaire, on the test scores and the retest scores, by performing a Spearman's Rho. **Table 6** shows the results.

Table 6: Spearman's Rho correlation coefficients between the three subscales (F-E, F-P and E-P) of the 'VHI aangepast aan de zangstem'

		Test	Retest
F-E	Spearman's Rho (r_s)	0.597	0.660
	p-value	p<0.001	p<0.001
F-P	Spearman's Rho (r_s)	0.710	0.729
	p-value	p<0.001	p<0.001
E-P	Spearman's Rho (r_s)	0.684	0.677
	p-value	p<0.001	p<0.001

Table 6 shows significant positive correlations between the three subscales (p<0.001). The strongest correlation was showed between the F-scale and P-scale (r_s= 0.710).

The internal consistency of each subscale

The internal consistency, or the homogeneity, of the questionnaire and of the items included in each subscale was determined using the Cronbach's alpha. Calculation was based on the scores of the test from the test group (N=85). The Cronbach's alpha for the total scores of the questionnaire was 0.875. **Tables 7** to **9** show the improved values of the Cronbach's alpha of the different subscales, considered by deleting items (one by one) who were less coherent with other items.

Table 7: Cronbach's alpha of the functional scale

Functional scale	New Cronbach's Alpha if item deleted			
Item F1	0.639	0.652	0.667	0.681
Item F3	**0.698**	/	/	/
Item F5	0.648	0.664	0.680	0.697
Item F6	0.695	0.708	**0.724**	/
Item F8	0.696	**0.709**	/	/
Item F11	0.618	0.633	0.644	0.673
Item F12	0.647	0.658	0.675	0.698
Item F16	0.674	0.688	0.699	0.721
Item F19	0.652	0.665	0.677	0.692
Item F22	0.645	0.656	0.669	0.685
Total Chronbach's alpha	**0.687**	**0.698**	**0.709**	**0.724**

Table 7 shows the improved values of the Cronbach's alpha ranging from r=0.687 to r=0.724, if the following items consecutively were deleted:

Item F3: 'Ik vermijd zingen met begeleiding (piano, orkest, instrumentale groep)'
Item F8: 'Mijn stemprobleem veroorzaakt een inkomensverlies'
Item F6: 'Ik heb het gevoel dat ik niet bij <projecten> word betrokken, omwille van mijn stem'

Table 8: Cronbach's alpha of the emotional scale

Emotional scale	Cronbach's Alpha if item deleted	
Item E7	0.738	0.743
Item E9	0.743	0.753
Item E15	0.766	/
Item E23	0.749	0.759
Item E24	0.755	0.761
Item E25	0.722	0.724
Item E27	0.748	0.758
Item E28	0.722	0.723
Item E29	0.748	0.750
Item E30	0.719	0.720
Total Chronbach's alpha	**0.761**	**0.766**

Table 8 displays the improved values of the Cronbach's alpha ranging from r=0.761 to r=0.766, if item E15: 'Ik hou niet van de klankkleur van mijn stem (timbre, schelheid, ruwheid,…) was removed.

Table 9: Cronbach's alpha of the physical scale

Physical scale	Cronbach's Alpha if item deleted			
Item P2	0.743	0.771	0.770	0.776
Item P4	**0.798**	/	/	/
Item P10	0.762	0.790	0.791	0.794
Item P13	0.772	**0.799**	/	/
Item P14	0.763	0.793	**0.802**	/
Item P17	0.727	0.760	0.763	0.766
Item P18	0.734	0.760	0.758	0.758
Item P20	0.741	0.767	0.768	0.771
Item P21	0.759	0.781	0.778	0.784
Item P26	0.755	0.779	0.779	0.784
Total Chronbach's alpha	**0.775**	**0.798**	**0.799**	**0.802**

Table 9 shows the improved values of the Cronbach's alpha from r=0.775 to r=0.802, if the following items consecutively were deleted:

Item P4: 'Ik heb moeilijkheden om mijn zinnen af te maken'

Item P13: 'Mijn spreekstem is slechter nadat ik gezongen heb'

Item P14: 'Ik voel irritatie of pijn in mijn larynx (strottenhoofd) wanneer ik zing'

Part 2:

Comparisons of the results between singers with (control group) and without (test group) vocal complaints

Validation of the 'VHI aangepast aan de zangstem'

To validate the 'VHI aangepast aan de zangstem' comparisons of the results (total scores and subscores) between singers with (control group) and without (test group) vocal complaints were made (see **Table 10**). A Mann-Whitney U-Test was performed. The 0-hypothesis was that there would be no difference between both groups.

Table 10: Comparisons of the mean difference between the scores of the test group and the control group

	Difference for total score		Difference for F-scale		Difference for E-scale		Difference for P-scale	
	Test	Retest	Test	Retest	Test	Retest	Test	Retest
Mean difference of the test group – control group	40.66	41.51	39.49	38.33	38.01	39.18	32.93	38.75
z-value	-3.386	-3.614	-3.318	-3.366	-3.200	-3.461	-2.754	-3.395
p-value	0.001	<0.001	0.001	0.001	0.001	0.001	0.006	0.001

The results in **table 10** show that the 0-hypothesis may be rejected because there were significant differences between all scores of both groups, where singers with voice disorders (control group) had systematically higher scores than singers without voice disorders (test group). All differences between both groups were significant (range: $p<0.001$ – $p=0.006$).

The Mann-Whitney U-Test was also performed to compare the mean difference between test and retest scores of both groups. The 0-hypothesis was that there would be no difference between both groups. **Table 11** shows the results.

Table 11: Comparisons of the mean difference between test and retest scores of the test group and control group.

	Difference total scores Test_Retest	Difference F-Score: Test_Retest	Difference E-Score: Test_Retest	Difference P-Score: Test_Retest
Mean difference of the test group-control group	27.60	6.90	25.48	23.46
z-value	-2.413	-0.615	-2.287	-2.069
p-value	0.016	0.538	0.022	0.039

The results show that the mean differences between test and retest scores are higher for the singers without voice disorders (test group) than for the singers without voice disorders (control group). The differences of the total scores and the physical and emotional scores are significant (range: p=0.016 to p=0.039). The difference of the functional scores is not significant (p=0.538) (**table 11)**.

Part 3: Effects of the different variables on the total score and subscores of the 'VHI aangepast aan de zangstem'

Variables

The influence of different variables on the total score and subscores of the test group (n=85) were investigated. The nonparametric Mann-Whitney U-Test for the nonrelated two samples was performed to calculate the effects of the variables 'gender', 'age', 'smoking', 'singing style' and 'following or having followed singing classes'. The results are shown in **table 12**.

Table 12: Mann-Whithey U-Test for the variables: 'gender', 'age', 'smoking', 'singing style' and 'following or having followed singing classes'

Variable	Subgroups		Total score	F-scale	E-scale	P-scale
Gender	man	n (%)	29 (34%)	29 (34%)	29 (34%)	29 (34%)
		Mean	41.3	44.9	44.5	36.6
	woman	n (%)	56 (66%)	56 (66%)	56 (66%)	56 (66%)
		Mean	43.6	42	42.2	46.3
		p-value	0.766	0.599	0.679	0.082
		z-value	-0.297	-0.525	-0.414	-1.739
Age	Younger than 50 years	n (%)	72 (85%)	72 (85%)	72 (85%)	72 (85%)
		Mean	44.5	43.3	44.5	44.8
	50 years or older	n (%)	13 (15%)	13 (15%)	13 (15%)	13 (15%)
		Mean	34.9	41.4	34.9	32.9
		p-value	0.199	0.800	0.191	0.106
		z-value	-1.284	-0.253	-1.306	-1.615
Smoking	non-smoker	n (%)	83 (98%)	83 (98%)	83 (98%)	83 (98%)
		Mean	12.0	3.6	3.8	4.6
	smoker	n (%)	2 (2%)	2 (2%)	2 (2%)	2 (2%)
		Mean	12.0	4.0	4.0	4.0
		p-value	0.954	0.681	0.691	0.988
		z-value	-0.058	-0.411	-0.397	-0.015
Singing style	Classical	n (%)	77 (91%)	77 (91%)	77 (91%)	77 (91%)
		Mean	12.0	3.7	3.6	4.7
	Classical + others	n (%)	8 (9%)	8 (9%)	8 (9%)	8 (9%)
		Mean	11.6	2.8	5.4	3.5
		p-value	0.886	0.484	0.303	0.472
		z-value	-0.143	-0.701	-1.030	-0.719
Following or having followed singing classes	No singing classes	n (%)	1 (1%)	1 (1%)	1 (1%)	1 (1%)
		Mean	15.0	7.0	1.0	7.0
	Singing classes	n (%)	84 (99%)	84 (99%)	84 (99%)	84 (99%)
		Mean	11.9	3.6	3.8	4.5
		p-value	0.791	0.194	0.549	0.436
		z-value	-0.265	-1.299	-0.599	-0.779

None of the variables, showed in **table 12**, were significant.

Before we could use the one-way analysis of variance (ANOVA), for calculating the effects of the variables 'grade of profession', 'type of profession' and 'classification of the singing voice', we verified whether the condition of equality of variances, performed by the homogeneity test of variance (Levene's test), was satisfied (see **table 13**). The three variables were divided into several subgroups. The results of the effects of the three variables are displayed in respectively **Table 13** to **16**.

Table 13: Levene's test of the variables: 'grade of profession', 'type of profession' and 'classification of the singing voice'

Variables		Total score	F	E	P
Grade of profession	**p-value**	0.052	0.469	0.103	0.075
Type of profession	**p-value**	0.444	0.502	0.426	0.396
Classification of the singing voice	**p-value**	0.400	0.308	0.111	0.156

The conditions of equality of variances, required for ANOVA, is satisfied ($p>0.01$) for all variables showed in **Table 13**.

Table 14 shows the results of the effect of the variable 'grade of profession'.

Table 14: One-way ANOVA of the variable: 'Grade of profession'

Grade of profession	n (%)	Mean	p-value
Total score			0.095
professional	54 (64%)	11.1	
semi-professional	15 (18%)	9.9	
amateur	8 (9%)	15.3	
others	8 (9%)	18.3	
F score			0.039
professional	54 (64%)	3.1	
semi-professional	15 (18%)	3.5	
amateur	8 (9%)	6.4	
others	8 (9%)	4.3	
E score			0.194
professional	54 (64%)	3.7	
semi-professional	15 (18%)	3.2	
amateur	8 (9%)	2.9	
others	8 (9%)	6.5	
P score			0.027
professional	54 (64%)	4.3	
semi-professional	15 (18%)	3.2	
amateur	88 (9%)	6.0	
others	8 (9%)	7.5	

The results show that the grade of profession variable has a significant influence on the scores of the F-scale and the P-scale. The subgroup amateur provides the highest scores on these scales. No significant influences were shown on the total scores or the scores on the E-scale. (**Table 14)**

Table 15 displays the results of the effect of the variable 'type of profession'.

Table 15: One-way ANOVA of the variable: 'type of profession'

Type of profession	n (%)	mean	p-value
Total score			0.537
Soloist	43	10.6	
Chorus member	9	11.4	
Soloist and chorus member	24	13.9	
Soloist and others	8	13.1	
Others	1	20.0	
F score			0.313
Soloist	43	2.9	
Chorus member	9	4.6	
Soloist and chorus member	24	4.3	
Soloist and others	8	4.0	
Others	1	5.0	
E score			0.383
Soloist	43	3.8	
Chorus member	9	2.3	
Soloist and chorus member	24	4.3	
Soloist and others	8	3.5	
Others	1	10.0	
P score			0.521
Soloist	43	3.9	
Chorus member	9	4.6	
Soloist and chorus member	24	5.3	
Soloist and others	8	5.6	
Others	1	5.0	

No significant results were demonstrated for the influence of the variable 'type of profession'.(**Table 15**)

The results of the last variable ‘classification of the singing voice’ that was investigated by performing the one-way ANOVA are shown in **table 16**.

Table 16: One-way ANOVA for the variable: ‘Classification of the singing voice’

Classification of the singing voice	n (%)	mean	p-value
Total score			0.946
Soprano	36 (42%)	12.6	
Mezzo-soprano	12 (14%)	10.5	
Alto	5 (6%)	15.4	
Tenor	15 (18%)	11.7	
Baritone	9 (11%)	11.6	
Bass	1 (1%)	11.0	
Others	7 (8%)	9.7	
F score			0.068
Soprano	36 (42%)	3.4	
Mezzo-soprano	12 (14%)	2.8	
Alto	5 (6%)	7.6	
Tenor	15 (18%)	4.1	
Baritone	9 (11%)	3.4	
Bass	1 (1%)	5	
others	7 (8%)	2.3	
E score			0.966
Soprano	36 (42%)	3.9	
Mezzo-soprano	12 (14%)	3.6	
Alto	5 (6%)	2.4	
Tenor	15 (18%)	4.2	
Bariton	9 (11%)	4.4	
Bass	1 (1%)	2.0	
others	7 (8%)	3.3	
P score			0.594
Soprano	36 (42%)	5.4	
Mezzo-soprano	12 (14%)	4.1	
Alto	5 (6%)	5.4	
Tenor	15 (18%)	3.3	
Bariton	9 (11%)	3.7	
Bass	1 (1%)	4.0	
others	7 (8%)	4.1	

No significant results were obtained for the ‘classification of the singing voice’ variable . as shown in **table 16**.

Discussion

Purpose

The purpose of this study was to develop a reliable and valid version of the 'VHI adapté aux chanteurs' (Morsomme et al., 2007), an instrument to evaluate the impact of voice disorders for the specific population of classical singers, in Dutch: 'VHI aangepast aan de zangstem'.

Translation

The first objective of this thesis was to translate the 'VHI adapté aux chanteurs', accomplished in French by Morsomme et al. (2007), into Dutch. The patient's history was also translated into Dutch. Two independent translations of the questionnaire, one by the author (student speech-language therapy A.D.) and one by a linguist P.S. (member of the department ISLV Editing and Translation Services of the University of Luik) (Appendix 1 and Appendix 2), were compared and modified into the final translation (Appendix 3).

Participants

The participants in this study were selected from the population of Dutch speaking, classical singers. Many arguments confirm that this population of classical singers is a homogenous group in contrary to a population of singers with other singing styles. The first one is that classical singers are more sensitive for a good functioning of their voice instrument and that the classical singing style enforces a big demand to the vocal performance. (Le Huche, 2001 ; Simon, 2004). The voice of a classical singer may not suffer any imperfection, contrary to other singing styles, in which breathiness or roughness might be an asset. Another reason is that classical singers often follow singing classes, which provides them to have general knowledge of their voice instrument (basic anatomy, specific jargon,…). Because of their education classical singers form a more homogenous group in contrast to the group of singers with different singing styles, which have very divers educations or no education at all (Simon, 2004). The final reasons are that classical singers are in general more assiduous in training and warm-up (Osta in Klein-Dallant, 2001) and they have a bigger rigor in following the rules of voice hygiene (Ormezzano, 2000). Moreover, different singing styles distinguish from classical singing at different physiological points: the singing technique and the use of registers might be different in different singing styles, which modify the acoustic results (Fournier, 1999 in Lesage, 2005). Also, the tessitura, the articulation and the posture of the singers are different (Ormezzano, 2000). In addition to all these arguments, Cohen et al.,

(2008) showed that singers who sing gospel had worse scores on the SVHI than singers with a different singing style (classical, rock, country, choral and pop).
Morsomme et al. (2007) also validated the questionnaire 'VHI adapté aux chanteurs' in the homogenous population of classical singers. In the first study of five, preliminary to the final version of the French 'VHI adapté aux chanteurs' by Morsomme et al. (2007), Balteau (2003) investigated his adaptation of the VHI by testing a heterogeneous group of singers with very different singing styles (classical, gospel, karaoke, rock, jazz, …). This heterogeneity of the population probably weakened the results of the internal consistency.

Taking these arguments into account, we opted for a homogenous population of classical singers. We could say that with a homogenous group our internal consistency will not be weakened and we could predict the sensibility of our test more precisely. On the other hand we cannot generalize our findings to the whole population of singers. The test will only be standardized for classical singers. This means that in the clinical practice we can only use the 'VHI aangepast aan de zangstem' on classical singers.

To standardize the 'VHI aangepast aan de zangstem' and to test the specificity of it, a sample of the population of classical singers without voice disorders was used. The test group consisted of 85 classical singers. Singers with subjective voice complaints were assigned to a different group. However, no objective measurements or auditory perceptual judgements of a speech and language therapist were taken. It would have been better to have objective measurements or auditory perceptual judgements, but on the other hand we found in the literature that classical singers are typically very sensitive to minor nuances and variations in their singing voice, identifying voice abnormalities during singing that are not perceived alike during speaking (Rosen & Murry, 2000; Murry et al., 2007). We also assumed that the singers would be honest to admit whether they had a voice problem or not. This was possible because the participants of this study were assured a complete anonymous process. To avoid false negatives in the test group, all singers who had any health problem (for example: a cough, fatigue, flu, …) during either test or retest were excluded. In that way a bias was prevented as much as possible.

To investigate whether the test is sensitive, a control group consisting of 5 classical singers with voice disorders was used. The amount of 5 subjects is disproportional to establish the findings of this study, in which the control group is involved. These findings are the results of the comparison of the mean difference between the scores of the test group and those of the

control group and the results of the comparison of the mean difference between test and retest scores of both groups. For several reasons we could not obtain more subjects in this group. The main reason is that there was lots of resistance in participation within these subjects, although an anonymous participation was assured. Sataloff (1997) showed that openly disclosing a voice disorder seems to remain taboo in the singing world. Other reasons could be that singers with voice disorders are hard to find. We assume that they consult more than one E.N.T. specialist or speech-language therapists for advice. Another reason might be that the prevalence of singers with severe voice disorders could be little. Rosen & Murry, (2000) and Murry et al., (2007), suggested that classical singers are typically very sensitive to minor nuances and variations in their singing voice, identifying voice abnormalities during singing that are not perceived alike during speaking. Moreover, classical singers have a bigger rigor in following the rules of voice hygiene (Ormezzano, 2000). This could be indicating that there are less voice problems in the population of classical singers.

Collection of the questionnaires was performed by email or by a personal appointment. Although we preferred to collect the questionnaires by a personal appointment, for practical reasons (e.g., contacting singers from the Netherlands) this was not always possible. Of the 311 singers that were contacted, only 97 returned the questionnaire. This means that we had a response rate of only 31%.

Standardization

To ensure the reproducibility of the questionnaire, an investigation of the test-retest reliability by calculating the Wilcoxon Signed Ranks Test, the Intraclass Correlation Coefficient (ICC) and the Spearman's Rho (r_s), on the results of the singers without voice disorders (test group) was made.

The scores of test and retest of the total score and subscores from the singers without voice disorders (test group) were compared by using a Wilcoxon matched-pairs Signed Ranks Test. The total score and the scores of the subscales (F-scale, E-scale and P-scale) were lower at the retest compared to the test. The mean differences were respectively 12 ; 8 ; 6 ; 9 and were all significant (resp. p=0.001 ; p=0.010 ; p= 0.012 ; p=0.008). Lesage (2005) also observed lower results at the retest for the total score and for the score of the physical subscale. Morsomme et al., (2005) as well as Morsomme et al., (2007) observed lower results at the retest of the total score and all subscores (F-scale, E-scale and P-scale). The latter suggested that the subjects pay more attention to the items and they search the best possibility that describes the problems

they sense, during the first exposure to the questionnaire. On the contrary, when the participants fill in the retest they are already confronted with the items and the effect of discovery is less. Moreover, Morsomme et al., (2007) suggested that the participants had the time to estimate their voice problems and handicap less grave at the retest. In agreement with Morsomme et al., (2007), we could say that singers might estimate their voice problems and handicap more in perspective at the retest, which could provide lower retest scores. We also propose not to complete the 'VHI aangepast aan de zangstem' with a relatively short succession, for example within the time interval of 14 days, in order not to diminish the results.

The Intraclass Correlation Coefficient (ICC) of the total score is 0.744. This means that the items loaded on a single measurement explains 74.4% of the total variance and the remaining 25.6% is attributed to bias. The ICC value of the subscales ranges from 0.651 to 0.738. The ICC values we obtained were all good (Fleiss, 1986) and significant ($p<0.001$). We can conclude that the 'VHI aangepast aan de zangstem' is a reliable instrument.

The Spearman's Rho correlation coefficient showed that the total scores and the functional, emotional and physical sub-scores between test and retest, correlated with respectively $r_s = 0.794$; $r_s = 0.773$; $r_s = 0.782$; $r_s = 0.747$. All these correlations were high significant ($p<0.001$). The Spearman's Rho correlation coefficient lies in the range of [-1, +1]. This means that the correlation of the scores of the test and the retest, are positively correlated and that the test-retest reliability of the questionnaire is very good.

A comparison between the different subscales (F-E, F-P and E-P) of the test and the retest was also performed by a Spearman's Rho. The results of the Spearman's Rho correlation coefficients ranged from $r_s=0.597$ to $r_s=0.729$ and had a high significance of $p<0001$. These results show that there is a positive correlation between the different subscores. The correlations are not extremely positive, which means that they are related to each other, but not redundant. We conclude that it is useful to have the different subscales in the questionnaire. These results are in agreement with the results Morsomme et al. (2007) showed, where a range between $r_s=0.613$ and $r_s=0.810$ was obtained.

The internal consistency, or the homogeneity, of the items included in each subscale was determined using the Cronbach's alpha. In the literature there is no consensus about the minimum value of the Cronbach's alpha coefficient. Nunnally (1978) has suggested that the Cronbach's alpha coefficient should be at least $\alpha=0.50$ for a single item to demonstrate

acceptable internal consistency. Later Mesbach (2002; in Simon,2004) proposed a Cronbach's alpha of α=0.70. In the original VHI, Jacobson et al. (1997) used a cut-off value of α=0.60 to eliminate items. In the SVHI, Cohen et al. (2007) also used a cut-off value of α= 0.60. As there is no consensus about the cut-off value of the Cronbach's alpha in the literature, the value has only been optimized. In this study, the total scores of the questionnaire showed a Cronbach's alpha of 0.875. The functional scale had the lowest value of α= 0.687. This value could be optimized to α=0.724 by deleting three items (Item F3, item F6 and item F8). The emotional scale could only be optimized from a Cronbach's alpha of α=0.761 to α=0.766 by deleting item E15. The physical scale could be optimized from α=0.775 to α=0.802 by deleting three items (Item P4, item P13 and item P14). The difference between the original values that were obtained and the optimized values ranged from 0.037 for the functional scale, over 0.027 for the physical scale to 0.005 for the emotional scale. Because of the minimal amending after deleting an item and because of the good initial values of the Cronbach's alpha, no item should be removed.

The internal consistency of the 'VHI adapté aux chanteurs' (Morsomme et al., 2007) was slightly better than the results in this study. They obtained a Cronbach's alpha of α>0.80 for each subscale, as in our study none of the subscales showed a Cronbach's alpha of α>0.80. Only the total score showed a Cronbach's alpha of α>0.80 in our study. This could be attributed to subtle differences in the use of the language, where the translation could have caused subtle differences in the value attributed to the words. For example the words 'me déplaît' in French in item E15 was translated into 'ik hou niet van' in Dutch, which reproduces the best suited translation, but still contains a subtle difference.

Validation

To validate the 'VHI aangepast aan de zangstem' comparisons of the results between singers with (control group) and without (test group) vocal complaints were made using the non-parametric Mann-Whitney U-Test. The 0-hypothesis was that there would be no difference between both groups. This hypothesis may be rejected because there were significant differences between all scores of both groups, where singers with voice disorders had systematically higher scores than singers without voice disorders. This means that the test is sensitive and specific. However, we must be careful with this interpretation because the control group consisted of only 5 subjects.

The mean difference between test and retest of the test group and control group was also performed by a Mann-Whitney U-Test. The 0-hypothesis was that there would be no difference between both groups. This 0-hypothesis apply only for the functional subscale where the difference was not significant (p=0.538). The differences between test and retest between both groups were significant for the total scores and the subscores of the emotional and physical scale (resp. p=0.016 ; p= 0.022 ; p= 0.039). In our previous results of the reliability in the test group, significantly lower total scores and subscores on the retest, compared to those on the test, were obtained. This difference is significantly greater than the difference between the test and retest scores of the control group for the total score and the emotional and physical subscales. A possible explanation is that singers with voice disorders might estimate their problem more precisely at the first exposure to the 'VHI aangepast aan de zangstem'. Because they are confronted daily with their problem, they are probably more aware of the impact and value of the different consequences. This might explain why there is less variability between the test scores and retest scores of the singers with voice problems. Another explanation is that the singers with voice disorders could have the intendancy to state their handicap. They are afraid of the diagnosis and the consequences of their problem. So they experience their handicap to be more severe. Three patients in the control group did not see a E.N.T. specialist in the time before or in the time interval of test and retest. The same anxiousness remains or is even increased, which could provide the same severity or even higher severity scores on the retest for this control group. Cohen et al. (2008) showed that in patients who did receive treatment, but whose voice was not improved after the treatment, the SVHI scores did not change. Two of the patients in the control group did see an E.N.T. specialist and knew the diagnosis of their voice disorder. It could be possible that they could estimate their problem more in perspective during the first testing so there is less variability in the answers of the retesting. Again we can agree with Cohen et al. (2008), as these two patients were tested after their treatment, their voice remained the same as well as their scores on the questionnaire. However, it is remarkable that the lower difference in the control group, of only the functional scale is not significant. Again, we must be careful with the interpretation of the results because of the small control group.

Variables

The effects of different variables on the total score and subscores of the test group (n=85) were investigated. The nonparametric Mann-Whitney U-Test for the nonrelated two samples of the variables 'gender', 'age', 'smoking', 'singing style' and 'following or having followed

singing classes' was performed. None of these variables were significant. The gender variable was represented by 34% men and 66% women. In the literature Simon (2004), Lesage (2005) and Surel (2006) showed that the gender variable was not significant. In contrary, Morsomme et al. (2007) found significant differences for the gender variable, where women had significantly higher total scores, emotional and physical subscores compared to men. They suggest that the hormonal cycles might play a role, due to an increased vocal fatigue, its associated roughness and limitation in the tessitura and that stronger emotions and sensitivity might also be observed.
Morsomme et al. (2007) also showed significant effects of the age variable on the F-scores in contrary to this study. They attribute this effect to the reality of concurrency rather than presbyphonia. The reason no significant results were obtained for the age variable in this study could be that the subgroups of the age variable was disproportionally (85% of the subjects were younger than 50 years and only 15% of the subjects were 50 years or older).

The results of the variables 'smoking', 'singing style' and 'following or having followed singing classes' were limited by the proportion of subjects. Only 2% of the singers in the test group smoked. Nine percent of the singers had another singing style, next to classical singing (see inclusion criteria). 'Singers who had not followed singing classes' was represented by only 1% of the sample. Our results are conform the results Morsomme et al. (2007) obtained for the variable 'smoking' and 'following or having followed singing classes', which were not significant.

The one-way analysis of variance (ANOVA) was performed for the variables 'grade of profession', 'type of profession' and 'classification of the singing voice'. The condition of equality of variances, required for ANOVA, was achieved (Levene's test: $p>0.01$).

The grade of profession variable was divided into professionals (64%), semi-professionals (18%), amateurs (9%) and others (9%) (e.g., students). Significant differences in the functional and physical scale were obtained for this variable. For these subscales the subgroup amateur had the highest scores. The same phenomenon was observed by Morsomme et al. (2007) in the difference between amateurs and professional singers. However, they also obtained significantly higher scores for amateurs on the emotional scale. Cohen et al. (2007) investigated the factors associated with the perception of singing voice handicap by using the SVHI. Amateurs had significant worse SVHI scores than professional singers. Rosen and Murry (2000) showed more voice handicap in recreational singers compared to professional

singers. Because of the economical impact (professional singers were defined as having singing as first source of income), professional singers may be more likely to seek medical attention earlier when problems occur, represented by lower singing voice handicap in contrary to amateur singers who consult later and have more important lesions. Roubeau et al., (2004) compared groups of non-singers, amateur and professional singers, where amateur singers demonstrated an intermediate vocal behavior to non-singers and professional singers. This can explain the significant differences in the grade of profession.

The type of profession variable was divided into 5 subgroups: soloist (51%), chorus member (11%), soloist and chorus member (28%), soloist and others (9%) (e.g., ensemble), others (1%). The type of profession variable was not significant for the total score or the subscores. Morsomme et al. (2007) obtained significant differences between the subgroups 'soloists' and 'chorus members' where soloists had increased scores on the E-scale. Simon (2004) showed the same phenomenon for the functional and physical subscales, establishing that soloist do not tolerate any imperfection of their voice in contrary to chorus members. Therefore, in this study, a Mann-Whitney U-Test was performed for these two subgroups but again the results were not significant. Our findings are in agreement with Lesage (2005) and Surel (2006), who did not obtain significant results for this variable.

No significant differences were obtained in the subgroups soprano (42%), mezzo-soprano (14%), alto (6%), tenor (18%), baritone (11%), bass (1%) and others (8%) (e.g., bass-baritone) of the variable 'classification of the singing voice'. The same observations were made by Morsomme et al. (2007). The fact that the singers had the same singing style and that they use their voice in a similar way, could explain these results.

Conclusion

Based on the results of this study, we can conclude that the 'VHI aangepast aan de zangstem' is a reliable and valid Dutch questionnaire, which evaluates the impact of voice disorders and which can be used in the clinical practice for the population of classical singers. Further research is necessary to validate this questionnaire in populations of singers, with different singing styles (for example rock, jazz, karaoke, gospel, musicals, pop,…), in order to expand the clinical use of the questionnaire to these populations.
Retests should not be performed shortly after a previous testing of the 'VHI aangepast aan de zangstem', as retesting is a variable that significantly lowers the scores (Wilcoxon matched-pairs Signed Ranks Test).
The 'VHI aangepast aan de zangstem' is a reliable questionnaire, as shown by good Intraclass Correlation Coefficients of all scores and good Spearman's Rho correlation coefficients between the total scores and subscores of test and retest and within the subscores.
The test is divided into three subscales: functional, emotional and physical subscales. It is useful to have these subscales as analysis showed that they are positively correlated, but not redundant. The internal consistency or homogeneity of each item in the subscales is good (Cronbach's alpha).
Although we showed significant differences between the singers with and without voice disorders, which means that the 'VHI aangepast aan de zangstem' is a valid instrument, we must be careful with this conclusion, because we had a small control group, consisting of 5 singers with voice disorders. Further research is necessary to confirm these results.

The effect of different variables on the results of the questionnaire was investigated by performing either the Mann-Whitney U-Test or the ANOVA. Most variables were not significant: 'gender', 'age', 'smoking', 'singing style' and 'following or having followed singing classes', 'type of profession' and 'classification of the singing voice'. Significant results were obtained for two variables: the retesting variable and the grade of profession variable. The first had significant influences on all scores of the questionnaire. The latter showed significant effects on only the functional and physical scales, the influences on the total score and the emotional scale were not significant. We must notice that the population was not always proportionally divided into the different subgroups of the variables, which could have provoked a bias. Further research is necessary.

Other interesting subjects for further research, are creating a ‘VHI-10 aangepast aan de zangstem’, analogue to the SVHI-10 (Cohen et al., 2009) and investigating the effects of treatment on the scores of the ‘VHI aangepast aan de zangstem’.

Acknowledgment

At the end of this thesis I would like to thank the people who have contributed to the accomplishment of it.

In the first place I would like to thank my promoter Prof. dr. Marc De Bodt, for giving me the professional advice that I needed to accomplish this thesis. Thank you for the trust you gave me during the whole process. It was very pleasant to work with you.

Prof. dr. Dominique Morsomme, as copromoter you were very enthusiastic about this research. You were always prepared to answer my questions and to give your professional advice. Thank you for the nice cooperation and the trust.

I would also like to thank all the subjects who cooperated in this research. Without them this research was not possible.

References

Balteau, A. (2003). *Essai d'adaptation d'une échelle d'auto-évaluation de la voix parlée (le "Voice Handicap Index" de Jacobson, 1997) à la voix chantée: validation d'une échelle sur une population de 53 chanteurs dysphoniques.* Mémoire de licence en logopédie non publié. U.C.L., Louvain-la-Neuve.

Baken, R.J. (1987). *Clinical Measurement of Speech and Voice.* London: Taylor & Francis Ltd.

Benninger, M.S., Gardner, G.M., Jacobson, B.H., Grywalski, C. (1997). New dimensions in measuring voice treatment outcomes. 789-794 In: Sataloff RT, *Professional Voice: The Science and Art of Clinical Care*. 2nd ed. San Diego, Californie: Singular Publishing Group.

Cohen, S.M., Jacobson B.H., Garrett, C.G., Noordzij, J.P., Stewart, M.G., Attia, A., Ossoff, R.H., Cleveland, T.F. (2007). Creation and Validation of the Singing Voice Handicap Index. *Annals of Otology, Rhinology and Laryngology. 166(6),* 402-406.

Cohen, S.M., Noordzij, J.P., Garrett, C.G., Ossoff, R.H. (2007). Factors associated with perception of singing voice handicap. *Otolaryngology Head and Neck Surgery. 138,* 430-434.

Cohen, S.M., Witsell, D.L., Scearce, L., Vess, G., Banka, C. (2008) Treatment Responsiveness of the Singing Voice Handicap Index. *Laryngoscope, 118,* 1705-1708.

Cohen, S.M., Statham, M., Rosen, C.A., Zullo, T. (2009). Development and Validation of the Singing Voice Handicap-10. *Laryngoscope, 119,* 1864-1969.

De Bodt, M., Heylen, L., Mertens, F., Vanderwegen, J., Van de Heyning, P. (2009). *Stemstoornissen.* Antwerpen- Apeldoorn: Garant.

De Bodt, M.S., Wuyts, F.L., Van de Heyning, P.H., Croux, C. (1997). Test-Retest Study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice, 11,* 74-80.

Dejonckere, P.H., Obbens, C., De Moor, G.M. & Wieneke, G.H. (1993). Perceptual Evaluation of Dysphonia: Reliability and Relevance. *Folia Phoniatr, 45,* 76-83.

Fleiss, J.L. (1986). Analysis of data from multiclinic trials. *Controlled Clinical Trials, 7(4),* 267-275.

Franic, D.M., Bramlett, R.E., Bothe, A.C. (2005). Psychometric Evaluation of Disease Specific Quality of Life Instruments in Voice Disorders. *Journal of Voice, 19(2),* 300-315.

Garcia-López, I., Núñez-Batalla, F., Bouzas, J.G., Górriz-Gil, C. (2010). Validación de la versión en español del índice de incapacidad vocal (S-VHI) para el canto. *Acta Otorrinolaringol Esp.61(4),*247-254.

Gliklich, R.E., Glovsky, R.M., Montgomery, W.W. (1999). Validation of a voice outcome survey for unilateral vocal cord paralysis. *Journal of Voice, 15,* 576-586.

Hartnick, C.J. (2002). Validation of a pediatric voice quality-of-life instrument: the pediatric voice outcome survey. *Otolaryngology Head and Neck Surgery, 128(8),* 919-922.

Hirano, M. (1981). *Clinical Examination of Voice.* New York: Springer Verlag.

Hogikyan, N.D., Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice, 13(4),* 557-569.

Hsiung, M.W., Pai, L., Wang, H.W. (2002). Correlation between Voice Handicap Index and Voice Laboratory Measurement in dysphonic patients. *European archives Oto-rhino-laryngology, 259,* 97-99.

Isshiki, N. & Takeuchi, Y. (1970). Factor Analysis of Hoarseness. *Studia Phonologica, 5,* 37-44.

Jacobson, B.H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M.S., Newman, C.W. (1997). The Voice Handicap Index (VHI): Development and Validation. *Journal of Speech-Language Pathology, 6,* 66-69.

Kearns, K.P. & Simmons, N.N. (1988). Interobserver reliability and perceptual ratings: more than meets the ear. *Journal of Speech and Hearing Research, 31,* 131-136.

Lamarche, A., Westerlund, J., Verduyckt, I., Ternström, S. (2010). The Swedish version of the Voice Handicap Index adapted for singers. *Logopedics Phoniatrics Vocology, 35,* 129-137.

Le huche, F., Allali, A. (2001). La voix. *Tome 1 : Anatomie et physiologie de la voix et de la parole.* Masson: Paris.

Lesage, A. (2005). *Essai d'adaptation d'une échelle d'auto-évaluation de la voix parlée (le "Voice Handicap Index" de Jacobson et al.,(1997)) à la voix chantée: validation d'une échelle sur une population de 15 chanteurs de variété dysphoniques, 37 chanteurs de variété normophoniques et 20 sujets non chanteurs sans plainte vocale.* Mémoire de licence en logopédie non publié. U.C.L., Louvain-La-Neuve.

Ma, E.P., & Yiu, E.M. (2001). Voice Activity and Participation Profile: Assessing the Impact of Voice Disorders on Daily Activities. *Journal of Speech, Language and Hearing Research, 44(3),* 511- 524.

Morsomme, D., Gaspar, M., Jamart, J., Remacle, M., Verduyckt, I. (2007). Adaptation du Voice Handicap Index à la voix chantée. *Revue de Laryngologie – Otologie - Rhinologie, 128(5),* 305-314.

Morsomme, D., Simon, Ch., Jamart, J., Remacle, M., Verduyckt, I. (2005). Proposition d'adaptation du Voice Handicap Index à la voix chantée. *Revue de Laryngologie – Otologie - Rhinologie, 126(5),* 305-313.

Murry, T., Zschommler, A., Prokop, J. (2007). Voice Handicap in Singers. *Journal of Voice, 23(3),* 376-379.

Nunnally, J.C. (1978. *Psychometric theory (2nd edition).* New York: McGraw-Hill.

Ormezzano, Y. (2000). *Le guide de la voix.* Paris: Ed Odile Jacob.

Rosen, C.A., Lee, A.S., Osborne, J., Zullo, T., Murry, T. (2004). Development and Validation of the Voice Handicap Index-10. *The Laryngoscope, 114,* 1549-1556.

Rosen, C.A. & Murry, T. (2000). Voice Handicap Index in Singers. *Journal of Voice, 14(3),* 370-377.

Sataloff, R.T. (1996). Rational Thought: The impact of Voice Science upon Voice Care. *Journal of Voice, 9,* 215-234.

Sataloff, R.T. (1997). *Professional Voice: The Science and Art of Clinical Care, 2nd Edition.* Plural Publishing Inc.

Schmitt, C., (2003). *Essai d'adaptation d'une échelle d'auto-évaluation de la voix parlée (le "Voice Handicap Index" de Jacobson et al.,(1997)) à la voix chantée: évaluation de la pertinence des items de l'échelle sur une population de 71 sujets (chanteurs et professionnels de la voix).* Mémoire de licence en logopédie non publié. U.C.L., Louvain-la-Neuve.

Simon, C. (2004). *Essai d'adaptation d'une échelle d'auto-évaluation de la voix parlée (le "Voice Handicap Index" de Jacobson et al.,(1997)) à la voix chantée: validation d'une échelle sur une population de 37 chanteurs classiques dysphoniques, 95 chanteurs classiques normophoniques et 20 sujets non chanteurs sans plainte vocale.* Mémoire de licence en logopédie non publié. U.C.L., Louvain-La-Neuve.

Surel, B. (2006). *Contribution à l'adaptation du Voice handicap Index aux chanteurs.* Mémoire pour le Certificat de Capacité d'Orthophoniste. Université Montpellier I.

Woodson GE, Zwirner, P, Murry T, Swenson M. (1992). Functional assessment of patients with spasmodic dysphonia. *Journal of Voice*, *6,* 338-343.

World Health Organization. (1980). *International Classification of Impairments, Disabilities, and Handicaps: a manual of classification relating to the consequences of disease.* Geneva, Switzerland: World Health Organization. 25-43

Wuyts, F.L., De Bodt, M.S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Raes, J. & Van de Heyning, P.H. (2000). The Dysphonia Severity Index: an objective measure of vocal quality based on a multiparameter approach. *Journal of the Acoustical society of America*, *6,* 255-266.

Wuyts, F.L., De Bodt, M.S., Van de Heyning, P.H. (1999). Is the reliability of a visual analogue scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. *Journal of Voice*, *13,* 508-517.

Appendices

Appendix I

Voice Handicap Index aangepast aan de zangstem (2005)

Dominique Morsomme – Martine Gaspar

Datum :

Score : F /40
E /40
P /40
Totaal : /120

Hieronder vindt u enkele zinnen die door de meerderheid van de zangers gebruikt worden om hun stem en de weerslag daarvan op hun dagelijkse leven te beschrijven. Duid aan hoe vaak u elke ervaring beleefde gedurende de afgelopen maand. « Altijd » komt overeen met « ja, dit is altijd een probleem » ; « nooit » komt overeen met « neen, geen enkel probleem ».

		NOOIT	BIJNA NOOIT	SOMS	BIJNA ALTIJD	ALTIJD
F1	Ik heb moeilijkheden om van het ene register over te schakelen op het andere.					
P2	Ik heb het gevoel mij te moeten inspannen/dat ik mijn stem moet forceren om te zingen.					
F3	Ik vermijd zingen met begeleiding (piano, orkest, instrumentale groep).					
P4	Ik heb moeilijkheden om mijn zinnen af te maken.					
F5	Mijn stem komt moeilijk boven de begeleiding uit.					
F6	Ik heb het gevoel dat ik weerhouden word van « projecten » omwille van mijn stem.					
E7	Mijn stemprobleem ergert mij.					
F8	Mijn stemprobleem veroorzaakt een inkomensverlies.					
E9	Mijn stemprobleem weegt op mijn humeur.					
P10	Mijn stem laat me met tussenpozen in de steek.					
F11	Ik kan de hoge noten zingend niet (meer) bereiken.					
F12	Mijn stem is onstabiel (verzwakt in de loop van het spreken of zingen).					
P13	Mijn spreekstem is slechter nadat ik gezongen heb.					
P14	Ik voel irritatie of pijn in mijn larynx (strottenhoofd) wanneer ik zing.					
E15	Ik hou niet van de klankkleur van mijn stem (timbre, schelheid, ruwheid,…).					
F16	Ik pas mijn stem moeilijk aan in functie van de vocale prestaties (plaats, afstand, omgeving, grootte van het publiek, te vertolken werk, sfeer).					
P17	De klank van mijn stem varieert in de loop van eenzelfde zangprestatie, repetitie of concert.					
P18	Er zit luchtgeruis op mijn stem.					
F19	Zelfs na op te warmen kom ik niet tot een « goede stem ».					
P20	De klaarheid van mijn stem is onvoorspelbaar.					
P21	Ik heb de gewoonte om me sterk in te spannen om te zingen.					
F22	Ik heb er moeite mee om mijn emoties zingend over te brengen.					
E23	Ik vind dat anderen mijn stemproblemen tijdens het zingen niet begrijpen.					
E24	Het zingen is voor mij stresserend.					
E25	Ik voel me slechter omwille van mijn stem.					
P26	Mijn stem lijkt krakerig en droog.					
E27	Ik word bang/angstig bij het idee te moeten zingen.					
E28	Zelfs wanneer ik niet zing, denk ik aan mijn stemproblemen.					
E29	Het gebeurt dat ik weiger te zingen.					
E30	Het gebeurt dat ik de moed verlies als ik aan mijn stemproblemen denk.					

Hoe klinkt uw stem <u>vandaag</u> ? zoals gewoonlijk
slechter dan gewoonlijk
beter dan gewoonlijk

Contactadressen: dominique.morsomme@clin.ucl.ac.be martine.gaspar@scarlet.be

Appendix II

Voice Handicap Index aangepast aan de zangstem (2005)

Dominique Morsomme – Martine Gaspar

Datum :

Score : F /40
E /40
P /40
Totaal : /120

Hieronder vindt u enkele zinnen die de meerderheid van de zangers gebruiken om hun stem en de impact ervan op hun dagelijkse leven te beschrijven. Duid aan hoe vaak elke ervaring de afgelopen maand op u van toepassing was. « Altijd » komt overeen met « ja, dit is altijd een probleem » ; « nooit » komt overeen met « neen, geen enkel probleem ».

		NOOIT	BIJNA	SOMS	BIJNA	ALTIJD
F1	Ik heb moeilijkheden om van het ene register op het andere over te schakelen.					
P2	Ik heb het gevoel ~~*mij te moeten inspannen/*~~*dat ik mijn stem moet forceren* om te zingen.					
F3	Ik vermijd zingen onder ~~*muzikale*~~ begeleiding (piano, orkest, instrumentale groep).					
P4	Ik heb *moeilijkheden~~/problemen~~* om mijn zinnen af te maken.					
F5	Mijn stem komt moeilijk boven de ~~*muzikale*~~ begeleiding uit.					
F6	*Ik heb het gevoel dat ik niet bij « projecten » wordt betrokken* omwille van mijn stem.					
E7	Mijn stemprobleem ergert mij.					
F8	Mijn stemprobleem veroorzaakt een inkomensverlies.					
E9	Mijn stemprobleem weegt op mijn humeur.					
P10	Mijn stem laat me met tussenpozen in de steek.					
F11	Ik kan de hoge noten zingend niet bereiken.					
F12	Mijn stem is onstabiel (verzwakt in de loop van het spreken of zingen).					
P13	Mijn spreekstem is slechter nadat ik gezongen heb.					
P14	Ik voel irritatie of pijn in mijn *larynx/strottenhoofd* wanneer ik zing. *(volgens mij is de larynx wel bekend jargon voor professionele zangers).* ***Trouvé les 2 sur internet. Je ne connais pas de chanteurs professionnels, ds le même contexte, j'ai également trouvé 'strot'.***					
E15	Ik hou niet van de kleur van mijn stem (timbre, schelheid, korrel,…).					
F16	Ik pas mijn stem moeilijk aan in functie van de vocale prestaties (plaats, afstand, omgeving, grootte van het publiek, te vertolken werk, sfeer).					

P17	De klank van mijn stem varieert in de loop van eenzelfde zangprestatie, repetitie of concert.					
P18	Er zit lucht op mijn stem.					
F19	Zelfs na op te warmen kom ik niet tot een « goede stem ».					
P20	De helderheid van mijn stem is onvoorspelbaar.					
P21	Ik heb de gewoonte ~~*om me sterk in te spannen/om*~~ *veel moeite te doen* om te zingen.					
F22	Ik heb er moeite mee om mijn emoties al zingend te uiten.					
E23	Ik ~~*denk/*~~*heb het gevoel* dat anderen mijn stemmoeilijkheden tijdens het zingen niet begrijpen.					
E24	Ik vind het stresserend om te zingen.					
E25	Ik voel me slechter omwille van mijn stem.					
P26	Mijn stem klinkt krakerig en droog.					
E27	Het idee te moeten zingen maakt me bang.					
E28	Zelfs wanneer ik niet zing, denk ik aan mijn stemprobleem.					
E29	Soms weiger ik te zingen.					
E30	Mijn stemprobleem maakt me soms wanhopig.					

Hoe klinkt uw stem vandaag?

zoals gewoonlijk
slechter dan gewoonlijk
beter dan gewoonlijk

Contactadressen: dominique.morsomme@clin.ucl.ac.be martine.gaspar@scarlet.be

Appendix III

Voice Handicap Index aangepast aan de zangstem (2010)

Anneleen D'haeseleer – Marc De Bodt – Dominique Morsomme

Datum :

Score : F /40
E /40
P /40
Totaal : /120

Hieronder vindt u statements die door de meerderheid van de zangers gebruikt worden om hun (zang)stem en de weerslag daarvan op het dagelijkse leven beschrijven. Duid aan in welke mate de statements bij u van toepassing waren in de afgelopen maand. « Altijd » komt overeen met « ja, dit is altijd een probleem » ; « nooit » komt overeen met « neen, geen enkel probleem ».

		NOOIT	BIJNA NOOIT	SOMS	BIJNA ALTIJD	ALTIJD
F1	Ik heb moeilijkheden om van het ene register over te schakelen op het andere.					
P2	Ik heb het gevoel mij te moeten inspannen/dat ik mijn stem moet forceren om te zingen.					
F3	Ik vermijd zingen met begeleiding (piano, orkest, instrumentale groep).					
P4	Ik heb moeilijkheden om mijn zinnen af te maken.					
F5	Mijn stem komt moeilijk boven de begeleiding uit.					
F6	Ik heb het gevoel dat ik niet bij « projecten » word betrokken, omwille van mijn stem.					
E7	Mijn stemprobleem ergert mij.					
F8	Mijn stemprobleem veroorzaakt een inkomensverlies.					
E9	Mijn stemprobleem weegt op mijn humeur.					
P10	Mijn stem laat me met tussenpozen in de steek.					
F11	Ik kan de hoge noten zingend niet (meer) bereiken.					
F12	Mijn stem is onstabiel (verzwakt in de loop van het spreken of zingen).					
P13	Mijn spreekstem is slechter nadat ik gezongen heb.					
P14	Ik voel irritatie of pijn in mijn larynx (strottenhoofd) wanneer ik zing.					
E15	Ik hou niet van de klankkleur van mijn stem (timbre, schelheid, ruwheid,…).					
F16	Ik pas mijn stem moeilijk aan in functie van de vocale prestaties (plaats, afstand, omgeving, grootte van het publiek, te vertolken werk, sfeer).					
P17	De klank van mijn stem varieert in de loop van eenzelfde zangprestatie, repetitie of concert.					
P18	Er zit luchtgeruis op mijn stem.					
F19	Zelfs na op te warmen kom ik niet tot een « goede stem ».					

P20	De klaarheid van mijn stem is onvoorspelbaar.					
P21	Ik heb de gewoonte veel moeite te doen om te zingen.					
F22	Ik heb er moeite mee om mijn emoties zingend over te brengen.					
E23	Ik vind dat anderen mijn stemprobleem tijdens het zingen niet begrijpen.					
E24	Het zingen is voor mij stresserend.					
E25	Ik voel me slechter omwille van mijn stem.					
P26	Mijn stem klinkt krakerig en droog.					
E27	Ik word bang/angstig bij het idee te moeten zingen.					
E28	Zelfs wanneer ik niet zing, denk ik aan mijn stemprobleem.					
E29	Het gebeurt dat ik weiger te zingen.					
E30	Het gebeurt dat ik de moed verlies als ik aan mijn stemprobleem denk.					

Hoe klinkt uw stem <u>vandaag</u> ? zoals gewoonlijk

slechter dan gewoonlijk
beter dan gewoonlijk

Contactadressen: anneleen.dhaeseleer@ugent.be marc.de.bodt@uza.be dominique.morsomme@clin.ucl.ac.be

Appendix IV

Anamnese die samen met de 'VHI aangepast aan de zangstem' wordt afgenomen

Uw initialen (voor verwerking van uw gegevens) : ……
Geboortedatum : …../…../…..
Leeftijd : ……..
Geslacht : …….
Roker? : NEE – JA (….. sigaretten per dag)
Telefoon : ……

Datum eerste testafname : …../…../…..
Datum tweede testafname : …../…../…..

Stemklachten:

Stemklachten: JA – NEE

Indien ja:

a) Welke?……
……
……
b) Hoe lang heeft u deze stemklachten reeds? ……
……
……
c) Volgt u logopedische therapie voor deze stemklachten? JA – NEE

Graad van professionalisme : (Duid aan wat best past)

1. Professioneel (zingen is de enige bron van inkomen) ○
2. Semiprofessioneel (de proefpersoon is loontrekkende en zanger) ○
3. Amateur ○
4. Andere…… ○

Type zang : (Duid aan wat best past)

Klassieke zang ○
Andere : ……. ○

Solist ○
Koorlid ○
Andere : …….. ○

Classificatie zangstem : sopraan – mezzosopraan – alt – tenor – bariton – bas-
Andere……

Volgt u of heeft u gevolgd : les in klassieke zang? JA - NEE