Logopedics Phoniatrics Vocology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ilog20

Validation of the acoustic breathiness index to the Brazilian Portuguese language

Marina Englert, Ben Barsties v. Latoszek, Youri Maryn & Mara Behlau

To cite this article: Marina Englert, Ben Barsties v. Latoszek, Youri Maryn & Mara Behlau (2021): Validation of the acoustic breathiness index to the Brazilian Portuguese language, Logopedics Phoniatrics Vocology, DOI: 10.1080/14015439.2020.1864467

To link to this article: https://doi.org/10.1080/14015439.2020.1864467

Published online: 06 Jan 2021.

LOGOPEDICS PHONIATRICS VOCOLOGY
https://doi.org/10.1080/14015439.2020.1864467

RESEARCH ARTICLE

Validation of the acoustic breathiness index to the Brazilian Portuguese language

Marina Englert[a], Ben Barsties v. Latoszek[b,c], Youri Maryn[d*] and Mara Behlau[a]

[a]Unifesp – "Universidade Federal de São Paulo" and CEV – "Centro de Estudos da Voz", São Paulo, Brazil; [b]Speech-Language Pathology, SRH University of Applied Health Sciences, Düsseldorf, Germany; [c]Department of Phoniatrics and Pediatric Audiology, University Hospital Münster, Westphalian Wilhelm University, Münster, Germany; [d]ENT Department, Sint-Augustinus GZA, Wilrijk, Belgium

ABSTRACT

Objective: To assess the concurrent validity and the diagnostic accuracy of the Acoustic Breathiness Index (ABI) in Brazilian Portuguese.

Methods: The counting numbers 1–20 and the vowel /a/ of 150 subjects were recorded (37 vocally healthy and 113 with dysphonia). The analyzed samples were the counting number 1–11 and 3 s of the sustained vowel. Nine voice specialists performed the perceptual judgment of the degree of breathiness. The Spearman Correlation and the receiver operating characteristic (ROC) curve were used to assess ABI's concurrent validity and diagnosis accuracy.

Results: Results from five listeners were chosen for the study analyses due to moderate and substantial intra-rater reliability (Cohen's Kappa values = 0.520–0.772) and moderate inter-rater reliability (Fleiss Kappa = 0.353). The ABI presented a high concurrent validity ($r = 0.746$); 55.6% of the breathiness vocal deviation can be explained by the acoustic analysis ($r^2 = 0.556$). The ROC curve presented good diagnostic accuracy (85.2%). At a threshold of 2.94, the sensitivity was 75.3% and the specificity was 93.4%.

Conclusion: The ABI is a valid tool for screening and patient's follow-up regarding breathy vocal qualities in the Brazilian Portuguese language.

ARTICLE HISTORY
Received 4 August 2020
Revised 7 November 2020
Accepted 10 December 2020

KEYWORDS
Voice; dysphonia; Acoustic breathiness index; voice quality; auditory-perceptual judgment; acoustic measurement

Introduction

Voice evaluation is common practice in the clinical routine of otolaryngology and speech language pathology [1,2]. This evaluation is usually performed by means of the perceptual judgment and the acoustic analysis [2,3].

Perceptual assessment is subjective as it is related to the listener's perception of the voice quality. It is highly influenced by training, professional experience, and internal standards of the raters [3–8]. The most commonly rated voice parameters are overall severity, roughness, and breathiness [9–11].

Acoustic analysis, on the other hand, is not influenced by the experience and internal standards of the listener. It quantifies degree of vocal deviation on the basis of acoustic signal markers [2,3]. According to recent meta-analysis [12], the perception of breathiness has been related to (all symbolic notations in this text are in accordance with the consensus described in Titze et al.) [13]: natural logarithm of period standard deviation (LNPSD), glottal-to-noise excitation ratio (GNE) 3000 Hz, differences between the amplitude of fundamental frequency peak and first formant 1 (A_1-L_{F1}), relative energy level of high frequency noise (Hfno), cepstral peak prominence (CPP), differences between the amplitudes of the first and second harmonics (A_1-A_2), smoothed cepstral peak prominence (CPPs), smoothed pitch perturbation quotient (sPPQ), harmonics-to-noise ratio (HNR) of Dejonckere, amplitude perturbation quotient-5 (APQ5), normalized noise energy (NNE) 1000–5000 Hz, and smoothed amplitude perturbation quotient (sAPQ). However, these acoustic measurements consider only one parameter of the voice and have lower validity and reliability when compared to multiparametric approaches that are more reliable [14–17]. Some examples of multiparametric approaches are the Dysphonia Severity Index (DSI) [14], the Acoustic Voice Quality Index (AVQI) [15,16], and the Cepstral Spectral Index of Dysphonia (CSID) [17]. These indexes were developed to quantify the more general perceptual dimension of overall severity, and not specific voice characteristics such as breathiness or roughness. Breathiness occurs due to the air leakage during glottal closure and it is commonly analyzed in the evaluation of voice quality; thus, it is included in several voice perceptual protocols [3,10,12] which reinforce the need of its objective analysis [18].

Attempts have been performed to create multivariate models for breathiness between the years of 1980 and 2014, however, they were limited to one single speech task, the

CONTACT Marina Englert marinaenglert@gmail.com Unifesp – "Universidade Federal de São Paulo" and CEV – "Centro de Estudos da Voz", São Paulo, Brazil
*Current address University of Ghent Université Catholique de Louvain (UCL) and University College Ghent (HoGent)

sustained phonation [12]. To reach higher ecological validity, it is better that the voice analysis is not only based on sustained vowel phonations but also incorporates continuous speech. An example of a method comprising both, is the Acoustic Voice Quality Index (AVQI) [15] for quantifying overall hoarseness severity. Analogous to AVQI, the Acoustic Breathiness Index (ABI) [18] was developed to quantify the degree of breathiness and to provide one single score based on both speech task.

With the implementation of continuous speech in these acoustic methods, it is noteworthy that differences among languages may influence their outcomes. Therefore, it is important to investigates the degree of language-specific differences that may occur [19,20]. The AVQI 02.02 has been validated to Finnish [21], Korean [22,23], Japanese [24], and Lithuanian [25]. The AVQI's latest version, 03.01, has already proven to be validated among different languages, Dutch [16], Japanese [26], Spanish [27], German [28], French [29], and Brazilian Portuguese [30]. It is noteworthy that all languages, expect Brazilian Portuguese, consider in the continuous speech the reading of a text, while in the Brazilian Portuguese counting numbers is used [30].

The ABI was first developed and tested for Dutch [18] and later validated in Spanish [27], German [28], Korean [31], and Japanese [32]. To the best of our knowledge, there is only one study with the ABI in Brazilian Portuguese [33]; however, this study aimed to compare the index outcome with two speech materials and not to validate the index to Brazilian Portuguese. Thus, the aim of the present study was to validate the ABI for the Brazilian Portuguese by assessing its concurrent validity and the diagnostic accuracy with a larger sample size than in the previous study.

Materials and methods

This study was approved by the Committee for Ethics in Research. All subjects agreed to participate and signed an informed consent form.

The presented study counted 150 participants: 17 men and 20 women were considered non-dysphonic (mean age = 34.2 years, SD in age = 13.3 years), and 39 men and 74 women were dysphonic patients (mean age = 52.4 years, SD in age = 18.7 years).

Only when participants reported not to have vocal complaints and had a score below 8 on the short version of the Voice Handicap Index (VHI-10), they were considered non-dysphonic. This index is an acceptable classifier for dysphonia; at a threshold of 7.5 it presents a sensitivity of 98.1% and specificity of 100% [34]. Laryngeal pathologies in the dysphonic group was diverse, as illustrated by the diagnoses and their frequency (Table 1).

All voice recordings were performed in a quiet clinical room, with environmental noise below sound level of 50 dBA [35]. An AKG C420 head-mounted condenser microphone, connected with a 174 AKG MPA V L cable to a Focusrite iTrack Solo, was used. All signals were digitized at a sample rate of 44.1 kHz and 16 bits of resolution as well as post-hoc edited with the program Audacity (version 2.0.6).

Table 1. Dysphonia diagnosis.

Diagnosis	*N*	%
Functional dysphonia	24	21.24
Parkinson disease	22	19.47
Nodules	13	11.50
Sulcus vocalis	8	7.08
Paralysis/paresis	8	7.08
Central neurological disorder (expect Parkinson disease)	5	4.42
Cyst	4	3.54
Polypoid mucosa (edema)	4	3.54
Post-laryngeal microsurgery	4	3.54
Spasmodic dysphonia	4	3.54
Acute laryngitis	2	1.77
Head and neck tumor post-surgery	2	1.77
Post-cordectomy	2	1.77
Presbylarynx	2	1.77
Intubation trauma	2	1.77
Vocal tremor	1	0.88
Hemilaryngectomy	1	0.88
Leukoplakia	1	0.88
Polyp	1	0.88
Post-radiotherapy	1	0.88
Reflux laryngitis	1	0.88
Vocal fold scar	1	0.88

The individuals were asked to count from 1 to 20 and to sustain the vowel [a:] at comfortable pitch and loudness.

Selection of the voice sample length

For the ABI analysis to reach a higher validity, the length of all voiced segments of the continuous speech concatenated together ideally equalize the length of the part of interest of the sustained vowel, i.e. 3 s [18,36]. With the Praat-script from Maryn et al. (2010) [15] for automated extraction of voiced segments, the amount of continuous speech that yields 3 s (with tolerance of 0.1 s below or above) of voiced segments was searched and verified. This procedure will be referred to as the customized hand-marked length and has already been used in the development of the AVQI v.03.01 [16].

Usually, the acoustic analysis of the AVQI and the ABI are performed considering the reading of a text. The standardized number of syllables for the reading ranges from 27 to 34 syllables and is the same for the analyses of both indexes [18,27,28]. However, as there is no standardized reading text for voice evaluation in Brazilian Portuguese, and because the general Brazilian population lacks fluent reading, counting numbers has been regarded as the speech task of choice in an earlier study in this language [37]. Optimal length for overall severity evaluation, i.e. equalized with the sustained vowel of 3 s, considering the AVQI in Brazilian Portuguese was found to be at 17 syllables, which corresponds with counting 1–11 [38]. In addition, the Brazilian Portuguese AVQI validation was performed with counting numbers 1–11 [30]. Therefore, the counting numbers recordings were edited from counting 1 to 11 and posteriorly compared with the duration and ABI score of the customized hand-marked length.

Acoustic analysis

The audacity program (version 2.0.6) was used to edit the recordings to contain counting from 1 to 11, and to extract

3 s of the middle portion of the vowel /a/. Thus, voice onset and offsets were avoided [39]. When the maximum phonation time was below 3 s, the entire phonation was considered. The standardized voice sample of counting number from 1 to 11, along with the sustained vowel /a/, was ran in the Praat program (version 6.0.40) using the ABI script [18]. The same procedure was performed using the customized hand-marked cutoff point.

The ABI gives one single score for the breathiness considering nine acoustic parameter: Smoothed cepstral peak prominence (CPPS); Jitter; Glottal to noise excitation ratio in maximum frequency of 4500 Hz, GNEmax-4500 Hz; parameter of relative level of high-frequency noise between energy 0 and 6 kHz and 6 to 10 kHz, Hfno-6000 Hz; harmonics to-noise ratio, HNR-D; difference in 1st and 2nd harmonic, H1-H2; Shim-dB; Shimmer; Period Standard Deviation, PSD, according to the formula [18]:

$$\text{ABI} = (5.0447730916 - [0.172 * \text{CPPs}] - [0.193 * \text{Jit}] - [1.283 * \text{GNEmax} - 4500\ \text{Hz}] - [0.396 * \text{Hfno} - 6000\ \text{Hz}] + [0.01 * \text{HNR} - \text{D}] + [0.017 * \text{H1} - \text{H2}] + [1.473 * \text{Shim} - \text{dB}] - [0.088 * \text{Shim}] - [68.295 * \text{PSD}]) * 2.9257400394$$

Perceptual judgement

Nine voice specialists perceptually rated breathiness considering the continuous speech and the sustained phonation. They had a minimum of 6 years and a maximum of 24 years of clinical experience (mean = 14.22 years of experience, SD = 6.01) in the field of voice disorders. The ratings were performed individually and the recording was presented in the EarPods (Apple) with stereo channels and noise isolation. All raters were blinded to diagnosis and identity of each voice sample.

The B from the GRBAS scale [9] was used. This is considered to be a reliable measure in the evaluation of overall breathiness [40]. The raters were asked to give a single score for the overall breathiness considering both sustained phonation and continuous speech. An ordinal four-point scale was used in which 0 = clear voice/no breathiness, 1 = slightly breathy, 2 = moderately breathy, and 3 = severely breathy. Before this rating task, anchor voices were presented to increase intrarater and interrater reliability [41,42]. The anchor voices had been previously selected from a voice bank by three speech language pathologist who had to agree on the type and degree of vocal deviation. The selected anchor voices represented different degrees of vocal breathiness (no, mild, moderate and severe deviation) from male and female voices. The raters had to listened to the voices before beginning their analysis and could reach out for them whenever they wanted to during the listening session.

Statistical analysis

The statistical analysis was performed using MS-Excel (MS-Office 2013), IBM SPSS (Statistical Package for Social Sciences, version 24.0, SPSS Inc., Chicago, IL) and the RStudio (version 1.1.453, packages R Commander, Rcmdr, version 2.4-4 and R Commander Miscellaneous Functions, RcmdrMisc, version 1.0-10). In all of the statistical tests, the level of significance was set at 5%.

The Wilcoxon signed posttest was used to compare the duration and the ABI results between the customized hand-marked and the standardized cutoff point. Cohen's and Fleiss Kappa coefficient (Ck and Fk) were used to analyze the intrarater and inter-rater reliability, respectively.

The concurrent validity of the ABI was assessed using the Spearman correlation. The diagnostic accuracy was assessed using the receiver operating characteristic curve (ROC curve) where the levels of accuracy classification are: excellent (0.9–1), good (0.8–0.9), fair (0.7–0.8), and poor (0.6–0.7) [34]; a threshold of mean $B > 0.5$ was considered as positive for breathiness in the voice sound. The chance of an individual having breathiness for positive and negative results were calculated with the likelihood ratio for positive and negative results (LR+and LR-, respectively). LR+ $\geq$ 10 and LR$-$ $\leq$ 0.1 are considered acceptable values [15,16].

Results

The duration and the ABI scores were statistically considered to be the same (p-value = 0.948) for the customized hand-marked and the standardized cutoff point (Table 2).

Two of the nine listeners presented insignificant Ck values (0.274 and 0.302); the other seven raters presented Ck ranging from 0.370 to 0.772. For the five most consistent raters, intra-rater B rating reliability ranged from Ck = 0.520 (i.e. fair) to 0.772 (i.e. substantial). With Fk = 0.353, the inter-rater reliability for these five selected raters was fair [43]. Therefore, based on fair to substantial intrarater and reasonable inter-rater consistencies, further analyses were based on average B from these five evaluators.

Figure 1 presents the distribution of the mean breathiness ratings as well as the scatterplots with all ABI and mean B coordinates. Dysphonic patients had voices with more perceived and measured vocal breathiness (Table 3).

With $r = 0.746$, ABI's concurrent validity was high. The coefficient of determination demonstrated that 55.6% of the variance in vocal breathiness ($r^2 = 0.556$).

The ABI presented good diagnostic accuracy for presence of breathiness (85.2%); at a threshold of 2.94, the sensitivity was 75.3% and the specificity 93.4% (Figure 2). The likelihood ratio for positive results was high: LR+ = 11.4.

Table 2. Comparison of the ABI scores and of the duration, in seconds, for the customized and standardized cutoff point.

	Mean	SD	Min	Max	Median	p-value
ABI Score						
Customized	3.09	1.53	0.51	8.71	2.87	0.2828
Standardized	3.07	1.52	0.41	8.68	2.75	
Duration						
Customized	3	0.05	2.91	3.09	3	0.948
Standardized	3.04	0.81	1.07	5.73	2.97	

Wilcoxon signaled posttest. SD: standard deviation.

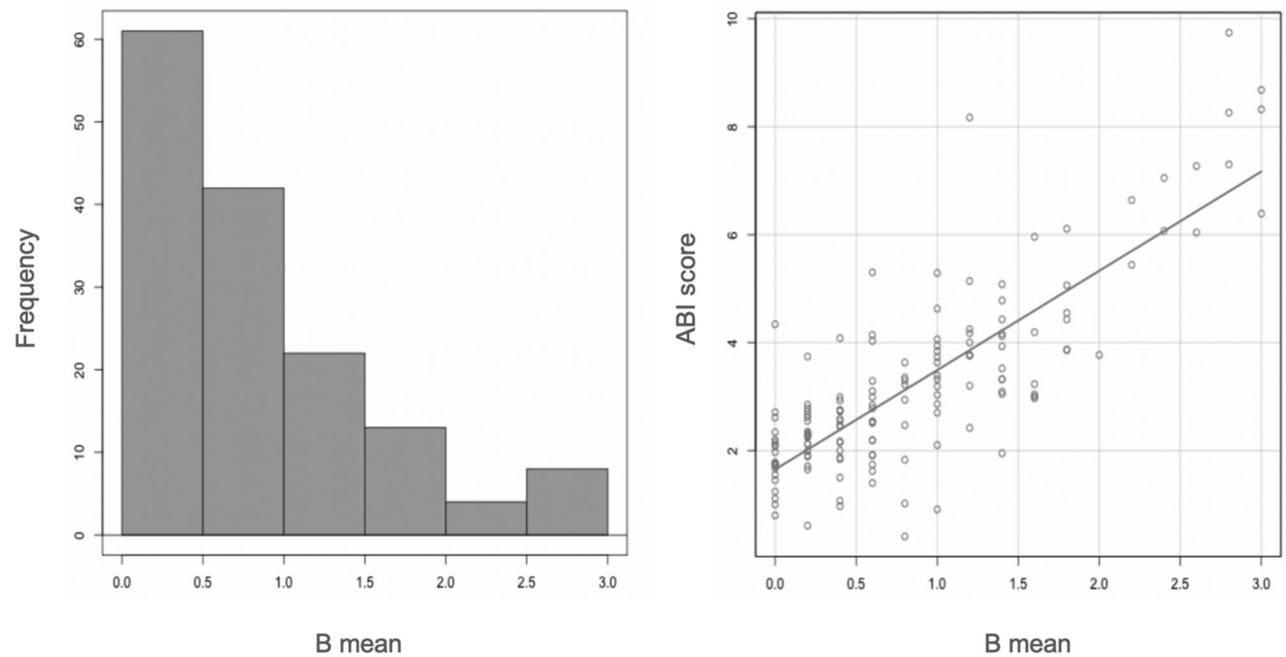

Figure 1. Frequency distribution and Scatterplots of breathiness.

Table 3. Numerical values of the ABI and B mean scores of the non-dysphonic individuals and dysphonic patients.

	Mean	SD	Min	Max	*N*	*p*-value
ABI						
Dysphonic	3.523	1.809	0.41	9.74	113	<0.001
Non-dysphonic	2.26	0.795	0.61	4.14	37	
B						
Dysphonic	1.033	0.762	0	3	113	<0.001
Non-dysphonic	0.286	0.331	0	1.4	37	

Two-sample Wilcoxon test.

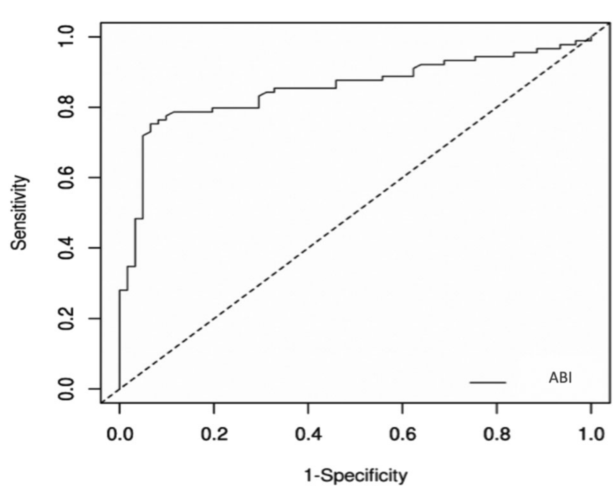

Figure 2. ROC curve.

The likelihood ratio for negative results was respectable: LR- = 0.264 [15,16].

Voice disorders with the highest ABI score were found in subjects with vocal fold paralysis/paresis, post-cordectomy, hemilaryngectomy, and some cases of intubation trauma, post-laryngeal microsurgery, spasmodic dysphonia, Parkinson disease, and functional dysphonia (Figure 3).

Discussion

The ABI has been developed to objectively quantify clinical breathiness ratings, especially in cases with vocal fold paralysis or paresis, vocal nodules, acute laryngitis, and vocal fold bowing [18]. The present study showed a high level of breathiness in cases of vocal fold paralysis/paresis and post-surgery and overall higher level of breathiness for dysphonic patients (Figure 3).

The ABI has already been validated to five different languages: Dutch, German, Spanish, Korean, and Japanese. Overall, ABI showed good concurrent validity and diagnostic accuracy [18,27,28,31,32]. The present Brazilian validation reinforced these findings, with good diagnostic accuracy and similar sensitivity and specificity values than found for other languages (ranging between 72% and 82.4%, and 92.9% to 95.4%, respectively) [18,27,28,31,32].

The LR+ > 10 stands for a high specificity of the ABI, thus, it is quite unlikely that someone with vocal breathiness is identified as having no vocal breathiness. On the other hand, the LR − of 0.264 presents less satisfactory sensitivity, therefore, there might be cases were breathiness is perceptually judge but not measured. This is in accordance with previous Spanish and German validations that found LR − values of 0.27 and 0.3 [27,28].

The ABI threshold for the Brazilian Portuguese was the lowest, 2.94, compared to other validation with values near to 3.4 points. These lower threshold for the Brazilian Portuguese language has also been observed in the studies with AVQI

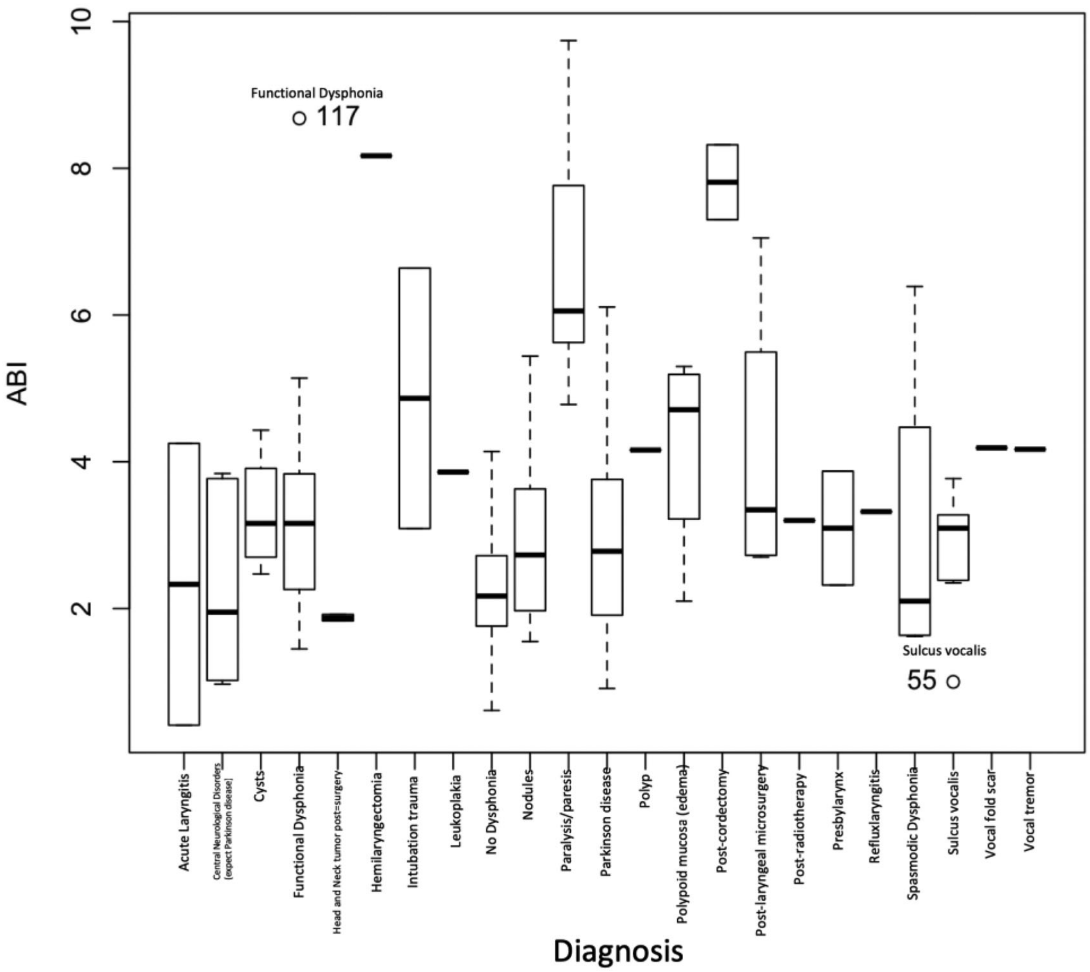

Figure 3. Boxplot for the ABI score for each diagnosis.

[33,37,38] and the previous study with the ABI [33]. Thus, this low threshold might be a language characteristic.

It is known that the vocal characteristics can be influenced by individuals from different cultures and that it can be language dependent [44–46]. The literature states that the Brazilian Portuguese language presents more breathy vocal qualities than the English language [47]. Hence, the amount of breathiness in the voice sound varies among languages. Therefore, it can be hypothesized that Spanish, German, Japanese, Korean, and Dutch have more breathiness than Brazilian Portuguese, which is reflected in higher ABI scores for these languages. Another hypothesis is related to the raters' different culture: the Brazilians raters can be less tolerant to perceived breathiness, which could explain the lower ABI threshold. Thus, there might be language and/or cultural explanations for this lower threshold. These hypotheses should be tested in further studies in order to properly presume if this difference is related to the raters' cultural background or to the Portuguese language.

It is noteworthy that the highest ABI scores were observed for patients with vocal fold paralysis/paresis and considerably higher score were observed for cases post-surgery (i.e. post-cordectomy and post-laryngeal microsurgery). Thus, the ABI seems to be a valid objective tool to assess cases of thyroplasty and to follow-up cases post-phonosurgery, as also observed in the Japanese validation study [32]. The ABI validation in the Brazilian Portuguese emphasize the validity of the index among different languages and prospects its use in the Portuguese speaking population for clinical voice assessment and follow-up.

Conclusion

The ABI is a valid tool in the Brazilian Portuguese language to measure the breathiness in the voice sound considering both continuous speech and sustained phonation. It can be used for clinical follow-up and pre versus post treatment/procedure evaluations, especially regarding diagnosis in which breathiness is expected.

Study conducted at *Universidade Federal de São Paulo* – Unifesp in partnership with the *Centro de Estudos da Voz* – CEV and University of Antwerp.

This study was approved by the Committee for Ethics in Research of the *Universidade Federal de São Paulo* (UNIFESP) under the protocol number 2.106.335. All subjects agreed to participate and signed an informed consent form.

Acknowledgments

The authors would like to acknowledge the patients and volunteers of the voice recording and all the voice specialists who performed the perceptual judgement of the voice quality.

Authorship

ME was responsible for the study design, data collection, audio editing, data analysis and writing of the manuscript; BBL and YM were responsible for the index constriction, study design, data analysis and revision of the manuscript; MB was the advisor, responsible for the study design, data analysis, and final revision of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Marina Englert is a speech language pathologist specializing in voice. Recently defended her PhD in Human Communication Disorders. Received seven awards regarding papers related to her masters and PhD research. Associated researcher at the Centro de Estudos da Voz (CEV).

Ben Barsties V. Latoszek is a speech-language pathologist and doctor of medical sciences (PhD) specializing in voice disorders. Associate professor for speech-language pathology with the emphasis in voice at the SRH University of Applied Health Sciences in Düsseldorf and Bonn, Germany. Research fellow at the department of phoniatrics and pediatric audiology at the University Hospital Münster and University of Münster in Münster, Germany. Works with voice patients at the Dormagen Therapy Centre in Dormagen, Germany.

Youri Maryn works as a clinical speech-language pathologist at the European Institute for ORL-HNS. Teaches on acoustic phonetics at the University of Ghent and on voice disorders at the Ghent University College and Université Catholique de Louvain. Post-doctoral researcher at the University of Antwerp. Serves as board member of the Flemish Association for Speech-Language Therapists. Publishes on voice disorder management and acoustics. Specific topics of interest are clinical voice assessment, voice disorder management, voice/speech acoustic and oral versus nasal speech production.

Mara Behlau is a speech-language pathologist and Ph.D. in Human Communication Disorders. Permanent professor of the Graduate Program in Human Communication Disorders at the Universidade Federal de São Paulo; also Associate Professor of Communication for business at the Insper, São Paulo, Brazil. She is the director of the "Centro de Estudos da Voz" – CEV (Center for Voice Studies), where she coordinates a graduate program to certificate voice specialists.

ORCID

Marina Englert http://orcid.org/0000-0003-0598-8313
Ben Barsties v. Latoszek http://orcid.org/0000-0002-0086-8163
Youri Maryn http://orcid.org/0000-0003-1491-4068
Mara Behlau http://orcid.org/0000-0003-4663-4546

References

[1] Awan SN, Roy N. Toward the development of an objective index of dysphonia severity: a four-factor acoustic model. Clin Linguist Phon. 2006;20(1):35–49.
[2] Roy N, Barkmeier-Kraemer J, Eadie T, et al. Evidence-based clinical voice assessment: a systematic review. Am J Speech Lang Pathol. 2013;22(2):212–226. doi:10.1044/1058-0360(2012/12-0014)
[3] Barsties B, Marc de B. Assessment of voice quality: current state-of-the-art. Auris Nasus Larynx. 2015;42(3):183–188.
[4] Gerratt BR, Kreiman J, Antonanzas-Barroso N, et al. Comparing internal and external standards in voice quality judgments. J Speech Hear Res. 1993;36(1):14–20.
[5] Chan KMK, Yiu EML. The effects of anchors and training on the reliability of perceptual voice evaluation. J Speech Lang Hear Res. 2002;45(1):111–126. doi:10.1044/1092-4388(2002/009)
[6] Yiu EML, Murdoch B, Hird K, et al. Perception of synthesized voice quality in connected speech by Cantonese speakers. J Acoust Soc Am. 2002;112(3 Pt 1):1091–1101.
[7] Sellars C, Stanton AC, Mcconnachie A, et al. Reliability of perceptions of voice quality: evidence from a problem asthma clinic population. J Laryngol Otol. 2009;123(7):755–763.
[8] Englert M, Madazio G, Gielow I, et al. Learning factor influence on the perceptual-auditory analysis. Codas. 2018;30(3):e20170107.
[9] Hirano M. Clinical examination of voice. New York (NY): Springer Verlag; 1981.
[10] Friedrich G, Dejonckere PH. The voice evaluation protocol of the European Laryngological Society (ELS) – first results of a multicenter study. Laryngorhinootologie. 2005;84(10):744–752.
[11] 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(2):124–132. Epub 2008 Oct 16.
[12] Barsties V Latoszek B, Maryn Y, Gerrits E, et al. A meta-analysis: acoustic measurement of roughness and breathiness. J Speech Lang Hear Res. 2018;61(2):298–323.
[13] Titze IR, Baken RJ, Bozeman KW, et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. J Acoust Soc Am. 2015;137(5):3005–3007.
[14] 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(3):796–809.
[15] Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. J Commun Disord. 2010;43(3):161–174.
[16] Barsties B, Maryn Y. External validation of the Acoustic Voice Quality Index Version 03.01 with extended representativity. Annals of otology. Ann Otol Rhinol Laryngol. 2016;125(7):571–583.
[17] Awan SN, Roy N, Zhang D, et al. Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a screening tool for voice disorders: development of clinical cutoff scores. J Voice. 2016;30(2):130–144.
[18] Barsties V Latoszek B, Maryn Y, Gerrits E, et al. The Acoustic Breathiness Index (ABI): a multivariate acoustic model for breathiness. J Voice. 2017;31(4):511.e11–e27.
[19] Maryn Y, De Bodt M, Barsties B, et al. The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. Eur Arch Otorhinolaryngol. 2014;271(6):1609–1619.
[20] Madill C, Nguyen DD, Cham K Y-N, et al. The impact of nasalance on cepstral peak prominence and harmonics-to-noise ratio. Laryngoscope. 2018; 129(8):E299–E304. [Epub ahead of print]
[21] Kankare E, Barsties V., Latoszek B, et al. The acoustic voice quality index version 02.02 in the Finnish-speaking population. Logopedics Phoniatrics Vocol. 2020;45(2):49–58.

[22] Maryn Y, Kim HT, Kim J. Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. J Voice. 2016;30(5):587–594. Epub 2015 Aug 25.

[23] Kim GH, Lee YW, Bae IH, et al. Comparison of two versions of the Acoustic Voice Quality Index for quantification of dysphonia severity. J Voice. 2018; 34(3):489.e11–489. pii: S0892-1997(18)30438-7.

[24] Hosokawa K, Barsties B, Iwahashi T, et al. Validation of the Acoustic Voice Quality Index in the Japanese language. J Voice. 2016;31(2):260.e1–260.e9.

[25] Uloza V, Petrauskas T, Padervinskis E, et al. Validation of the Acoustic Voice Quality Index in the Lithuanian language. J Voice. 2017;31(2):257.e1–257.e11.

[26] Hosokawa K, Barsties V Latoszek B, Iwahashi T, et al. The Acoustic Voice Quality Index Version 03.01 for the Japanese-speaking population. J Voice. 2019;33(1):125–125.e12.

[27] Delgado Hernández J, León Gómez NM, Jiménez A, et al. Validation of the Acoustic Voice Quality Index Version 03.01 and the Acoustic Breathiness Index in the Spanish language. Ann Otol Rhinol Laryngol. 2018;127(5):317–326.

[28] Barsties V Latoszek B, Lehnert B, Janotte B. Validation of the Acoustic Voice Quality Index Version 03.01 and Acoustic Breathiness Index in German. J Voice. 2018;18:30231–30235.

[29] Pommée T, Maryn Y, Finck C, et al. Validation of the Acoustic Voice Quality Index, Version 03.01, in French. J Voice. 2018; pii(18):30514–30517. [Epub ahead of print]

[30] Englert M, Barsties V Latoszek B, Maryn Y, et al. Validation of the Acoustic Voice Quality Index, Version 03.01, to the Brazilian Portuguese language. J Voice. 2019;29:30248-6. [Epub ahead of print]

[31] Kim GH, von Latoszek BB, Lee YW. Validation of Acoustic Voice Quality Index Version 3.01 and Acoustic Breathiness Index in Korean population. J Voice. 2019; S0892-1997(19)30342-X. [in press]

[32] Hosokawa K, von Latoszek BB, Ferrer-Riesgo C, et al. Acoustic Breathiness Index for the Japanese-speaking population: validation study and exploration of affecting factors. J Speech Lang Hear Res. 2019;62(8):2617–2631.

[33] Englert M, Lima L, Behlau M. Acoustic Voice Quality Index and Acoustic Breathiness Index: analysis with different speech material in the Brazilian Portuguese. J Voice. 2019. pii: S0892-1997(19)30036-0. [Epub ahead of print]

[34] Behlau M, Madazio G, Moreti F, et al. Efficiency and cutoff values of self-assessment instruments on the impact of a voice problem. J Voice. 2016;30(4):506.e9–506.e18.

[35] Patel RR, Awan SN, Barkmeier-Kraemer J, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association Expert Panel to develop a protocol for instrumental assessment of vocal function. Am J Speech Lang Pathol. 2018;27(3):887–905.

[36] Barsties B, Maryn Y. The influence of voice sample length in the auditory-perceptual judgment of overall voice quality. J Voice. 2017;31(2):202–210.

[37] Englert M, Lima L, Constantini AC, et al. Acoustic Voice Quality Index- AVQI – for Brazilian Portuguese speakers: analysis of different speech material. CoDAS. 2019;31(1): e20180082.

[38] Englert M, Lima L, Latoszek BBV, et al. Influence of the voice sample length in perceptual and acoustic voice quality analysis. J Voice. 2020;S0892-1997(20)30256-3. Epub ahead of print.

[39] Behlau M. Uma análise das vogais do português brasileiro falado em São Paulo: perceptual, espectrográfica de formantes e computadorizada de frequência fundamental [dissertação]. São Paulo: Universidade Federal de São Paulo; 1984.

[40] Walton C, Carding P, Conway E, et al. Voice outcome measures for adult patients with unilateral vocal fold paralysis: a systematic review. Laryngoscope. 2019;129(1):187–197. Epub 2018 Sep 19.

[41] Rossi M, Pavlovic C, Espesser R. Reducing context effects in the subjective evaluation of speech [abstract]. In 120th Meeting Acoustical Society America. Provence, France: Inst de Phonetique, Univ de Provence; 1990. p. 12.

[42] Awan SN, Lawson LL. The effect of anchor modality on the reliability of vocal severity ratings. J Voice. 2009;23(3):341–345.

[43] Landis R, Koch G. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–174.

[44] Bebout L, Arthur B. Cross-cultural attitudes toward speech disorders. J Speech Hear Res. 1992;35(1):45–52.

[45] Yiu EM, Murdoch B, Hird K, et al. Cultural and language differences in voice quality perception: a preliminary investigation using synthesized signals. Folia Phoniatr Logop. 2008;60(3): 107–119.

[46] Ghio A, Cantarella G, Weisz F, et al. Is the perception of dysphonia severity language-dependent? A comparison of French and Italian voice assessments. Logoped Phoniatr Vocol. 2015; 40(1):36–43.

[47] Engelbert APPF, Denise C, Kluge DC, et al. [Línguas diferentes, vozes distintas: evidências da fala de bilíngues em português e inglês]. R Ilha Do Desterro A J of En Language, Lit in English and Cult Studies. 2016;69(1):033–048. doi:10.5007/2175-8026.2016v69n1p33