We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Cross-Cultural Adaption of the GRBAS and CAPE-V Scales for Portugal and a New Training Programme for Perceptual Voice Evaluation

Luis M.T. Jesus, Ana Inês Tavares and Andreia Hall

Additional information is available at the end of the chapter

Abstract

Several methods have been proposed for the perceptual evaluation of voice quality, but the GRBAS and Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) scales are the most widely used and recommended as part of standardised voice evaluation protocols. In this study, cross-cultural adaptation and translation of the GRBAS (the first translation from the original Japanese version) and CAPE-V scales to European Portuguese were carried out following international guidelines. Results from a study of the intra- and inter-rater reliability of the perceptual evaluation of voices with the GRBAS and CAPE-V scales, before and after a training programme, designed according to the most recent American Speech-Language-Hearing Association and Japan Society of Logopedics and Phoniatrics guidelines, are also reported.

Keywords: voice, voice disorders, perceptual voice evaluation, GRBAS, CAPE-V, American Speech-Language-Hearing Association, simulation training

1. Introduction

Phonation results from the interaction of the vocal folds with the airflow and the air column above them [1, 2]. When air particles pass through the glottis and their speed increases, this reduces the pressure between the vocal folds triggering a suction effect that brings the vocal folds closer to each other, followed by an elastic recoil that promotes a new glottic adduction, thus enabling the production of voice. The sound that results from the vibration of the vocal folds, which is modified by the resonance cavities, is called voice. This audible sound

is the product of a complex relation between the pressure and velocity of expiratory airflow, intensity, the different patterns of abduction and adduction of the vocal folds, the vocal tract configuration and resulting resonances [3].

Voice disorders can have a significant negative impact on a person's life, because the voice is an important tool for communication [4]. There is a voice perturbation whenever the vocal quality, intensity, fundamental frequency (f_0) or vocal flexibility are altered for the age, sex and culture of the speaker ([5], p. 5). Any difficulty or alteration in the vocal emission that prevents the natural production of the voice is called dysphonia. Dysphonia manifests itself through the following changes: Perturbations in vocal quality, emission effort, vocal fatigue, loss of vocal power, uncontrolled variations of fundamental frequency, lack of intensity and projection, loss of vocal efficiency, low vocal resistance and unpleasant sensations during vocal emission. These result in the alteration of one or several acoustic characteristics of the voice.

1.1. Vocal evaluation

Vocal evaluation is considered the first stage of intervention and rehabilitation. Voice assessment has the following main objectives [3]: To know the vocal behaviour of the person, identify the causes of the vocal problem, describe the vocal characteristics of the individual, identify vocal habits, to characterise the relation between body and personality. Perceptual evaluation of voice is routinely used in clinical practice but still poses some inter- and intra-subject problems because it is subjective and often not correlated with the severity of the pathology [6].

According to Chan and Yiu [7], the perceptual evaluation is a subjective process, in which the intra- and inter-rater reliability can vary. Pontes et al. [8] also point out that the perceptual evaluation of vocal quality assumes a subjective character, which varies according to the evaluator, with its internal standards on voice quality, with their perception skills, discrimination and experience with regard to the evaluation of voice. Nevertheless, auditory perception-based assessments can be performed rapidly, are non-invasive, do not require electronic equipment, so results are readily available [9].

Several scales have been proposed for the perceptual evaluation of voice quality [10], but the GRBAS ([11], pp. 181–209; [12], pp. 83–84) and Auditory-Perceptual Evaluation of Voice (CAPE-V) [13–15] are the most widely used [16] and recommended as part of standardised voice evaluation protocols [3, 17–19].

The GRBAS scale ([11], pp. 181–209; [12], pp. 83–84) defines five parameters for vocal quality classification: Grade (G), Rough (R), Breathy (B), Astenic (A), Strain (S). The parameter G corresponds to the grade of alteration of vocal quality; R is psychoacoustic vocal fold vibration irregularity impression, corresponding to the vocal fold vibration irregularity, fluctuation in the value of f_0 and amplitude of the sound of the glottal source; the parameter B refers to the psychoacoustic impression of air passage through the glottis, thus relating to turbulence; the parameter A assesses the weakness or lack of energy in the voice, thus characterising a weak intensity of source sound glottic, or lack of harmonics; finally, the S that characterises the state hyperfunction of phonation. The scale is scored from 0 to 3 for each of its five parameters: 0, normal or absence of hoarseness; 1, slight; 2, moderate; 3, severe.

The CAPE-V scale [13–15] uses six features to evaluate voice quality: Overall severity, roughness, breathiness, strain, pitch, loudness. The parameter *overall severity* captures a global impression of voice disturbance, *roughness* allows clinicians to register source irregularities, the perception of *breathiness* results from air escape, and *strain* is related to the perception of vocal effort. The perceived f_0 and intensity are registered as the *pitch* and *loudness* parameters, respectively. Comments about *resonance* and *additional features*, such as *falsetto* or *tremor*, can also be registered. The scale is scored from 0 to 100 on 100-mm Likert scales for each of its six parameters: Mildly Deviant (MI), Moderately Deviant (MO), and Severely Deviant (SE) qualitative attributes are distributed uniformly along the scale, and the consistent (C) and intermittent (I) labels can be associated with each parameter.

The first European Portuguese (EP) translation [20] of the CAPE-V included six new EP sentences designed to elicit the production of every oral vowel in EP, easy onset with/s/, only voiced phonemes, hard glottal attack, nasal phonemes and voiceless stops ([20], p. 971). The proposed new sentences (thoroughly reviewed by a Speech Scientist, a Linguist and three experienced speech and language therapists (SLTs)) and used in its original form in the study reported in this book chapter were: <A Marta e o avô vivem naquele casarão rosa velho> [ɐ ˈmaɾtɐ i u ɐˈvo ˈvivɐ̃j nɐˈkeli kɐzɐˈɾɐ̃w ˈʁɔzɐ ˈvɛʎu]; <Sofia saiu cedo da sala> [suˈfiɐ sɐˈiw ˈsedu dɐ ˈsalɐ]; <a asa do avião andava avariada> [ɐ ˈazɐ du ɐviˈɐ̃w ɐ̃ˈdavɐ ɐvɐɾiˈadɐ]; <agora é hora de acabar> [ɐˈgɔɾɐ ɛ ˈɔɾɐ di ɐkɐˈbaɾ]; <a minha mãe mandou-me embora> [ɐ ˈmiɲɐ mɐ̃j mɐ̃ˈdomɨ ẽˈbɔɾɐ]; <o Tiago comeu quatro peras> [u tiˈagu kuˈmew kuˈatɾu ˈpeɾɐʃ].

1.2. Auditory-perceptual training of evaluators

The continuous training of the evaluators is recommended, in order to guarantee the reliability and validity of a perceptual evaluation of voice. Both the intra- and inter-rater reliability may vary because perceptual voice assessment is a subjective process, but it is generally accepted that the inter-rater reliability is a greater concern. Kreiman et al. [21, 22] argued that reliability variation can be attributed to the different internal standards acquired by evaluators.

Helou et al. [23] conducted a study with 10 experienced evaluators and 10 inexperienced evaluators, who rated 10 male voices and 10 female voices with CAPE-V. The results revealed that inexperienced evaluators rate voices more severely than the experienced evaluators. Inexperienced evaluators also had lower intra- and inter-rater reliability than those with experience. Experienced evaluators rated the voices similarly.

Studies by Kreiman et al. [21] and Gerratt et al. [24] used natural voice samples and/or voice samples synthesised as anchors, and showed that inter- and intra-rater reliability in the perceptual assessment of voice improved with training.

Anchors are considered references that listeners (evaluators) can use to compare with the signals they are invited to judge [7]. In the study by Eadie and Smith [25], 20 inexperienced and 20 experienced evaluators rated 20 samples of normal voices. The results of this study showed that the anchors reduce inter-rater variability.

Silva et al. [26] analysed the impact of auditory-perceptual training on the evaluation of voice performed by speech and language therapy students. Seventeen students analysed samples of normal and dysphonic voices with the GRBAS scale. All students had auditory training during a total of nine weekly sessions, each about 15 min long. The evaluation of voice samples was performed before and after the training, and in four other moments during the meetings. Student ratings were compared with an assessment by three voice specialists. The results showed that the students' success rate at the pre-training moment was considered between regular and good. A maintenance of the number of hits throughout the evaluations performed, for most of the scale parameters, was also observed. Regarding the post-training moment, a better analysis was observed, mainly for the Astenic (A) parameter.

Training judges/listeners has been shown to 'increase the extent to which they share common standards for different' ([10], p. 63) voice qualities, so the current study also includes the intra- and inter-rater reliability analysis of the perceptual evaluation of voices with the GRBAS and CAPE-V scales, before and after a training programme.

1.3. Cross-cultural adaptation and translation of the GRBAS and CAPE-V scales

To the best of our knowledge, there are no standard assessment instruments to perceptually evaluate voice quality in EP, so clinicians in Portugal use various translations of GRBAS ([27, 28], pp. 66–69) and CAPE-V [29], and generally have no access to EP versions of the original instructions published by the Japan Society of Logopedics and Phoniatrics ([11], pp. 181–209) or the American Speech-Language-Hearing Association [13]. They therefore use various procedures and non-standardised definitions of the parameters [30–32]. This results in different voice assessment methods hamper the development, objectivity and specificity of therapeutic plans, thus compromising the efficiency and efficacy of intervention strategies.

We believe the access to the original author's definitions of core concepts behind the development of health instruments, contributes towards the standardisation of evaluations procedures, and considerable improvements in intra- and inter-rater reliability. The translation and adaptation processes of the whole tool (not just the score sheets) should follow international guidelines [33] for cross-cultural adaptation of health assessment instruments. Evidence-based practice would thus be enhanced, and comparisons across countries would be facilitated. A broader evidence base for effective service delivery planning based on results from large-scale randomised controlled trials requires that the same assessment instruments are validated in different cultures.

Cross-cultural adaptation of instruments is necessary when the new target population differs from the original in which the assessment tool is used regarding culture or cultural background, country and language. There are specific guidelines [34, 35] to conserve the sensibility of the assessment tool in the original culture [36]. The steps that must be followed, if relevant to the specific assessment tool, are [37, 38]: translation, synthesis of the translations, back-translation, committee review and pre-testing.

The first stage of a cross-cultural adaptation must be the production of several translations by, at least, two independent translators. In a second stage, the two translators synthesise

the results of the translations, producing one common translation [37]. It is then necessary to back-translate the assessment tool (third stage), which means translating back from the final language into the source language, producing as many back-translations as translations, based on the synthesised translation [36]. In the fourth stage, an expert committee compares the source and the final version. The fifth stage consists of a cognitive debriefing that tests alternative wording, understandability, and interpretation of the translation [35].

In this study, the cross-cultural adaptation and translation of the GRBAS and CAPE-V scales to EP were carried out following these international guidelines [33].

2. Method

Ethical approval was obtained from all authorities required by Portuguese bylaws for clinical research: National data protection committee, independent ethics committees. Informed consent was collected from all participants prior to any data collection.

2.1. Cross-cultural adaptation and translation of the GRBAS and CAPE-V scales

The American Speech-Language-Hearing Association (ASHA) Special Interest Division 3 was contacted by the first author of this book chapter, and the 'Request to Translate & Distribute' CAPE-V was granted to the University of Aveiro by ASHA on 28 January 2008. The Portuguese version of the CAPE-V has been available from the Advanced Communication and Swallowing Assessment (ACSA) http://acsa.web.ua.pt/platform since 2010.

At a later stage, the same author (Luis M.T. Jesus) contacted the Japan Society of Logopedics and Phoniatrics (JSLP), and the University of Aveiro was granted permission to complete a new translation and adaptation of the GRBAS on the 10 January 2013. Professor Eiji Yumoto, on behalf of the JSLP, concluded that there were no detailed descriptions of the GRBAS scale in other languages besides Japanese, so he generously granted the University of Aveiro access to a detailed description of the GRBAS scale, published in 1979 by the JSLP. This book [11] was written in Japanese, and Professor Minoru Hirano described the scale only briefly in his publication ([12], pp. 83–84).

The work reported in this book chapter reports work that is part of the validation process of a voice evaluation protocol developed by our research team [18–20] and freely available from the ACSA platform. University of Aveiro's Voice Evaluation Protocol [18] includes the assessment of voice quality, glottal attack, respiratory support, respiratory-phonatary-articulatory coordination, digital laryngeal manipulation (laryngeal crepitation) and laryngeal tension. It also allows the self-assessment of voice quality and instrumental evaluation. Results from various instrumental evaluation techniques (videostroboscopy, aerodynamics, electroglottography and electromyography) can be registered by the protocol, including an extensive acoustic analysis based on sustained productions of /a, i, u, ɔ/, CAPE-V sentences and reading a passage. The complete protocol provides data to test different methods applied to voice function assessment. The focus of our research was performance improvement of assessment methods used by voice clinicians.

2.2. Training programme

The user interface, terminology, audio and video samples of the GRBAS CD and DVD developed by JSLP [39] were used as a standard reference to design (according to the most recent ASHA and JSLP guidelines), the training programme described in the subsequent text.

Forty-five EP speakers from the Advanced Voice Function Assessment Databases (AVFAD) (see Jesus et al. [40] in this book for a detailed description) were used as auditory stimuli. Fifteen participants were selected for anchors, 15 for training and 15 for evaluation.

The voices from the AVFAD were selected based on the auditory perception in a quiet room by a speech and language therapist (the second author of this book chapter) using *VLC media player 2.1.3. rincewind* running on a laptop connected to a pair of *NGS 2.1* loudspeakers. The same speech and language therapist classified the representative samples (anchors) of all selected voices with the GRBAS and CAPE scales.

Judges listened to sustained productions of vowels/a, i/played in a quiet environment with the same volume through a pair of *Sennheiser HD380Pro* headphones connected to the internal soundcard of a laptop computer.

Ten female speech and language therapists were asked to rate the severity of dysphonic voice stimuli using the GRBAS and CAPE-V scales. Each judge first evaluated 15 voices without any training and then went through a training programme based on two 1-h sessions.

During the first session, judges read detailed written instructions and the original description of the scale, and then, classified anchor voices that included several grades of severity for vocal quality. During the second session, a new set of voices (training voices) were classified, and judges could listen to one anchor for each five voices that were classified. At the end of the second session, all judges had access to a feedback document, but they could not change their classifications. One week later, the same judges classified a new set of 15 voices.

IBM SPSS Statistics 23 was used to calculate the Intraclass Correlation Coefficient (ICC) of responses and to run a one-way analysis of variance (ANOVA) with repeated measures.

3. Results

3.1. Cross-cultural adaptation and translation of the GRBAS and CAPE-V scales

Two EP translations of the original American English CAPE-V scoresheet and instructions were produced by two independent translators. This led to the detection of errors and divergent interpretations of ambiguous items in the original tool [33]. The translators were fluent in both languages (with the target language as their mother tongue), knowledgeable of the two cultures, and experts in the content measured by the instrument (they were both SLTs). Then, both translators synthesised the results of the translations, producing one common translation, which was used to back-translate the assessment tool, producing as many back-translations as translations (two). An expert committee compared the source and the final versions and produced a pre-final version for field testing, based on all translations and back-translations.

During CAPE-V's cognitive debriefing (final stage of the cross-cultural adaptation and translation), alternative wording and interpretation of the translation were tested in five clients with voice pathology by the second author of this book chapter (a SLT), at the University of Algarve. Finally, the translation was revised taking into account the feedback obtained.

Hirano's ([12], pp. 83–84) GRBAS description was translated by the same group of experts involved in CAPE-V's cross-cultural adaption, using exactly the same process and stages [33].

A professional Japanese translator (Tomoko Suga) certified by the Japanese embassy in Lisbon (Portugal) translated the original Takahashi ([11], pp. 181–209) instructions. The translator was fluent in both languages (Japanese and EP) and knowledgeable of the two cultures. Since the original Japanese used by Takahashi ([11], pp. 181–209) is quite different, in some respects, from what is used nowadays and scientific terminology has changed considerably, only the core descriptions by Takahashi ([11], pp. 181–209) were retained in the Portuguese version, and the translation had to be thoroughly revised by an expert committee that included the original Japanese translator, the first author of this book chapter and an SLT blind to the purposes of the study.

The same group of five clients with voice pathology recruited for CAPE-V's cognitive debriefing was involved in Takahashi's ([11], pp. 181–209) and Hirano's ([12], pp. 83–84) GRBAS analysis of the level of comprehensibility of the instructions and the final items, cognitive equivalence of the translation, translation alternatives and items that were eventually inappropriate or confusing.

CAPE-V's cognitive debriefing results showed no inconsistencies but GRBAS' instructions analysis revealed that the number of vowels required by later protocol is not the same as currently suggested by ASHA [13] or in most recent voice assessment procedure based on sustained vowels.

According to Takahashi ([11], pp. 181–209), clinicians should perceptually evaluate five sustained vowels [a, ɛ, i, ɔ, u] and choose to register, on a table that is part of the score sheet, the one that they attribute the highest GRBAS parameters scores.

The use of sustained vowels usually results in articulatory stability and allows the clinician to focus on the typology a voice source signal that is more regular and stable than in connected speech, facilitating the perceptual assessment of voice quality [41, 42].

Given the fact that the CAPE-V protocol [13] proposes [a, i] as the sustained vowels to be used during assessment, and that both GRBAS and CAPE-V are used by the University of Aveiro's Voice Evaluation Protocol [18], following cognitive debriefing, these were proposed as the basis of perceptual evaluation. According to ASHA ([13], p. 3), [i] is used because it is the only sound speakers that can produce during laryngeal videoendoscopy, and /a/ is used because it differs from /i/ in terms of its degree of tenseness: /a/ is a lax vowel and /i/ is a tense vowel in most English dialects. Portuguese does not have the tense-lax contrast, but still, the close /i/ versus open /a/ distinction could be used to monitor the effect of an enlarged pharyngeal cavity for close vowels ([43], p. 627) and the lowering of laryngeal structures ([43], p. 633) for open vowels.

The translations were revised taking into account the results from the cognitive debriefing process, and the assessment tools can now be administered to a representative sample of the

population. Both Takahashi's ([11], pp. 181–209) and Hirano's ([12], pp. 83–84) instructions in Portuguese are now included in the University of Aveiro's Voice Evaluation Protocol [18] manual and available from the ACSA platform.

3.2. Training programme

The GRBAS CD and DVD [39] were thoroughly analysed (see **Figure 1**); terminology, audio and video samples therein were used as a standard reference to design the training programme (described earlier).

This resulted in a first prototype of a *PowerPoint* presentation, shown in **Figure 2**, based on Japanese audio and video samples [39] that guided the final *PowerPoint* presentation design that then used audio samples from the Portuguese AVFAD [40].

The anchors were available during the training programme on a *PowerPoint 2010* presentation with a total of 15 slides formatted as shown in **Figure 3**.

The ICC is a measure of inter-rater reliability that describes the similarity between the responses observed within a given set. The ICC value varies between 0 and 1, the closer to 1, the more consistent are the results.

The ICC mean of 10 raters for the parameters of the GRBAS scale is presented in **Table 1**.

All ICC values (pre- and post-training) are very high which indicates a good agreement between judges. We could only observe a post-training increase of ICC values for the *Strain* parameter which suggests that the training programme was not very effective. However, since the pre-training values are already very high, it is harder to observe an increase of the values after training as a result of the training programme. The *Strain* parameter is the only one with a mean ICC value below 0.900 pre-training, so a possible cause for the difficulty in observing the expected effect of training in the other parameters could be related to the fact that the ICC values pre-training are above a certain threshold. Still, even with evaluator distinct pre-training standards (different GRBAS parameter values), all changed classifications post-training.
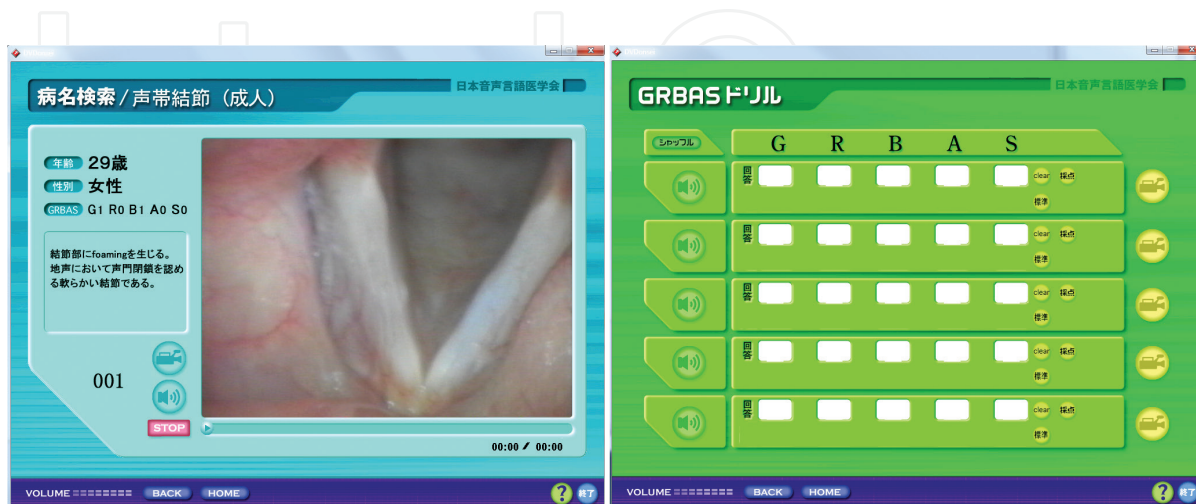

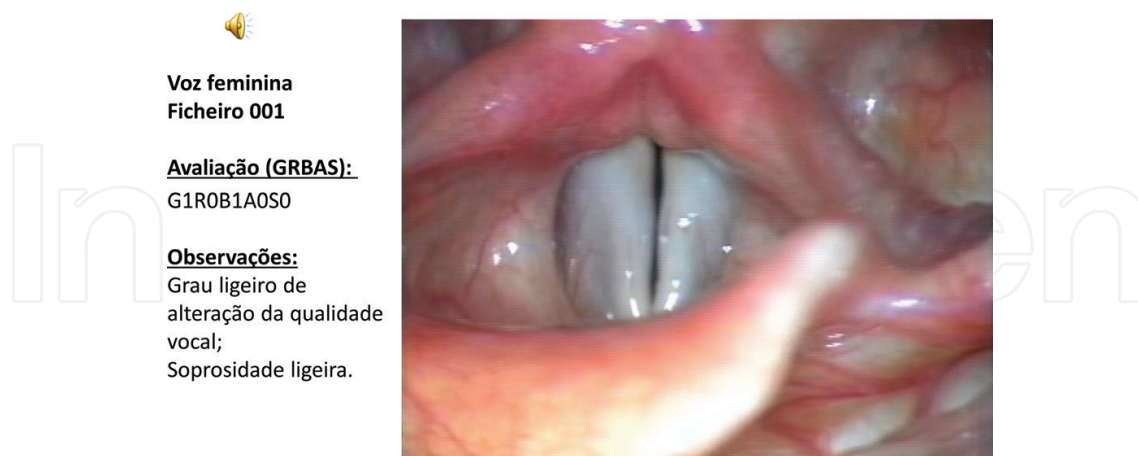**Figure 1.** Case (left) and drill (right) user interfaces from the GRBAS DVD [39].

Figure 2. GRBAS *PowerPoint* prototype presentation using samples from JSLP [39].

Figure 3. Screenshot from the *PowerPoint* presentation used to run the training programme. Samples (anchors) of female voices with GRBAS and CAPE-V classifications are shown.

	Pre-training	Post-training
G	0.963	0.959
R	0.949	0.947
B	0.907	0.886
A	0.937	0.832
S	0.886	0.907

Table 1. ICC for the parameters of the GRBAS scale pre- and post-training.

	Time	Judge	Time × Judge
G	0.000	0.008	0.000
R	0.000	0.598	0.000
B	0.001	0.069	0.000
A	0.000	0.000	0.001
S	0.000	0.002	0.006

Table 2. p-Values of the repeated measure ANOVA for the parameters of the GRBAS scale with the judges as between subject factor and time as within subjects factor.

We also ran a one-way ANOVA with repeated measures for the GRBAS scale. This analysis allowed us to test if the evaluators changed the classifications as a function of time (pre- to post-training) and how this change relates to possible differences between them. The results are shown in **Table 2**.

From **Table 2**, it can be seen that time is a significant effect for all parameters and the interaction also. As for the differences between raters, as a main effect, only for the *Rough* and the *Breathy* parameters, these are not significant.

The ICC mean of 10 raters for the parameters of the CAPE-V scale is presented in **Table 3**.

	Pre-training	Post-training
Overall severity	0.975	0.970
Roughness	0.961	0.942
Breathiness	0.899	0.946
Strain	0.926	0.916
Pitch	0.895	0.788
Loudness	0.913	0.782

Table 3. ICC for the parameters of the CAPE-V scale pre- and post-training.

Similar to what was observed for the GRBAS scale, all ICC values (pre- and post-training) are quite high which indicates a good agreement between judges. We could only observe a post-training increase of ICC values for the *Breathiness* parameter which suggests that the training programme was not very effective. However, when analysing the CAPE-V parameter values, all evaluators changed in the same direction from pre- to post-training, that is, all 10 evaluators either presented higher or lower values post-training for a specific parameter.

We also ran a one-way ANOVA with repeated measures for the CAPE-V scale. The results are shown in **Table 4**.

From **Table 4**, it can be seen that both the effects and their interaction are significant for all the parameters.

	Time	Judge	Time × Judge
Overall severity	0.000	0.027	0.000
Roughness	0.000	0.022	0.000
Breathiness	0.000	0.000	0.000
Strain	0.000	0.000	0.000
Pitch	0.042	0.004	0.021
Loudness	0.022	0.001	0.015

Table 4. *p*-Values of the repeated measure ANOVA for the parameters of the CAPE-V scale with the judges as between subject factor and time as within subjects factor.

4. Conclusions

One of the major contributions of this work was the development of the first non-Japanese version of the original manual of the GRBAS scale and the first Portuguese version of the detailed design considerations, description and instructions of CAPE-V. This research followed international guidelines for the translation and cultural adaptation of health assessment tools.

The GRBAS and CAPE-V scales are now part of the following comprehensive and unique set of resources developed for clinicians at the University of Aveiro in Portugal: A standardised voice case history form [44, 45]; a voice evaluation protocol [18]; a reference voice database [40]. All of these are freely available from the ACSA platform.

The manuals developed during this project had a crucial impact on the training of judges. The ICC values were generally very high, which could be the result of the written instructions and detailed description of the scales, which is a possible cause for the small training effect. The definition of the *Breathiness* parameter benefited particularly from the availability of these instructions. Problems related to the use of the Portuguese term for *Grade* 'grau de rouquidão' being erroneously interpreted as the CAPE-V term 'rouquidão' (*Roughness*), as previously reported by Jesus et al. ([16], p. 62), have been circumvented by the manual, training and samples of voices that represent specified grades of severity.

We also ran a one-way ANOVA with repeated measures for the GRBAS and CAPE-V scales. This analysis allowed us to test if the evaluators changed the classifications as a function of time (pre- to post-training) and how this change relates to possible differences between them.

Regarding the analysis of variance, taking into account the time factor as the main object of study, results showed pre- to post-training differences. The evaluators had individual and distinct standards, and changed the classifications, allowing us to conclude that their internal standards have been modified.

Increasing the level of experience of the evaluators, or the number of training sessions, could have contributed to reducing the variability of the results.

Acknowledgements

Author details

Luis M.T. Jesus[1]*, Ana Inês Tavares[2] and Andreia Hall[1]

*Address all correspondence to: lmtj@ua.pt

1 University of Aveiro, Aveiro, Portugal

2 University of Algarve, Algarve, Portugal

References

[1] Titze I, Alipour F. The Myoelastic Aerodynamic Theory of Phonation. Denver: The National Center for Voice and Speech; 2006

[2] Seikel J, King D, Drumright D. Anatomy & Physiology for Speech, Language, and Hearing. 4th ed. Clifton Park. Delmar; 2010

[3] Barsties B, Bodt M. Assessment of voice quality: Current state-of-the-art. Auris Nasus Larynx. 2015;**42**:183-188

[4] Brinca L, Pinto P, Tavares A, Batista A, Gonçalves I, Moreno M. The prevalence of laryngeal pathologies in an academic population. Journal of Voice. 2015;**29**(1):130.e1-130.e9

[5] Aronson A, Bless D. Clinical Voice Disorders. New York, NY: Thieme; 2011

[6] Oguz H, Demirci M, Safak M, Arslan N, Islam A, Kargin S. Effects of unilateral vocal cord paralysis on objective voice measures obtained by Praat. European Archives of Otorhinolaryngology. 2007;**264**:257-261

[7] Chan K, Yiu E. The effect of anchors and training on the reliability of perceptual voice evaluation. Journal of Speech, Language, and Hearing Research. 2002;**45**:111-126

[8] Pontes P, Vieira V, Gonçalves M, Pontes A. Características das vozes roucas, ásperas e normais: análise acústica espectográfica comparativa. Revista Brasileira de Otorrinolaringologia. 2002;**68**(2):182-188

[9] Webb A, Carding P, Deary I, MacKenzie K, Steen N, Wilson J. The reliability of three perceptual evaluation scales for dysphonia. European Archives of Oto-Rhino-Laryngology. 2004;**261**:429-434

[10] Kreiman J, Gerratt B. Perceptual assessment of voice quality: Past, present, and future. Perspectives on Voice and Voice Disorders. 2010;**20**:62-67

[11] Takahashi H. Assessment of auditory impression of dysphonia. In: The Japan Society of Logopedics and Phoniatrics, editor. Voice Examination Methods. Tokyo: Interna; 1979

[12] Hirano M. Acoustic analysis of the voice signal. In: Clinical Examination of Voice. New York, NY: Springer-Verlag; 1981

[13] ASHA. CAPE-V Form and Procedures. American Speech-Language-Hearing Association (ASHA) Special Interest Division 3, Voice and Voice Disorders; 2006

[14] Kempster G, Gerratt B, Abbott K, Kraemer J, Hillman R. Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. American Journal of Speech-Language Pathology. 2009;**18**:124-132

[15] Zraick R, Kempster G, Connor N, Thibeault S, Klaben B, Bursac Z, Thrush C, Glaze L. Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). American Journal of Speech-Language Pathology. 2011;**20**(1):14-22

[16] Jesus L, Barney A, Couto P, Vilarinho H, Correia A. Voice quality evaluation using CAPE-V and GRBAS in European Portuguese. In: Proceedings of the 6th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA 2009); Florence, Italy. 2009. pp. 61-64

[17] Dejonckere P, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Woisard V. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngolo. European Archives of Oto-Rhino-Laryngology. 2001;**258**(2):77-82

[18] Jesus L, Vilarinho H, Santos R. Protocolo de Avaliação da Qualidade Vocal da Universidade de Aveiro [University of Aveiro's Voice Evaluation Protocol]. University of Aveiro, Portugal.

2010-2015. INPI Registration Number 463522 and IGAC Registration 27/05/2010 (Version 2 IGAC Registration Number 355/2012 and Version 2.3.6 IGAC Registration Number 3474/2015). Available from: Advanced Communication and Swallowing Assessment (ACSA) http://acsa.web.ua.pt/

[19] Alves M. Validade e Fiabilidade do Protocolo de Avaliação da Qualidade Vocal da Universidade de Aveiro [Validity and reliability of University of Aveiro's Voice Evaluation Protocol] [M.Sc. thesis]. Mestrado em Ciências da Fala e da Audição [M.Sc. in Speech and Hearing Sciences]. Portugal: University of Aveiro; 2015

[20] Jesus L, Barney A, Santos R, Caetano J, Jorge J, Couto P. Universidade de Aveiro's voice evaluation protocol. In: Proceedings of Interspeech; Brighton, UK. 2009. pp. 971-974

[21] Kreiman J, Gerratt B, Precoda K, Berke G. Individual differences in voice quality perception. Journal of Speech and Hearing Research. 1992;**35**:512-520

[22] Kreiman J, Gerratt B, Kempster G, Erman A, Berke G. Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. Journal of Speech and Hearing Research. 1993;**36**(1):21-40

[23] Helou L, Solomon N, Henry L, Coppit G, Howard R, Stojadinovic A. The role of listener experience on Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) ratings of post-thyroidectomy voice. American Journal of Speech-Language Pathology. 2010;**19**:248-258

[24] Gerratt B, Kreiman J, Barroso N, Berke G. Comparing internal and external standards in voice quality judgments. Journal of Speech and Hearing Research. 1993;**36**(1):14-20

[25] Eadie T, Smith K. The effect of listener experience and anchors on judgments of dysphonia. Journal of Speech, Language, and Hearing Research. 2011;**54**:430-447

[26] Silva R, Zenari M, Nemr K. Impacto de treinamento auditivo na avaliação perceptivo-auditiva da voz realizada por estudantes de Fonoaudiologia. Revista da Sociedade Brasileira de Fonoaudiologia. 2012;**24**(1):19-25

[27] Pinho S, Pontes P. Escala de Avaliação Perceptiva da Fonte Glótica: RASAT. São Paulo: INVOZ; 2002

[28] Pinho S, Pontes P. Músculos Intrínsecos da Laringe e Dinâmica Vocal. Série Desvendando os Segredos da Voz Vol. 1. Rio de Janeiro: Revinter; 2008

[29] Behlau M. Consensus auditory-perceptual evaluation of voice (CAPE-V), ASHA 2003. Revista da Sociedade Brasileira de Fonoaudiologia. 2004;**9**(3):187-189

[30] Silvestre I, Guimarães I, Teixeira A. Voice quality on women with vocal nodules diagnosis: Preliminary study. Otorrinolaringologia e Cirurgia Cérvico-Facial (Revista Portuguesa). 2011;**49**(2):69-77

[31] Freitas S, Pestana P, Almeida V, Ferreira A. Audio-perceptual evaluation of Portuguese voice disorders—An inter- and intrajudge reliability study. Journal of Voice. 2014;**28**(2):210-215

[32] Brinca L, Batista A, Tavares A, Pinto P, Araújo L. The effect of anchors and training on the reliability of voice quality ratings for different types of speech stimuli. Journal of Voice. 2015;**29**(6):776.e7-776.e14

[33] Valente A, Jesus L, Roberto M, Leahy M, St. Louis K. Cross-cultural adaptation of the Public Opinion Survey of Human Attributes (POSHA-S): The European Portuguese (EP) translation and adaptation process as a model. In: Santiago D, editor. Stuttering: Risk Factors, Public Attitudes and Impact on Psychological Well-Being. New York, NY: Nova Science; 2015. pp. 93-110

[34] Hutchinson A, Bentzen N, Zahn C, editors. Cross Cultural Health Outcome Assessment: A User's Guide. Ruinen: European Research Group on Health Outcomes; 1997

[35] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Lorenz A, Erikson P. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health. 2005;**8**(2):94-104

[36] Geisinger K. Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment. 1994;**6**(4):304-312

[37] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. Spine. 2000;**25**(24):3186-3191

[38] Guillemin F, Bombardier C, Beaton DE. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. Journal of Clinical Epidemiology. 1993;**46**(12):1417-1432

[39] JSLP. GRBAS CD and DVD. The Japan Society of Logopedics and Phoniatrics (JSLP), editor. Tokyo: Interuna Publishers; 2005

[40] Jesus L, Belo I, Machado J, Hall A. The Advanced Voice Function Assessment Databases (AVFAD): Tools for voice clinicians and speech research. In: Fernandes F, editor. Advances in Speech-language Pathology. Rijeka: InTech; 2017 pp.

[41] Revis J, Giovanni A, Wuyts F, Triglia J. Comparison of different voice samples for perceptual analysis. Folia Phoniatrica et Logopaedica. 1999;**51**(3):108-116

[42] Morsomme D, Jamart J, Wéry C, Giovanni A, Remacle M. Comparison between the GIRBAS scale and the acoustic and aerodynamic measures provided by EVA for the assessment of dysphonia following unilateral vocal fold paralysis. Folia Phoniatrica et Logopaedica. 2001;**53**(6):317-325

[43] Pinho C, Jesus L, Barney A. Weak voicing in fricative production. Journal of Phonetics. 2012;**40**(5):625-638

[44] Jesus L, Vilarinho H, Ferreira M. Protocolo de Anamnese Vocal (Jovens e Adultos) da Universidade de Aveiro [University of Aveiro's Standard Voice Case History Form]. University of Aveiro, Portugal. 2010. INPI Registration Number 463523 and IGAC

Registration 27/05/2010. Available from: Advanced Communication and Swallowing Assessment (ACSA) http://acsa.web.ua.pt/

[45] Ferreira M, Jesus L, Couto P, Vilarinho H. University of Aveiro's Standardised Voice Case History Form. Revista de Saúde Pública 48 (Special Number); 2014. p. 297