Measuring the Dysphonia Severity Index (DSI) in the Program *Praat*

*,†,‡Youri Maryn, §Dominique Morsomme, and ‡,‖Marc De Bodt, *‡Antwerp, †‖Ghent, and §Liège, Belgium

Summary: Background. The original Dysphonia Severity Index (ie, DSI) weighs and combines four voice markers in a single number to size dysphonia gradation in the clinic: percent jitter (from Multi-Dimensional Voice Program; KayPENTAX Corp., Montvale, NJ), softest intensity and highest fundamental frequency (both from Voice Range Profile; KayPENTAX Corp.), and maximum phonation time. To be more generally applied, however, implementation of DSI in the program *Praat* (Paul Boersma and David Weenink, Institute for Phonetic Sciences, University of Amsterdam, The Netherlands) would be advantageous for all voice clinicians. The presented project was therefore designed (1) to develop such a *Praat* application and (b) to test its concurrent validity.

Methods. The four voice markers for the original DSI, as well as ten additional voice markers in *Praat*, were administered on a total of 49 subjects in three different clinical voice centers. A crossover research design was implemented to counterbalance for possible exercise effects. First, stepwise multiple linear regression was applied to build a statistical model with the best combination of *Praat* predictors for the original DSI. Second, correlation statistics were applied to compare *Praat*'s DSI with the original DSI.

Results. Both DSI versions correlated strongly. A custom script was therefore written for automated DSI determination in *Praat*.

Conclusion. With this script, every voice clinician can easily determine DSI in the *Praat* program.

Key Words: Dysphonia Severity Index–Praat–Multi-Dimensional Voice Program–Voice Range Profile–Variability.

INTRODUCTION

The Dysphonia Severity Index (DSI) is one of the outcomes of a multicentric study of the Belgian study group on voice disorders on 68 normophonic and 319 dysphonic subjects.[1] With 13 dependent acoustic and aerodynamic variables[a] in proportional odds logistic regression analysis, Wuyts et al[1] constructed a model containing a weighted combination of four factors to best predict auditorily perceived dysphonia severity (ie, grade or G): maximum phonation time (MPT), highest fundamental frequency (F_0High) and softest intensity (ILow) in the Voice Range Profile Model 4326 (VRP, KayPENTAX Corp., Montvale, NJ), and jitter percent (Jitt) in the Multi-Dimensional Voice Program Model 4305 (MDVP, KayPENTAX Corp.). The following was the resulting equation: $DSI = 0.13 \times MPT + 0.0053 \times F_0High - 0.26 \times ILow - 1.18 \times Jitt + 12.4$. Based on the classification table of observed versus predicted dysphonia severity, the classification accuracy of DSI ranged from 49.9% (with perfect agreement or "only diagonal" as the stringent criterion) to 98.4% (with "diagonal plus one off-diagonal" as the less stringent criterion). Furthermore, with −0.79, the Pearson correlation coefficient between the DSI and the Voice Handicap Index was reasonable. However, although G was the aimed variable to be predicted by this statistical model, measures of correlation and proportional relationship between DSI and G were not mentioned in Wuyts et al,[1] and therefore DSI's intrinsic criterion-related concurrent validity could not be interpreted.

Since its initial publication, the DSI has been investigated on multiple methodological aspects. Aichinger et al[2] compared DSI data from the lingWAVES system (WEVOSYS, Forchheim, Germany) and the DiVAS system (XION, Berlin, Germany), and found significant differences for F_0High and MPT, but not for ILow and jitter, and also not for the umbrella DSI. It should be mentioned, however, that both systems use different acoustic algorithms than the systems (ie, VRP and MDVP) applied in the original study. For correct interpretation of DSI scores, it is important to know about the eventual differences between subject groups. Hakkesteegt et al[3] showed significant decrease in DSI with advancing age in females as well as males. Awan and Ensslen[4] compared DSI data from trained vocalists (ie, those who previously participated in private voice lessons, choir activities, or musical theater under the guidance of a trained musical director) and untrained vocalists (ie, those without formal voice and singing lessons or directed singing experience). With F_0High, ILow, Jitt, and consequently also DSI being significantly higher in the group of trained vocalists, Awan and Ensslen[4] indicated alternative norm references for the DSI when used within specific groups of vocalists. A similar conclusion arises from Jayakumar and Savithri,[5] who evaluated the influence of geographical and ethnic variations on the DSI and found noticeable differences between Indian and European populations, and concluded that specific normative references are to be established for different populations. These kinds of differences are to be taken into account when interpreting DSI outcomes. Hussein Gaber et al[6] compared the DSI data—determined with a hardware and software

Accepted for publication January 4, 2017.

Conflict of interest: The authors declare that they do not have conflicts of interest.

[a]The 13 dependent variables were jitter percent, shimmer percent, noise-to-harmonics ratio, highest fundamental frequency, lowest fundamental frequency, range in fundamental frequency in Hertz, range in fundamental frequency in semitones, loudest intensity, softest intensity, range in intensity, maximum phonation time, vital capacity, and phonation quotient.

From the *Department of Otorhinolaryngology and Head and Neck Surgery, European Institute for ORL-HNS, Sint-Augustinus General Hospital, Antwerp, Belgium; †Department of Speech-Language Pathology and Audiology, University College Ghent, Ghent, Belgium; ‡Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium; §Speech and Language Department, Voice Unit, University of Liège, Liège, Belgium; and the ‖Faculty of Medicine & Health Sciences, Ghent University, Ghent, Belgium.

Address correspondence and reprint requests to Youri Maryn, Department of Otorhinolaryngology and Head and Neck Surgery, European Institute for ORL, Sint-Augustinus General Hospital, Oosterveldlaan 24, Wilrijk, Antwerp 2610, Belgium. E-mail: youri.maryn@gza.be

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2017.01.002

configuration very similar to the original system used in Wuyts et al[1]—with auditory-perceptual ratings of roughness, breathiness, and hoarseness (ie, overall degree of voice deviance), and obtained Spearman's correlation coefficients of –0.686, –0.609, and –0.487, respectively. Nemr et al[7] determined the DSI using the *Praat* program (Paul Boersma and David Weenink, Institute for Phonetic Sciences, University of Amsterdam, The Netherlands) for F_0High and Jitt and a dB meter for ILow. They compared the DSI data with ratings of overall dysphonia severity, breathiness, and roughness according to the protocol of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V[8]). In the group of 24 dysphonic subjects, the Pearson correlation coefficients were –0.503, –0.459, and –0.392, respectively. Furthermore, Hakkesteegt et al[9] showed that the DSI significantly differentiates between averaged dysphonia grade scores. Pebbili et al,[10] using methods similar to the original ones, provided norm-referencing data for the DSI in children without voice complaints and problems. Additionally, Awan et al[11] examined the test-retest variability of the DSI between three moments using the *TF32* software (Paul Milenkovic, Madison, WI), and although data decreased slightly across the three testing sessions no significant time-related effects emerged.

The initial purpose of the DSI was to objectively and quantitatively measure dysphonia severity based on very different sustained vocalizations: as comfortable, long, soft, and high as possible. Because most of the yielded measures (ie, MPT, F_0High, and ILow) are not directly related to overall voice sound quality, it is not so surprising that their combination approximates only a limited portion of auditory-perceptual ratings of, eg, hoarseness or overall dysphonia severity.[6,7] However, with measures strongly related to vocal capacity, the DSI is assumed to reflect more adequately the limitations in vocal functioning and may be applied as a global measure of vocal function or performance and voice disorder.[4]

Inconsistency in DSI between computer systems and programs, as found by Aichinger et al,[2] may arise from, eg, differences in pitch detection algorithm prior to determining jitter[12] or insufficiently calibrated sound level measures.[13] It is therefore important for the DSI to stick as much as possible to its original materials and methods. However, this corresponds with substantial financial investments, often only available in specialized voice centers within greater ENT departments. For its application to become more generalized, implementation of the DSI in a freely available yet high-standard program such as *Praat* would be beneficial to all voice clinicians. The present study was therefore designed (1) to explore the applicability of such an implementation (ie, using the *Praat* program to measure DSI), (2) to compare the DSI resulting from both protocols (ie, initial method with VRP and MDVP versus method with *Praat*), and (3) to approach the concurrent validity of DSI in *Praat*.

METHODS

Participants

The present multicentric research is based on data that were gathered from 19 subjects from the Sint-Jan General Hospital in Bruges (ie, Center 1), 16 subjects from the University Hospital of Liège (ie, Center 2), and 14 subjects from the University Hospital of Antwerp (ie, Center 3). Over a period of 2 months, these 49 subjects agreed to participate in this study that mainly consisted of providing voice samples for two independent DSI administrations. There were 28 females and 21 males, and their ages ranged from 10 to 82 years, with a mean of 43.6 years and a standard deviation of 19.8 years. Table 1 summarizes the variety of laryngeal diagnoses in this study group. This group of subjects is considered to be a representative sample of a population seeking help in a voice clinic. It reflects different age and gender groups, and it covers nonorganic as well as organic laryngeal pathologies.

TABLE 1.
List of the Absolute and Relative Occurrence of the Laryngeal Diagnoses of the 49 Participants of This Study

Laryngeal Diagnosis	Absolute Number	Relative Number
Nodules	8	16
UVFP	8	16
Normal	7	14
Functional dysphonia/MTD	5	10
Polypoid mucosa	4	8
After microphonosurgery for cyst or polyp	4	8
Vocal fold atrophy	3	6
Cyst/pseudocyst	2	4
Vocal fold immobility	2	4
After medialization thyroplasty for UVFP	2	4
After elongation thyroplasty for VF	1	2
Sulcus glottidis	1	2
Papillomatosis	1	2
Dysarthria	1	2

Abbreviations: UVFP, unilateral vocal fold paralysis; MTD, muscle tension dysphonia; VF, voice feminization.

Sound recordings

All voice samples were recorded with an AKG head-worn condenser microphone (AKG Acoustics GmbH, Vienna, Austria) placed between 5 cm and 10 cm from the mouth, and at approximately 45° off-axis from the direct airstream from the mouth. This microphone was connected to the Computerized Speech Lab (ie, CSL) hardware (KayPENTAX Corp., Montvale, NJ) with the switch to manually control the intensity of the sound recordings set in the middle of its range. This external CSL module was coupled to a desktop computer. Vocal sounds were recorded in an anechoic audiometry booth in Centers 1 and 2, and in a relatively quiet laboratorium room in Center 3. The samples were saved in a .wav format, and task-specific software was used to yield the various acoustic voice measures from these samples. Table 2 summarizes the recording hardware and software of the three data acquisition systems in this study.

Calibration of sound intensity measures

To compare the estimates of all vocal sound intensities (ie, not only of the softest intensities or measures of ILow) across the three recording settings, the sound intensity measurements in the *Praat* program (ie, the measured sound intensity or I_M)

TABLE 2.
Main Information Regarding Hardware and Software of the Three Computer Systems Used in This Study

Hardware/Software	Center 1	Center 2	Center 3
Computer type	Fujitsu Siemens Scenic P300 desktop computer	Custom desktop computer	Dell OptiPlex GX745 desktop computer
Microphone type	AKG C420, condenser, head-worn	AKG C520, condenser, head-worn	AKG C520, condenser, head-worn
Mouth-to-microphone angle	±45°	±45°	±45°
Mouth-to-microphone distance	±10 cm	±10 cm	±10 cm
External hardware	CSL Model 4500	CSL Model 4500	CSL Model 4500
Multi-Dimensional Voice Program	Model 5105, Version 2.6.2	Model 5105, Version 3.0.3	Model 5105, Version 3.0.3
Voice Range Profile	Model 4326, Version 2.6.2	Model 4326, Version 3.0.3	Model 4326, Version 3.0.3
Praat	Version 5.3.53	Version 5.3.53	Version 5.3.53

were calibrated based on the output of a CR:832C sound level meter (Cirrus Research plc, Hunmanby, North Yorkshire, UK) (ie, the expected sound intensity or I_E) and according to the following steps. First, an acceptable recording level of the computer's soundcard was sought by balancing between (1) decreasing its software switch so that the loudest vocal sounds were not clipped, and (2) increasing it so that the softest vocalizations were still recorded and could be adequately differentiated from background noise. All following sounds were recorded without any additional adjustments of the switches on the external CSL module and in the computer's configuration software. Second, audiometric speech noise (ie, acoustic noise specifically created from speech material to mask the spoken stimuli during speech audiometry), derived from all monosyllabic words from the audiometry lists of the Dutch Association for Audiology,[14,15] was used as calibration signal. Other signals have been advocated for the calibration of an audio recording system, such as a 1 kHz tone presented at 94 dB.[16] However, audiometric speech noise can be presented as long as needed at fixed intensities, and it is much more similar to natural speech than a tone. To proceed with the calibration protocol, the audiometric speech noise was radiated by an Inspire T12 loudspeaker Model MF1625 (Creative Technology Ltd., Singapore) at a minimum of four different intensities within the phonatory range and captured by the microphones of the clinical recording systems and the sonometer. The two microphones were held equidistantly at 10 cm from the loudspeaker. Third, for every intensity at which the calibration signal was played, both the output in dB_C of the sonometer and the output in dB of the *Praat* program were noted in an *Excel* spreadsheet (Microsoft Corp., Redmond, WA). For example, the calibration in Center 3 yielded the following corresponding data for the sonometer and the recording system with *Praat*, respectively: 50 dB_C and 37 dB, 61 dB_C and 39 dB, 71 dB_C and 48 dB, 81 dB_C and 55 dB, and 87 dB_C and 61 dB. Fourth, to convert the measured sound intensity into the real sound intensity, the $y=(a\times x)+b$ transformation formula was constructed based on regression analysis. With the calibration data from Center 3, this resulted in the formula $I_E=(1.426\times I_M)+1.543$. Similarly, the calibration procedure in Center 1 resulted in the formula $I_E=(0.989\times I_M)+8.339$, and in Center 2 it produced the formula $I_E=(1.0001\times I_M)-1.279$. Fifth, to automate the conversion of I_M data into I_E data, these formulae were implemented in the customized scripts for intensity analysis in the *Praat* program. Subsequently, a voice signal with, eg, I_M = 64 dB in Center 1 was automatically converted into the corresponding I_E = 72 dB_C in the *Praat* program, and the same was done with I_M = 49 dB in Center 3. In summary, following Maryn and Zarowski,[13] this calibration enabled an equalization of the sound intensities of the recordings across the three recording settings.

Dysphonia Severity Index: Original version

The first DSI method (ie, DSI_{ALFA}) in the present study simulated as much as possible the materials and procedures as originally applied and disseminated by Wuyts et al.[1] It consisted of establishing the following four measures in all participants across the three voice clinics: maximum phonation time (MPT), jitter percent (Jitt), lowest intensity (ILow), and highest fundamental frequency (F_0High). The participants were encouraged to execute these tasks with maximum performance, whereas the clinicians provided maximal verbal and nonverbal support. For the MPT, the participants were asked to sustain the vowel [a:] as long as possible after inhaling as deep as possible. Start and stop of this phonation were manually demarcated immediately afterwards in a narrowband spectrogram, and the duration between these two markers was noted. This was repeated three times and the longest duration counted as the MPT. For the Jitt, the subjects were asked to produce a sustained [a:] at a comfortable pitch and loudness for at least 3 seconds. The Jitt was selected from the statistical report of the Multi-Dimensional Voice Program (ie, MDVP; KayPENTAX Corp.) based on the medial 3 seconds of the sustained vowel. For the ILow, the participants were requested to phonate as quiet as possible while their voice signals were recorded and phonetographically fed back with the Voice Range Profile (ie, VRP; KayPENTAX Corp.). The sound intensity level of the quietest voice signal (ie, the mark located the lowest in the phonetogram) out of three repetitions was appointed as the ILow. For the F_0High, the participants were asked to phonate at their highest possible F_0 while their voice signals were recorded and phonetographically fed back with the VRP. The voice signal with the highest fundamental frequency (ie, the mark located most right in the phonetogram) out of three repetitions was selected as the F_0High. In this study, the alfa version of the DSI was calculated from these data with the original equation from Wuyts et al[1]: $DSI_{ALFA} = 12.4 + 0.13 \times MPT + 0.0053 \times F_0High - 0.26 \times ILow - 1.18 \times Jitt$.

Dysphonia Severity Index: Second version

To assist the general feasibility of the original DSI, this study sought for a beta application to determine the DSI in the *Praat* program (ie, DSI_{BETA}). The following methods were designed to simulate as closely as possible the original methods and signal processing strategies.

The 49 participants were therefore asked to repeat the three tasks to yield several acoustic measures in *Praat*. First, to assign a substitute measure for the Jitt, a *Praat* script was written to derive the following acoustic measures from recordings of the medial 3 seconds of a sustained [a:] phonation: jitter local ($Jitt_{LOC}$), jitter local absolute ($Jitt_{ABS}$), jitter rap ($Jitt_{RAP}$), and jitter ppq5 ($Jitt_{PPQ}$). Because comfortable sustained phonation was expected at fundamental frequency between 70 Hz and 600 Hz in normal as well as pathological voices, these values counted as "pitch floor" and "pitch ceiling," respectively. The additional parameters were kept at their standard value: maximum period factor = 1.3, maximum amplitude factor = 1.6, silence threshold = 0.03, and voicing threshold = 0.45. Second, to find a surrogate for the F_0High, a *Praat* script was written to search automatically for the highest fundamental frequency in recordings of three attempts to phonate at maximum pitch (ie, F_0Max). During this task, the F_0Max was expected to be found between 70 Hz and 1.3 kHz, and therefore these two values counted as "pitch floor" and "pitch ceiling" in cross-correlation method, respectively. Furthermore, to restrict the fundamental frequency analysis only to signal fragment with sufficient periodicity, the parameter "voicing threshold" was set sufficiently high to 0.80. Consequently, only signal fragments with an autocorrelation value of 0.80 or higher were considered voiced and were retained for further analysis, and all fragments with an autocorrelation value below 0.80 were considered unvoiced and dismissed from further analysis. This strategy prohibited unvoiced fragments to be included in the fundamental frequency analysis and guaranteed valid F_0Max measurements. Additional parameters in the cross-correlation method were retained at their standard value: time step = 0 second (ie, it uses a default time step of 0.25/pitch floor = 0.0036 seconds), maximum number of candidates = 15, very accurate = no (ie, it uses a default Hamming window of 3/pitch floor = 0.043 seconds), silence threshold = 0.03, octave cost = 0.01, octave-jump cost = 0.35, and voiced or unvoiced cost = 0.14. Third, to appoint a proxy for the ILow, another *Praat* script was written to search automatically for the lowest sound intensity in recordings of three attempts to phonate at the softest loudness (ie, IMin). Because the softest phonation was expected at a fundamental frequency between 70 Hz and 600 Hz in normal as well as pathological voices, these values counted as "pitch floor" and "pitch ceiling," respectively. Furthermore, for the same reason as in the F_0Max estimation (ie, to prohibit unvoiced signal fragments to be counted as voiced and thus to be included in the analysis), a voicing threshold = 0.80 was applied in the cross-correlation fundamental frequency analysis as well as the same standard parameter values as in the F_0Max script. Before taking the IMin of these voiced signals, the measured sound intensity data were converted into the corresponding expected data based on the abovementioned setting-specific calibration formulae. Fourth, the same MPT from the original DSI method was also used in the development of the beta version of the DSI.

The data of these "beta" measures (ie, $Jitt_{LOC}$, $Jitt_{ABS}$, $Jitt_{RAP}$, $Jitt_{PPQ}$, F_0Max, Imin, and MPT) were subjected to further statistical analysis.

Research design and statistical analyses

Sequencing two DSI methods carries the risk of extra exercise with better performance for the second DSI method. For example, Flament et al[17] found that F_0High was especially susceptible to encouragement and exercise. The possible effects of repetition were therefore counterbalanced with a crossover research design in which 25 participants followed the sequence "DSI_{BETA} after DSI_{ALFA}," and 24 participants followed the reverse sequence (ie, "DSI_{ALFA} after DSI_{BETA}").

All statistical analyses in the present study were completed using *SPSS* for Windows version 12.0 (SPSS Inc., Chicago, IL). First, stepwise multiple linear regression was executed to construct a statistical model representing the best combination of the "beta" predictors for the DSI_{ALFA}. A multiple regression equation was constructed based on the unstandardized coefficients of the statistical model. To simplify clinical interpretation, the model was linearly rescaled so that the outcomes of the equation resulted in a range of scores with the same minimum and maximum as the DSI_{ALFA}. This final model was called DSI_{BETA}. Second, the proportional relationship between DSI_{ALFA} and DSI_{BETA} was investigated with the parametric Pearson product-moment correlation coefficient (ie, r_P). The variance in DSI_{ALFA} that is explained by the variance in DSI_{BETA} was examined with the coefficient of determination (ie, r_p^2).

Finally, a new *Praat* script incorporating all required "beta" steps was developed to automate and standardize the calculation of the DSI_{BETA}. This script is represented in Appendix 1.

RESULTS

Descriptive data

Table 3 provides the descriptive data of MPT, F_0High, ILow, Jitt, and DSI_{ALFA} of the three voice centers (separated and together) involved in this study, following the authentic methods for determining DSI and according to Wuyts et al.[1] Table 4 offers the descriptive data of IMin, F_0Max, $Jitt_{LOC}$, $Jitt_{ABS}$, $Jitt_{RAP}$, $Jitt_{PPQ}$, and DSI_{BETA} of these voice centers (separated and together), following the custom-made *Praat* script (Appendix 1). Because MPT data for the beta method were copied from the alfa method, they are not iterated in Table 4.

Correlation between corresponding measures

Table 5 lists the r_Ps between corresponding measures (eg, F_0High and F_0Max). With $r_P = 0.816$ ($P = 0.000$), the correlation was strongest between the two measures of the maximum F_0 in the VRP. This indicates that two thirds (ie, 66.6%) of variance in F_0High is accounted for by variance in F_0Max, or vice versa. A nonsignificant $r_P = 0.248$ ($P = 0.085$) correlation was found between the two measures of the minimum sound level (ie, ILow from VRP and IMin from *Praat*'s script) in the VRP. This correlation

TABLE 3.
Descriptive Information of the Alfa Measures

	Min	Max	Mean	SD
Center 1 (N = 19)				
MPT, s	6.89	23.63	14.63	4.82
ILow, dB	43.00	71.00	50.89	6.07
F_0High, Hz	110.00	1108.73	686.35	279.83
Jitt, %	0.41	7.01	2.31	1.70
DSI_{ALFA}	–9.96	7.23	1.94	4.47
Center 2 (N = 16)				
MPT, s	6.38	29.48	14.82	6.36
ILow, dB	50.00	63.00	56.00	3.16
F_0High, Hz	207.65	739.99	438.16	155.82
Jitt, %	0.53	2.52	1.58	0.67
DSI_{ALFA}	–2.83	4.18	0.40	1.68
Center 3 (N = 14)				
MPT, s	2.00	14.90	8.56	3.70
ILow, dB	45.00	66.00	52.21	5.59
F_0High, Hz	207.00	932.00	448.92	225.04
Jitt, %	0.60	3.80	1.97	1.14
DSI_{ALFA}	–4.40	5.60	0.10	2.66
Combined (N = 49)				
MPT, s	2.00	29.48	12.94	5.74
ILow, dB	43.00	71.00	52.94	5.51
F_0High, Hz	110.00	1108.73	537.47	254.98
Jitt, %	0.41	7.01	1.97	1.29
DSI_{ALFA}	–9.96	7.23	0.91	3.32

Abbreviations: DSI, Dysphonia Severity Index; MPT, maximum phonation time; SD, standard deviation.

was the weakest across the different pairs of corresponding measures. This means that only 6.2% of variance in ILow is explained by the variance in IMin, or vice versa. Regarding indices of jitter, the strongest correlation occurred between Jitt and $Jitt_{PPQ}$ ($r_P = 0.617$, $P = 0.000$), with 38.1% of variance in Jitt being justified by variance in $Jitt_{PPQ}$, or vice versa. In a later stage of this study (see next paragraph), this $Jitt_{PPQ}$ was selected to the surrogate for DSI_{ALFA}'s Jitt in DSI_{BETA}.

Beta version of DSI

With F_0Max, IMin, $Jitt_{LOC}$, $Jitt_{ABS}$, $Jitt_{RAP}$, $Jitt_{PPQ}$, and MPT as dependent variables, stepwise multiple regression analysis revealed that a combination of four "beta" variables best predicted the original DSI. The equation for the new model, based on the unstandardized coefficients of the regression, was as follows: $DSI_{BETA} = -0.149 + 0.140 \times MPT + 0.00450 \times F_0Max - 0.0329 \times IMin - 4.530 \times Jitt_{PPQ}$. The outcomes of this equation ranged from –9.625 to 5.067. This equation was linearly rescaled, so its scale resembled as much as possible the DSI_{ALFA} scale with values between –9.960 and 7.230. The resulting equation was $DSI_{BETA} = 1.127 + 0.164 \times MPT + 0.0053 \times F_0Max - 0.038 \times IMin - 5.30 \times Jitt_{PPQ}$.

The proportional relationship between the DSI_{ALFA} and the DSI_{BETA} is clarified by the scatterplot in Figure 1. Based on the $r_P = 0.857$ ($P = 0.000$) between these data and the corresponding $r_p^2 = 0.734$, it can be stated that the variance in DSI_{BETA} represented 73.4% of the variance in DSI_{ALFA}.

TABLE 4.
Descriptive Information of the Beta Measures

	Min	Max	Mean	SD
Center 1 (N = 19)				
IMin, dB	51.5	79.72	57.50	6.73
F_0Max, Hz	98.19	1156.66	812.40	295.72
$Jitt_{LOC}$, %	0.17	2.98	0.72	0.70
$Jitt_{ABS}$, s	0.01	0.40	0.05	0.09
$Jitt_{RAP}$, %	0.09	1.49	0.39	0.36
$Jitt_{PPQ}$, %	0.09	1.96	0.44	0.46
DSI_{BETA}	–10.25	6.93	3.26	4.08
Center 2 (N = 16)				
IMin, dB	36.11	60.79	50.02	7.12
F_0Max, Hz	267.70	835.70	476.31	148.44
$Jitt_{LOC}$, %	0.27	2.61	0.61	0.60
$Jitt_{ABS}$, s	0.01	0.20	0.04	0.05
$Jitt_{RAP}$, %	0.15	1.46	0.34	0.34
$Jitt_{PPQ}$, %	0.17	0.99	0.33	0.24
DSI_{BETA}	–2.70	6.04	2.39	2.41
Center 3 (N = 14)				
IMin, dB	43.10	73.05	60.78	8.75
F_0Max, Hz	237.04	1069.66	547.13	289.20
$Jitt_{LOC}$, %	0.21	2.54	0.67	0.60
$Jitt_{ABS}$, s	0.01	0.20	0.05	0.05
$Jitt_{RAP}$, %	0.12	1.65	0.40	0.39
$Jitt_{PPQ}$, %	0.13	1.07	0.37	0.26
DSI_{BETA}	–4.51	5.33	1.11	2.87
Combined (N = 49)				
IMin, dB	36.11	79.72	55.99	8.55
F_0Max, Hz	98.19	1156.66	626.87	292.18
$Jitt_{LOC}$, %	0.17	2.98	0.67	0.63
$Jitt_{ABS}$, s	0.01	0.40	0.05	0.07
$Jitt_{RAP}$, %	0.09	1.65	0.38	0.35
$Jitt_{PPQ}$, %	0.09	1.96	0.38	0.34
DSI_{BETA}	–10.25	6.93	2.36	3.33

Abbreviations: DSI, Dysphonia Severity Index; SD, standard deviation.

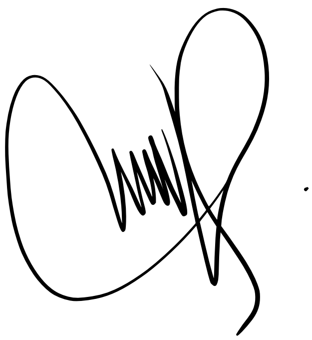
TABLE 5.
Pearson Correlation Coefficients Between Alfa Measures and Their Corresponding Beta Measures, Their Significance Levels, and Coefficients of Determination

Alfa Measure	Beta Measure	r_P	*P*	r_p^2
F_0High	F_0Max	0.816	0.000	0.666
ILow	IMin	0.248	0.085	0.062
MPT	MPT	N/A	N/A	N/A
Jitt	$Jitt_{LOC}$	0.551	0.000	0.304
Jitt	$Jitt_{ABS}$	0.422	0.003	0.178
Jitt	$Jitt_{RAP}$	0.542	0.000	0.294
Jitt	$Jitt_{PPQ}$	0.617	0.000	0.381

Abbreviations: MPT, maximum phonation time; N/A, not applicable because data in both methods were identical.

DISCUSSION

The DSI is a well-established, objective, and multivariable method to measure dysphonia severity. The vocological research group of the University of Ghent, eg, has used it extensively to evaluate the



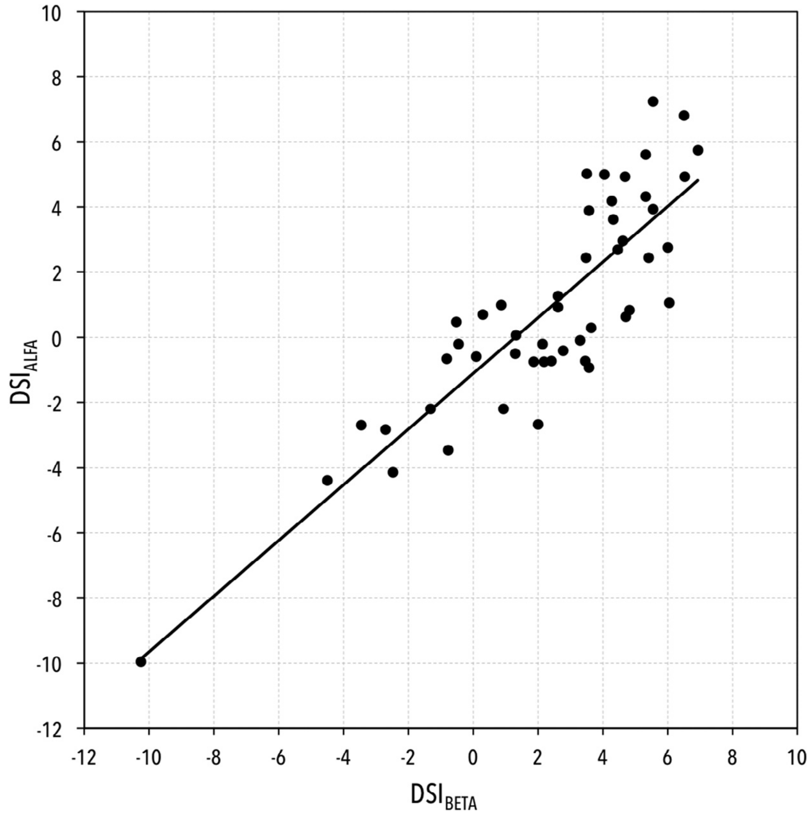


FIGURE 1. Scatterplot illustrating the proportional relationship between the DSI_{ALFA} and the DSI_{BETA}. The linear regression line is defined by $y = 0.8548x+1.1086$ and corresponds with $r_p^2 = 0.735$. DSI, Dysphonia Severity Index.

outcomes of various treatment techniques and methods.[18–22] However, having to use its original materials and methods to rule out intersystem variability,[2] it cannot be applied by all voice clinicians due to organizational (ie, financial or otherwise) constraints. Because *Praat* is freely downloadable and commonly used for both clinical and research purposes, this study was designed to develop a procedure to administer the DSI with *Praat* and to compare the outputs of this procedure with their original counterparts.

In its final stage, in which two methods were administered on 49 patients across three different Belgian voice centers, the project yielded a customized macro to be implemented in *Praat* for single-program determination of DSI_{BETA} as a summary of the four individually weighted measures: IMin, F_0Max, $Jitt_{PPQ}$, and MPT. Prior to running this script, (1) MPT has to be determined, (2) samples from comfortable, softest, and highest vowel phonations have to be recorded, and (3) renamed to read "Sound ppq", "Sound il", and "Sound fh", respectively. After initiating the script, the clinician has to complete the sound level calibration formula as well as the previously determined MPT in a fill-in form. Examples of output of this script in a normophonic case and a dysphonic case are illustrated in Figures 2 and 3. Once MPT is determined and appropriate samples are recorded (and sometimes edited), DSI_{BETA} can easily and rapidly be calculated without additional manual interference and having to leave the *Praat* program.

DSI_{BETA}'s concurrent validity (ie, how proportionally related is DSI_{BETA} when compared with DSI_{ALFA}?) is acceptable yet incomplete as it carries $r_P = 0.851$ with DSI_{ALFA}. MPT and the corresponding factors F_0High and F_0Max contributed most to this correlation. This was completely understandable for MPT, as these data were identical in the two DSIs. This was predictable for highest F_0, as multiple studies already indicated measures of fundamental frequency to be immune to recording system, environmental noise, and their combination.[25–31] However, why $r_P = 0.816$ between F_0High and F_0Max subsides compared with correlations between mean/habitual/comfortable F_0 measures in these studies remains difficult to explain. Differences in window length, smoothing factor, interval duration, pitch detection algorithm, etc between VRP programs and *Praat*'s/*Praat* script's algorithm are hypothesized to explain this finding.

The corresponding ILow and IMin factors, as well as the Jitt and $Jitt_{LOC}$ factors, on the other hand, added less to the correlation between the two DSIs. For the minimum sound level measures, the low $r_P = 0.248$ may also be explained by differences in window length, interval duration, pitch detection algorithm, and calibration. Finally, the $r_P = 0.422$ between Jitt and $Jitt_{LOC}$ is situated somewhere between correlation values found in earlier studies juxtaposing Jitt and $Jitt_{LOC}$ (ie, $r_P = 0.366$ with outlying and extreme data removed in Maryn et al,[12] $r_P = 0.36$ with a minority of jitter values exceeding 0.3% removed in Amir et al,[29] and $r_P = 0.899$ in Oğuz et al).[32] However, because of superior correlation, $Jitt_{PPQ}$ instead of $Jitt_{LOC}$ was chosen as stand-in for DSI_{ALFA}'s Jitt.

The following are the strengths of this study. First, data were gathered in three voice clinics or rooms, each with different environmental acoustics and slightly differing recording equipment. The findings therefore transcend any particular data acquisition situation and are hypothesized to be more generalizable. Second, sound

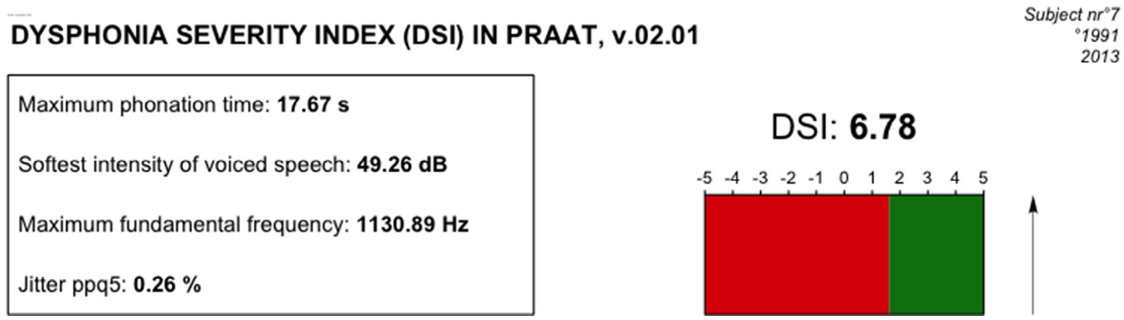


FIGURE 2. Example of the graphical output of the *Praat* script for the DSI_{BETA}. Subject 7 was a woman of 21 years with a normal larynx. With DSI_{ALFA} = 1.6 as cutoff point between normal and disordered voice functions (based on Wuyts et al,[1] and denoted by the boundary between the green and the red boxes), DSI_{BETA} = 6.78 confirms normophonia. Indeed, when compared with normative data from De Bodt et al,[23] her MPT is sufficiently long (normative mean = 16.2 seconds), sound intensity of her softest voiced phonation is adequately low (normative mean = 51.2 dB), and F_0 of her highest pitch is satisfactorily high (normative mean = 867 Hz). Furthermore, compared with normative data from Trauwaen et al,[24] her $Jitt_{PPQ}$ was relatively low (normative mean = 0.24%). DSI, Dysphonia Severity Index; MPT, maximum phonation time. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

levels in *Praat* were calibrated according to Maryn and Zarowski,[13] and the completion form of the DSI_{BETA} script has the option to automatically proceed with calibrated sound levels. Third, in contrast to, eg, Aichinger et al[2] who investigated nonoriginal DSI methods with lingWAVES and DiVAS, the present study determined DSI_{ALFA} with materials and methods identical to those described by Wuyts et al.[1] Consequently, the criterion upon which *Praat*'s DSI_{BETA} was validated mimicked the authentic DSI method, enabling the same guidelines for interpretation.

This study also has limitations. First, the subject group size was relatively small in total, as well as per voice center, and the age range was wide. Although fairly representative for clinical voice center population, the number of participants was low, and therefore not all dysphonia severity levels may have been included. Second, while CSL's MDVP or VRP was recording, it was impossible to make recordings with another program (eg, *Praat*) on the same computer. Because at the moment this study was running, there were no three exemplars (ie, for the three sites or hospitals) of the same recording equipment available during clinical activities and in the assessment rooms, and thus samples for DSI_{ALFA} and DSI_{BETA} could not be recorded simultaneously. It was therefore decided to proceed with successive recordings while trying to counterbalance any order effect with a cross-over design. Nevertheless, vocalizations for both methods may have been different, causing true phonatory differences that may have been misinterpreted as recording and analysis-based differences in a later stage of the study. Third, as it was not the scope of this project, and because only sustained phonations were recorded (ie, there were no recordings of continuous speech), this study did not provide insight into the proportional relationship between DSI and auditory perception or interpretation of overall voice quality. How well DSI measures the degree of overall dysphonia still remains unclear. Additional validation of DSI methods—not only the original and *Praat*'s version, but also other versions (eg, from lingWAVES and DiVAS)—as clinical measures of dysphonia severity is therefore necessary.

In conclusion, the main goal of this project was to provide simple access to reliably measure DSI in a feasible way. Although DSI_{BETA} correlated strongly with DSI_{ALFA} and every voice clinician can easily determine DSI_{BETA} in the *Praat* program, future research is still needed to refine its methods and to additionally validate its clinical utilization.

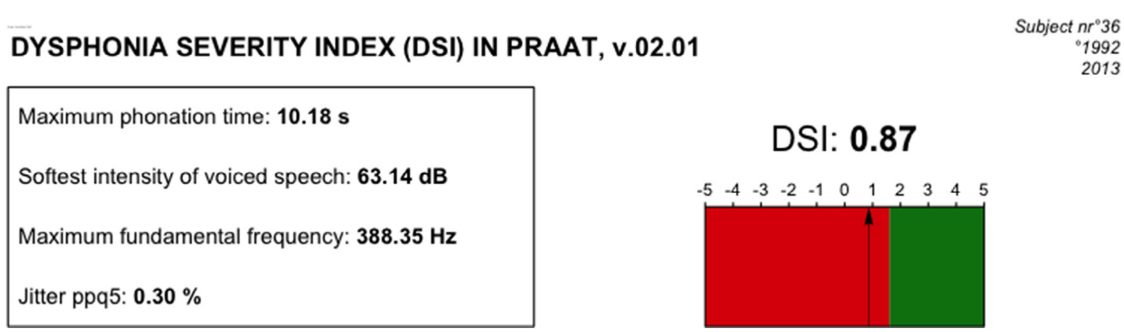


FIGURE 3. Example of the graphical output of the *Praat* script for the DSI_{BETA}. Subject 36 was a woman of 21 years diagnosed with polypoid vocal fold mucosa. With DSI_{ALFA} = 1.6 as demarcation between normal and disordered voice functions (based on Wuyts et al,[1] and denoted by the boundary between the green and the red boxes), DSI_{BETA} = 0.87 confirms dysphonia. Indeed, when compared with normative data from De Bodt et al,[23] her MPT is relatively but not too short (normative mean = 16.2 seconds), sound intensity of her softest voiced phonation is increased (normative mean = 51.2 dB), F_0 of her highest pitch is too low (normative mean = 867 Hz), and finally compared with normative data from Trauwaen et al,[24] her $Jitt_{PPQ}$ was too high (normative mean = 0.24%). DSI, Dysphonia Severity Index; MPT, maximum phonation time. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

APPENDIX 1. BETA VERSION OF THE SCRIPT FOR THE CALCULATION OF THE DYSPHONIA SEVERITY INDEX IN THE PROGRAM *PRAAT* (VERSION 5.3.53 OR HIGHER), © YOURI MARYN (14/09/2013)

```
# Part 1 of Praat-script: sheet with introductory text, instructions and
# the possibility to adjust the calibration factor and to complete the
# MPT and patient information.

form Dysphonia Severity Index in Praat, v.02.01
comment >>>  The original Dysphonia Severity Index (i.e., DSI) is developed
comment by Wuyts et al. (2000). It is a multivariate measure of the degree of
comment of dysphonia and consists of a set of four weighted measures:
comment maximum phonation time, highest fundamental frequency (from the
comment Voice Range Profile of the KayPentax CSL), softest intensity (also from the
comment Voice Range Profile of the KayPentax CSL) and percent jitter (from the
comment Multi-Dimensional Voice Program of the KayPentax CSL).
comment >>> The output and the measures of this script are not equal to but highly
comment resemble the output and the measures of the original DSI: maximum
comment phonation time (i.e., 'mpt'), highest fundamental frequency or f0-high (i.e.
comment 'fh'), softest intensity of voiced speech or I-low (i.e., 'il') and jitter
comment ppq5 (i.e., 'ppq'). Their weighted combination is strongly correlated to
comment the original DSI-data.
comment >>> Make sure that the intensity measurements are calibrated. Does the
comment calibration method necessitate the implementation of a calibration factor?
choice choose: 2
button no
button yes:
sentence Calibration_factor 1*self+10
comment >>> For this script it is important to (1) have a recording of phonation(s) as
comment high as possible and name it 'fh', (2) have a recording of phonation(s) as soft
comment as possible and name it 'il', (3) have a recording of at least three medial
comment seconds of a sustained vowel [a:] (i.e., without voice onset and voice offset)
comment and name it 'ppq', and (4) fill in the maximum phonation time here:
positive Maximum_phonation_time_(s)
comment >>> Additional information (optional):
sentence name_patient
sentence left_dates_(birth_-_assessment)
sentence right_dates_(birth_-_assessment)
comment
comment © Youri Maryn (14/09/2013)
endform
```

```
mpt = maximum_phonation_time
Erase all
Select inner viewport... 0.5 7.5 0.5 4.5
Axes... 0 1 0 1
Black
Text special... 0.5 centre 0.6 half Helvetica 12 0 Please wait an instant. Depending on the
duration and/or the sample rate of the recorded
Text special... 0.5 centre 0.4 half Helvetica 12 0 sound files, this script takes more or less
time to process the sounds and calculate the DSI-2.

# Part 2 of Praat-script: determination of IMIN. The sound recording
# with the softest phonations should be named “Sound im”.

select Sound im
To Pitch (cc)... 0 70 15 no 0.03 0.8 0.01 0.35 0.14 600
select Sound im
plus Pitch im
To PointProcess (cc)
select PointProcess im_im
To TextGrid (vuv)... 0.02 0.01
select Sound im
plus TextGrid im_im
Extract intervals where... 1 no "is equal to" V
Concatenate
if choose = 1
To Intensity... 60 0 yes
elsif choose = 2
To Intensity... 60 0 yes
Formula...   'calibration_factor$'
endif
minimumIntensity = Get minimum... 0 0 None

# Part 3 of Praat-script: determination of F0HIGH. The sound recording
# with the highest phonations should be named “Sound fh”.

select Sound fh
To Pitch (cc)... 0 70 15 no 0.03 0.8 0.01 0.35 0.14 1300
maximumF0 = Get maximum... 0 0 Hertz None
```

```
# Part 4 of Praat-script: determination of JITTER PPQ5. The sound recording
# with sustained [a:] phonation should be named "Sound ppq".
select Sound ppq
durationVowel = Get total duration
durationStart=durationVowel-3
if durationVowel>3
Extract part... durationStart durationVowel rectangular 1 no
Rename... ppq2
elsif durationVowel<=3
Copy... ppq2
endif
To Pitch... 0 70 600
select Sound ppq2
plus Pitch ppq2
To PointProcess (cc)
select Sound ppq2
plus Pitch ppq2
plus PointProcess ppq2_ppq2
voiceReport$ = Voice report... 0 0 70 600 1.3 1.6 0.03 0.45
jitterPpq5Pre = extractNumber (voiceReport$, "Jitter (ppq5): ")
jitterPpq5 = jitterPpq5Pre*100

# Part 5 of Praat-script: calculation of DSIbeta.

dsi2=1.127+0.164*mpt-0.038*minimumIntensity+0.0053*maximumF0-5.30*jitterPpq5

# Part 6 of Praat-script: drawing all the information and relevant graphs.
# To insert this output in another program, just use to 'paste'-function in e.g.
# the text editor after this script is terminated.

Erase all
Solid line
Line width... 1
Black
Helvetica
Font size... 1
Select inner viewport... 0.5 7.5 0.1 0.15
Axes... 0 1 0 1
```

```
Text... 0 Left 0.5 Half Script: Youri Maryn, PhD
12
Select inner viewport... 0.5 7.5 0 0.5
Axes... 0 1 0 1
Text... 0 Left 0.5 Half ##DYSPHONIA SEVERITY INDEX (DSI) IN PRAAT, v.02.01#
Font size... 8
Select inner viewport... 0.5 7.5 0 0.5
Axes... 0 1 0 3
Text... 1 Right 2.3 Half %%'name_patient$'%
Text... 1 Right 1.5 Half %%°'left_dates$'%
Text... 1 Right 0.7 Half %%'right_dates$'%
Font size... 10
Select inner viewport... 0.5 7.5 0.5 2
Axes... 0 7 4 0
Text... 0.05 Left 0.5 Half Maximum phonation time: ##'mpt:2' s#
Text... 0.05 Left 1.5 Half Softest intensity of voiced speech: ##'minimumIntensity:2' dB#
Text... 0.05 Left 2.5 Half Maximum fundamental frequency: ##'maximumF0:2' Hz#
Text... 0.05 Left 3.5 Half Jitter ppq5: ##'jitterPpq5:2' \% #
Select inner viewport... 0.5 3.8 0.5 2
Draw inner box
Font size... 8
Arrow size... 1
Select inner viewport... 4 7.5 1.25 2
Axes... -10 10 1 0
Paint rectangle... green 1.6 5 0 1
Paint rectangle... red -5 1.6 0 1
Draw arrow... dsi2 1 dsi2 0
One mark top... -5 yes yes no
One mark top... -4 yes yes no
One mark top... -3 yes yes no
One mark top... -2 yes yes no
One mark top... -1 yes yes no
One mark top... 0 yes yes no
One mark top... 1 yes yes no
One mark top... 2 yes yes no
One mark top... 3 yes yes no
One mark top... 4 yes yes no
One mark top... 5 yes yes no
Font size... 16
```

```
Select inner viewport... 4 7.5 0.5 1.15
Axes... 0 1 0 1
Text... 0.5 Centre 0.5 Half DSI: ##'dsi2:2'#
Select inner viewport... 4.875 6.625 1.25 2
Draw inner box
Select inner viewport... 0.5 7.5 0 2
Copy to clipboard
select all
minus Sound im
minus Sound fh
minus Sound ppq
Remove
```

REFERENCES

1. Wuyts FL, De Bodt MS, Molenberghs G, et al. The Dysphonia Severity Index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*. 2000;43:769–809.
2. Aichinger P, Feichter F, Aichstill B, et al. Inter-device reliability of DSI measurement. *Logop Phoniatr Vocol*. 2012;37:167–173.
3. Hakkesteegt M, Brocaar M, Wieringa M, et al. Influence of age and gender on the Dysphonia Severity Index. A study of normative values. *Folia Phoniatr Logop*. 2006;58:264–273.
4. Awan SN, Ensslen AJ. A comparison of trained and untrained vocalists on the Dysphonia Severity Index. *J Voice*. 2010;24:661–666.
5. Jayakumar T, Savithri S. Effect of geographical and ethnic variation on Dysphonia Severity Index: a study of Indian population. *J Voice*. 2012;26:e11–e16.
6. Hussein Gaber AG, Liang F-Y, Yang J-S, et al. Correlation among the Dysphonia Severity Index (DSI), the RBH voice perceptual evaluation, and minimum glottal area in female patients with vocal fold nodules. *J Voice*. 2014;28:20–23.
7. Nemr K, Simões-Zenari M, de Souza GS, et al. Correlation of the Dysphonia Severity Index (DSI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and gender in Brazilians with and without voice disorders. *J Voice*. 2016;30:765.e7–765.e11.
8. Kempster GB, Gerratt BR, Verdolini Abbott K, et al. Consensus Auditory-Perceptual Evaluation of Voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol*. 2009;18:124–132.
9. Hakkesteegt M, Brocaar M, Wieringa M, et al. The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *J Voice*. 2008;22:138–145.
10. Pebbili GK, Kidwai J, Shabnam S. Dysphonia Severity Index in typically developing Indian children. *J Voice*. 2017;31:125.e1–125.e6.
11. Awan SN, Miesemer SA, Nicolia TA. An examination of intrasubject variability on the Dysphonia Severity Index. *J Voice*. 2012;26:814.e21–814.e25.
12. Maryn Y, Corthals P, De Bodt M, et al. Perturbation measures of voice: a comparative study between multi-dimensional voice program and Praat. *Folia Phoniatr Logop*. 2009;61:217–226.
13. Maryn Y, Zarowski A. Calibration of clinical audio recording and analysis systems for sound intensity measurement. *Am J Speech Lang Pathol*. 2015;24:608–618.
14. Bosman A. *Speech perception by the hearing impaired* [PhD thesis]. Utrecht, The Netherlands: University of Utrecht; 1989.
15. Bosman A, Wouters J, Damman W. Realisatie van een CD voor spraakaudiometrie in Vlaanderen [Realization of a CD for speech audiometry in Flanders]. *Tijdschr Logop Audiol*. 1995;9:218–225.
16. Howard DM, Murphy DT. *Voice Science, Acoustics and Recording*. San Diego, CA: Plural Publishing Inc.; 2008.
17. Flament E, Roels S, Maryn Y. *De invloed van aanmoediging bij automatische fonetografie [The influence of encouragement on automatic phonetography]* [Bachelor thesis]. Bruges, Belgium: Katholieke Hogeschool Brugge-Oostende; 2005.
18. Van Lierde K, Claeys S, De Bodt M, et al. Long-term outcome of hyperfunctional voice disorders based on a multiparameter approach. *J Voice*. 2007;21:179–188.
19. Van Lierde K, De Bodt M, D'haeseleer E, et al. The treatment of muscle tension dysphonia: a comparison of two treatment techniques by means of an objective multiparameter approach. *J Voice*. 2010;24:294–301.
20. D'haeseleer E, Claeys S, Van Lierde K. The effectiveness of manual circumlaryngeal therapy in future elite vocal performers: a pilot study. *Laryngoscope*. 2013;123:1937–1941. doi:10.1002/lary.24050.
21. Meerschman I, D'haeseleer E, De Cock E, et al. Effectiveness of chewing technique on the phonation of female speech-language pathology students: a pilot study. *J Voice*. 2015;30:574–578. doi:10.1016/j.jvoice.2015.06.016.
22. Meerschman I, Bettens K, Dejagere S, et al. Effect of two isolated vocal-facilitating techniques chant talk and pitch inflections on the phonation of female speech-language pathology students: a pilot study. *J Voice*. 2015;30:771.e17–771.e25. doi:10.1016/j.jvoice.2015.11.007.
23. De Bodt M, Heylen L, Mertens F, et al. *Stemstoornissen. Handboek voor de Klinische Praktijk [Voice disorders. Handbook for clinical practice]*. 6th ed. Antwerp, Belgium: Garant; 2015.
24. Trauwaen I, Caelenberghe E, Verstraete J, et al. *Normatieve studie inzake akoestische stemkwaliteit [Normative study regarding acoustic voice quality]* [Bachelor thesis]. Bruges, Belgium: Katholieke Hogeschool Brugge-Oostende; 2009.
25. Bielamowicz S, Kreiman J, Gerratt BR, et al. Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res*. 1993;39:126–134.
26. Karnell MP, Hall KD, Landahl KL. Comparison of fundamental frequency and perturbation measurements among three analysis systems. *J Voice*. 1995;9:383–393.
27. Carson CP, Ingrisano DRS, Eggleston KD. The effect of noise on computer-aided measures of voice: a comparison of CSpeechSP and the multi-dimensional voice program software using the CSL 4300B module and multi-speech for windows. *J Voice*. 2003;17:12–20. doi:10.1016/S0892-1997(03)00031-6.
28. Smits I, Ceuppens P, De Bodt MS. Comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. *J Voice*. 2005;19:187–196.
29. Amir O, Wolf M, Amir N. A clinical comparison between two acoustic analysis softwares: MDVP and Praat. *Biomed Signal Process Control*. 2009;4:202–205.
30. Mat Baki M, Wood G, Alston M, et al. Reliability of OperaVOX against Multidimensional Voice Program (MDVP). *Clin Otolaryngol*. 2014;40:22–28.
31. Maryn Y, Ysenbaert F, Zarowski A, et al. Mobile communication devices, ambient noise, and acoustic voice measures. *J Voice*. 2016;doi:10.1016/j.jvoice.2016.07.023. Submitted.
32. Oğuz H, Kiliç MA, Şafak MA. Comparison of results in two acoustic analysis programs: Praat and MDVP. *Turk J Med Sci*. 2011;41:835–841.