Research Paper: The Relationship Between Auditory Perceptual Evaluation and Acoustic Measurements of the Voice in Dysphonia: Some Issues About the Task Effect on Perceptual Rating

Seyyedeh Maryam Khoddami[1*], Nazila Salary Majd[1]

1. Department of Speech Therapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.

Use your device to scan and read the article online

Citation Khoddami SM, Salary Majd N. The Relationship Between Auditory Perceptual Evaluation and Acoustic Measurements of the Voice in Dysphonia: Some Issues About the Task Effect on Perceptual Rating. Journal of Modern Rehabilitation. 2018; 12(2):113-122.

doi:

Funding: See Page 120

Article info:
Received: 25 Sep 2017
Accepted: 10 Jan 2018
Available Online: 01 Apr 2018

ABSTRACT

Introduction: This study aimed to investigate construct convergent validity of the Persian version of consensus auditory perceptual evaluation of voice (called ATSHA), using the acoustic measurements. Moreover, the effect of voice tasks on the perceptual ratings was studied.

Materials and Methods: The study data were gathered from a total of 40 dysphonic patients (Mean±SD age=36.79±8.26 years). Perceptual voice evaluation was performed using the ATSHA during sustained vowels and sentence reading tasks. The acoustic features including fundamental frequency (F0), intensity, jitter, shimmer, and Harmonics-to-Noise Ratio (HNR) were extracted using *Praat* application. To assess construct validity of ATSHA, correlation between perceptual and acoustic measures were studied. The effect of tasks was investigated by mean comparison and Pearson correlation.

Results: The results demonstrate that ATSHA has significant correlation with all acoustic measures except the frequency ($r=-0.08-0.35$; $P\geq0.05$). There was no significant correlation between pitch and the acoustic measures of intensity and jitter ($r=-0.31$; $P=0.05$ and $r=0.24$; $P=0.14$, respectively). The highest correlation observed between the overall severity and the HNR ($r=-0.85$; $P<0.001$). The correlation between the perceptual scores in both tasks was high ($r=0.82-0.99$, $P<0.05$).

Conclusion: The ATSHA is a valid scale for perceptual judgment on intensity, jitter, shimmer, and HNR. However, this scale could not estimate the frequency of voice in dysphonia. The current study demonstrate that vowel prolongation and sentence reading has no noticeable effect on the perceptual ratings in dysphonia.

Keywords:
Auditory perception, Acoustic, Voice, Dysphonia

*** Corresponding Author:**
Seyyedeh Maryam Khoddami, PhD.
Address: *Department of Speech Therapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.*
Tel: *+98 (21) 77636042*
E-mail: *khoddami@tums.ac.ir*

1. Introduction

The voice is a multidimensional phenomenon [1, 2], and voice problems is usually due to multifactorial etiologies [3]. Although several subjective and objective methods have been proposed for the voice assessment [4, 5], there is no standard measure for the voice function [6]. The auditory-perceptual assessment is the most common subjective method which describes voice parameters just by listener's judgment [7, 8]. However, acoustic analysis is an objective issue that measures the same vocal parameters by dedicated instruments or integrated application [3]. A minimum voice assessment includes perceptual judgment and acoustic measures of voice parameters [9].

There are many reports about advantages and disadvantages of both perceptual judgment and acoustic measurement. The main advantage of perceptual judgment is due to its easy to use nature [10] that makes it the most common method in the voice clinical assessment [7, 8, 10]. However, there are some concerns about its subjective nature and reliability [10, 11]. Acoustic measures are the objective counterparts of perceptual voice assessment and sometimes can support them [12-15]. There are some limitations about using the acoustic measurements in the assessment of voice. There is no consensus among specialists for identifying the most sensitive instrument in the acoustic measurements [16]. Moreover, there are some problems to use them because of the nature of the acoustic measurements. In particular, the acoustic measurement needs especial instruments (sometimes expensive ones), calibration, and proper education in order to use them accurately. It seems that none of the perceptual and acoustic assessment methods are perfect, but together they can complement each other's shortcomings [10].

As mentioned before, the main advantage of perceptual voice evaluations has made them the most common method in the voice clinics. However, Speech-Language Pathologists (SLPs) and ENT specialists usually document their perceptual judgments using rating scales to minimize some limitations of the perceptual assessment. Up to now, several voice rating scales have been developed which GRBAS is the most popular one [17]. Completed version of the GRBAS is the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) that was developed in a consensus meeting in 2002. CAPE-V has 6 items to describe voice, including overall severity, roughness, breathiness, strain, pitch, and loudness in different vowels and speech tasks [18]. It is a perceptually valid and reliable scale that follows detailed protocols for voice sampling and recording [10, 18]. Moreover, several documents confirm its more sensitivity and reliability over the GRBAS [19, 20].

Because of all above-mentioned advantages, CAPE-V was selected for cross-culturally adaptation to Persian language and psychometrically studying by Salary Majd et al. (2014) in Iran. The authors reported that the Cronbach α of the Persian version of CAPE-V (called the ATSHA) was high. Furthermore, intra-rater and inter-rater reliability of ATSHA were high for all perceptual parameters except for pitch and loudness. They concluded that ATSHA could be used as a valid and reliable Persian scale for perceptual evaluation of the voice [21]. To the best of our knowledge, no study has ever extracted construct convergent validity of the ATSHA thorough investigating correlation between its voice perceptual parameters and the acoustic measures.

In the present study, we were going to consider auditory-perceptual voice assessment by ATSHA and acoustic measurements using *Praat* application in Persian dysphonic patients. In addition, the authors aimed to study the relationship between all perceptual parameters of the ATSHA and target acoustic measures in dysphonic patients in order to investigate construct convergent validity of the ATSHA. Our hypothesis is that a remarkable correlation exists between perceptual evaluation and acoustic measurement of the voice. Furthermore, the effect of different tasks, including vowel prolongation and sentence reading on perceptual voice assessment will be studied in patients with dysphonia.

2. Materials and Methods

Study design

This was a cross-sectional study.

Study participants

The perceptual and acoustic data were gathered from dysphonic patients who were native Persian speakers. The study sample comprised 40 patients with voice problems (25 male and 15 female, with a Mean±SD age of 36.79±8.26 years). The patients were randomly selected from those with voice complaints who attended the otorhinolaryngology clinics of Amir-A'lam Hospital in Tehran, Iran. All patients were diagnosed as having dysphonia by an otorhinolaryngologist and a SLP based on the voice history, videostroboscopic examination, and informal perceptual voice assessment. The selected patients have different severity of dysphonia (mild, moderate, and severe) based on the informal voice assessment of grade of dysphonia.

Khoddami SM, et al. The Relationship Between Auditory Perceptual Evaluation and Acoustic Measurements of the Voice in Dysphonia: Some Issues About the Task Effect on Perceptual Rating. JMR. 2018; 12(2):113-122.

Study procedure

All patients were instructed to sustain vowels /a/ and /i/ approximately 5 seconds and then read 6 sentences of the ATSHA at their habitual pitch and loudness level [21]. The vowel tasks were repeated 3 times, and the longest trials were selected for perceptual and acoustic analyses. For auditory-perceptual voice assessment, both vowel prolongation and sentence reading were used while for the acoustic measurements just the vowel tasks were applied. If the patients could not read the sentences, they would be asked to repeat after the examiner.

The voice samples were recorded using a VAIO notebook (Sony, Model VPCEE23FX); the PRO TOOLS software; a condenser microphone (USB AVID, VOCAL STUDIO) at a sampling frequency of 20-100 kHz. The voice samples were recorded with a constant mouth-to-microphone distance of 10 cm. The voice recording is performed in an acoustic room in the hospital; voice recording lasted about 5 minutes for each participant. All recorded voices were presented to a blind rater thorough a headphone (AKG K7) for perceptual evaluation. To extract the acoustic parameters, all voice samples analyzed blindly by *Praat*, version 5.1.10 [22].

The perceptual judgment was performed in a silent environment by one SLP who was completely familiar with ATSHA. The rater was an experienced SLP with more than 15 years' experience in clinical voice assessment and treatment. The rater had normal hearing and was native speaker of standard Persian. The longest vowel /a/ and /i/ as well as 6 sentences of the ATSHA were presented to the rater two times using a headphone at a comfortable loudness level in a quiet room. The recorded samples were presented blind and randomly. Then, the ATSHA was completed by the rater according to the CAPE-V guideline [18]. The rater asked to record score 0 to 100 to indicate abnormality perceived via auditory judgment for individual voice parameter in the ATSHA. Based on the ATSHA and CAPE-V guideline, the higher score means greater abnormality in each perceptual parameter [18, 21].

Acoustic measurement. The longest vowel /a/ produced by each patient was analyzed using *Praat* to extract the acoustic parameters. Before analysis, the voices were edited using the PRO TOOLS software and the first and last 1.5 seconds of each recording were eliminated because of their larger irregularity and then saved as wave files. Afterwards, the acoustic parameters, including fundamental frequency, intensity, jitter (local), shimmer (local), and Harmonics-to-Noise Ratio (HNR) were extracted.

Statistical analysis

Descriptive statistical analyses were performed and the perceptual data of the ATSHA as well as the acoustic data of *Praat* were reported. Normal distribution of the clinical data was studied using Kolmogorov–Smirnov test. The relation between the perceptual and acoustic assessments was tested by calculating the Pearson correlation coefficient between the ATSHA parameters and the acoustic measures during the sustain vowel /a/. To test the effect of the tasks on the ATSHA scores, Paired t-test was used and Pearson correlation coefficient was calculated. Statistical analyses were performed in SPSS V. 19.0 (SPSS, Inc., Chicago, IL). The threshold for statistical significance was set at $P<0.05$.

3. Results

Characteristics of the participants

The participants comprised 40 dysphonic patients (male=30, female=10). All patients were diagnosed as having a functional dysphonia with and without laryngeal lesions (n=30; male=20, female=10) or neurologic dysphonia (n=10; male=5, female=5).

Perceptual ratings by the ATSHA

Descriptive data

The results of the ATSHA are presented with respect to the vowel and reading tasks in the patients with dysphonia (Table 1). As shown in the Table 1, the overall severity and pitch obtained the highest and lowest scores according to the rater judgment, respectively.

Regarding the voice feature differences between male and female, the mean scores of the ATSHA are displayed in Figure 1 with respect to the gender of the patients. The values reported in this section have been calculated by averaging the ATSHA scores on both tasks, including sustain vowels and sentences reading. According to Figure 1, the experienced SLP assigned the mean score of 41.5 and 54.68 for the overall severity in females and males, respectively. Also, the highest score has been recorded for the overall severity but pitch obtained the lowest score in both females and males. The males achieved higher scores in all parameters, except the breathiness, based on the scores given by the experienced SLP (Figure 1).

Comparison based on the tasks

According to the Table 2, of perceptual parameters of the ATSAH, only the roughness and strain were significantly scored different by the experienced SLP on the vowel and reading tasks (P<0.05). Furthermore, the Pearson correlation was used to measure the effect of task on the perceptual ratings. The correlation between the ATSHA scores given by the experienced SLP in both tasks was high (r=0.82-0.99, P<0.05). The perceptual parameter of overall severity obtained the highest correlation and the strain showed the lowest correlation.

Acoustic measurements by *Praat*

Descriptive data

The mean values and standard deviations of the acoustic parameters in the dysphonic patients are presented in Table 3. In Table 4, the results of relationship between the perceptual ratings and acoustic measurements by Pearson correlation analyses are presented. The correlation analyses will be reported for individual perceptual parameter of the ATSHA as follow as: Overall severity, Roughness, Breathiness, Strain, Pitch, and Loudness.

The Pearson correlation analyses demonstrated a significant correlation between the overall severity and all acoustic parameters except for the frequency (r=0.1; P=0.4).

The correlation between the roughness and all acoustic parameters was significant, although there was no significant correlation between the roughness and the frequency (r=0.02; P=0.99). The Pearson correlation analyses revealed a significant correlation between the breathiness and all acoustic parameters except for the frequency (r=0.27; P=0.1). There was a significant correlation between the strain and all acoustic parameters except for the frequency (r=0.08; P=0.65).

The Pearson analyses found no significant correlation between the pitch and the acoustic parameters of intensity (P=0.03), F0 (P=0.05), and jitter (P=0.14). However, significant correlation was observed between the pitch and shimmer (r=0.45; P=0.006) and between the pitch and HNR (r=-0.38; P=0.01). There was a significant correlation between the loudness and all acoustic parameters except for the frequency (r=0.17; P=0.29). Furthermore, the highest correlation value was seen between the perceptual parameters and the HNR, and the lowest value was assigned to the jitter, although as reported above the pitch obtained the highest correlation value with shimmer (r=0.45; P=0.006).

4. Discussion

The current study aimed to investigate construct convergent validity of the Persian version of CAPE-V (called the ATSHA) using the acoustic measurements.

Table 1. Mean±SD of the ATSHA ratings by the experienced SLP With respect to the tasks in the dysphonic patients (n=40)

Perceptual Parameters	Tasks	Mean	SD
Overall severity	Sustained vowels	53.68	24.37
	Sentences reading	54.68	23.46
Roughness	Sustained vowels	51.40	26.41
	Sentences reading	47.28	26.15
Breathiness	Sustained vowels	23.36	20.25
	Sentences reading	26.16	22.60
Strain	Sustained vowels	40.96	18.42
	Sentences reading	25	19.78
Pitch	Sustained vowels	17.92	20.99
	Sentences reading	17.64	20.56
Loudness	Sustained vowels	33.20	23.25
	Sentences reading	34.92	22.12

Abbreviations: SLP: Speech-Language Pathologist; SD: Standard Deviation

JMR

Khoddami SM, et al. The Relationship Between Auditory Perceptual Evaluation and Acoustic Measurements of the Voice in Dysphonia: Some Issues About the Task Effect on Perceptual Rating. JMR. 2018; 12(2):113-122.

Moreover, the effect of voice tasks on the perceptual ratings was studied. This study showed that all perceptual parameters of the ATSHA, except for the pitch, have construct validity based on the majority of the acoustic analyses done. Of the target acoustic parameters, fundamental frequency had no relationship with perceptual parameters of the ATSHA. However, HNR demonstrated the highest relationship with all perceptual parameters of the ATSHA. Although there are significant differences between scores of the roughness and strain in sustained vowel and reading tasks, high correlation observed between perceptual ratings during vowel prolongation and

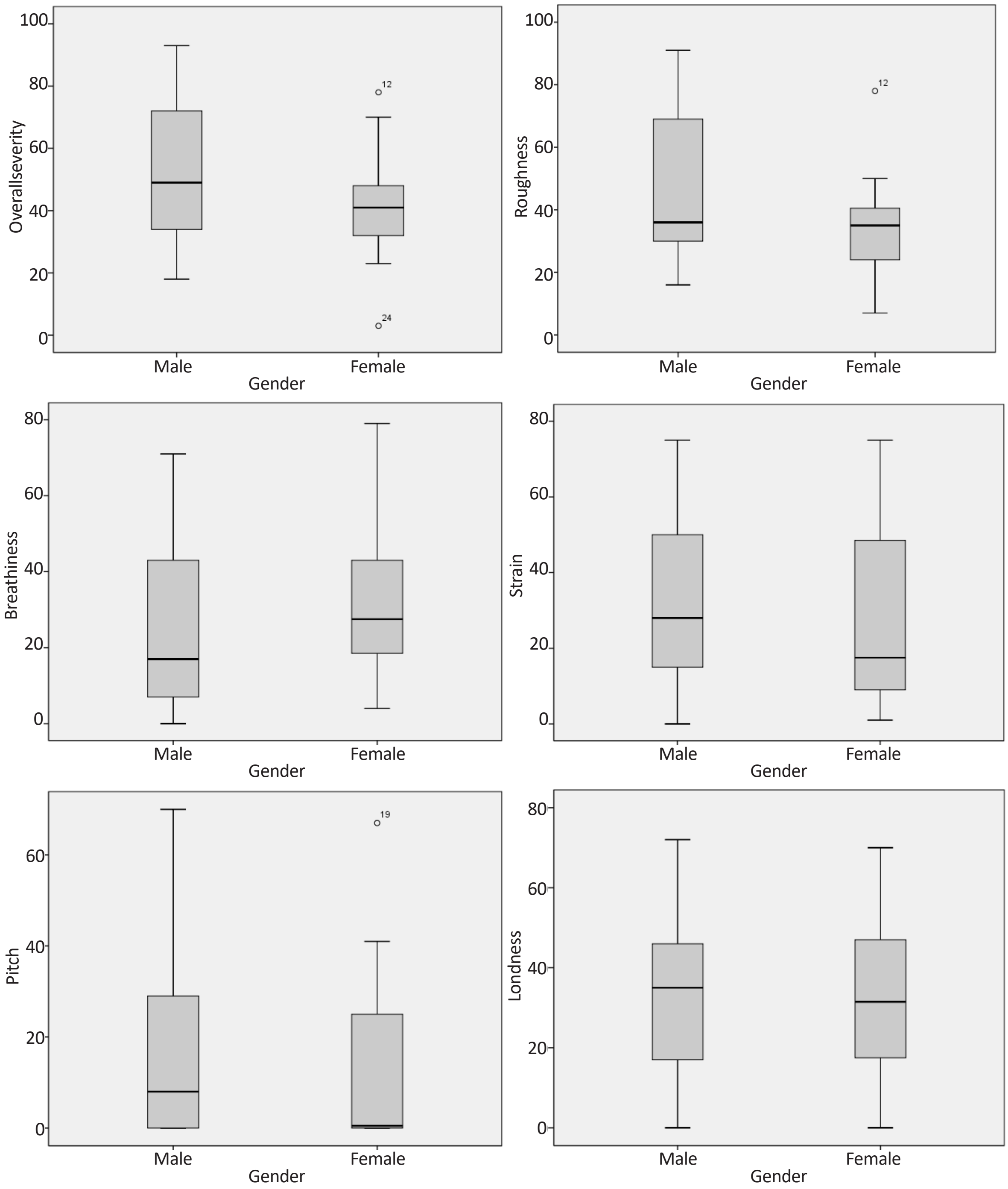

Figure 1. Box-and-whisker plot for the ATSHA in the dysphonic patients (n=40)

The dark line shows the median; The box indicates the middle 50% of the distribution and the whiskers show the remaining 25% at the bottom and top of the distribution. Circles indicate the atypical values.

Table 2. Comparison of the ATSHA ratings by the experienced SLP based on the tasks in the dysphonic patients (n=40)

Perceptual Parameters (Sustained Vowels * Sentences Reading)	Mean	SD	MD	P	95% CI	
					Lower	Upper
Overall severity 1* Overall severity 2	0.260	4.179	49	0.662	-0.92	1.448
Roughness 1* Roughness 2	6.120	10.211	49	0.000*	3.21	9.02
Breathiness 1* Breathiness 2	-0.40	9.258	49	0.976	-2.67	2.59
Strain 1* Strain 2	15.100	12.970	49	0.000*	11.41	18.78
Pitch 1* Pitch 2	-0.38	3.827	49	0.486	-1.46	0.70
Loudness 1* Loudness 2	-0.20	12.123	49	0.991	-3.46	3.42

JMR

Abbreviations: CI: Confidence Interval; SD: Standard Deviation; MD: Mean Difference; SLP: Speech-Language Pathologist; 1: Sustained vowels; and 2: Sentences reading; P<0.05

reading tasks suggest that the effect of tasks is not remarkable on the perceptual assessment of voice by the ATSHA.

The CAPE-V has been accepted as a widespread research and clinical scale for the evaluation of voice [10, 18]. Since development of original version of the CAPE-V in 2002, several versions of this scale were developed and their validity and reliability studied [21-25]. Salary Majd et al. (2014) who studied cultural adaptation and reliability of the Persian version of the CAPE-V (called the ATSHA), reported that the ATSHA has high internal consistency, as well as high intra-rater and inter-rater reliability for all perceptual voice parameters except for the pitch and loudness. Therefore, they recommended this profile as a valid and reliable Persian scale to evaluate perceptual features of the voice in patients with dysphonia [21].

Regarding the lack of documentation of construct convergent validity of the ATSHA and because of multidimensional entity of the voice, the present study was designed to survey the correlation between this perceptual scale and the acoustic measurements. The results of present study indicate that the ATSHA has acceptable construct validity. To be more exact, this study indicate that the ATSHA is able to have acceptable relationship with the voice acoustic measures extracted by *Praat* application in the patients with functional and organic voice disorders.

The authors found no study to investigate concurrent validity of the CAPE-V with the acoustic parameters elicited in the current study. The construct validity of the Italian, Spanish, and Turkish versions of CAPE-V has been confirmed by the GRBAS which is another known auditory scale in the field of voice studies. Mozzanica et al. (2013), Núñez-Batalla et al. (2015), and Özcebe et al. found significant correlation between the CAPE-V and GRBAS; they reported the highest correlation was related to the parameter of overall severity [23-25]. This finding is in agreement with the current study in which the overall severity had moderate to high correlations with the target acoustic measures. Indeed, the highest correlation was seen between the overall severity and HNR (r=-0.85; P=0.000).

As mentioned before, no study has ever surveyed the relation between the CAPE-V or ATSHA and the F0, intensity, perturbation parameters, and HNR. The results

Table 3. Mean±SD of the acoustic measures in the dysphonic patients (n=40)

Dysphonic Patients	Mean±SD				
	F0 (Hz)	Intensity (dB)	Jitter (%)	Shimmer (%)	HNR
Male (25)	142±45.76	58±6.84	0.01±0.01	0.07±0.05	17.58±7.22
Female (15)	194±49.79	54±6.32	0.008±0.008	0.05±0.05	18.32±7.90

Abbreviations: F0: Fundamental Frequency; HNR: Harmonics-to-Noise Ratio; SD: Standard Deviation

JMR

Table 4. Correlations between perceptual and acoustic voice evaluation in the dysphonic patients (n=40)

Acoustic Measures	Overall Severity	Roughness	Breathiness	Strain	Pitch	Loudness
F0 (Hz)	0.10(0.4)	0.02(0.99)	0.27(0.1)	0.08(0.65)	0.35(0.03)	0.17(0.29)
Intensity (dB)	-0.5(0.002)*	-0.54(0.001)*	-0.51(0.001)*	-0.6(0.000)*	-0.31(0.05)	-0.53(0.001)*
Jitter (%)	0.59(0.001)*	0.48(0.003)*	0.49(0.002)*	0.49(0.002)*	0.24(0.14)	0.34(0.03)*
Shimmer (%)	0.65(0.000)*	0.55(0.000)*	0.61(0.000)*	0.59(0.000)*	0.45(0.006)*	0.55(0.000)*
HNR	-0.85(0.000)*	- 0.75(0.000)*	-0.79(0.000)*	-0.74(0.000)*	-0.38(0.01)*	-0.71(0.000)*

*Statistically significant correlation at the 0.05 level (2-tailed)

JMR

of the current study regarding the relationship between the perceptual and acoustic assessments are in agreement with those found by Awan et al. (2010) and Vaz Freitas et al. (2015) [26, 27]. Awan et al. (2010) who studied the relationship between acoustic spectral/cepstral measures (to calculate noise based on the long-term spectral analysis) and perceptual ratings by the CAPE-V, reported a strong correlation between perceptual and acoustic results of dysphonia severity rating [26].

In another study, Vaz Freitas et al. (2015) tried to investigate the correlation between the GRBAS scale and the acoustic measures using different voice analysis software programs including *Praat*. They found the power of relationship between the GRBAS scale and the acoustic measures varies from weak to moderate. They reported the local shimmer and HNR can be strong predictive acoustic parameters for perceptual voice assessment [27]. Although the CAPE-V and GRBAS have different vocal parameters, both are perceptual scales and acceptable relationship between them are expected. However, acceptable relationship between the ATSHA as a perceptual scale and *Praat* application as an instrument to measure acoustic features of the voice is noteworthy and can be considered to confirm construct validity of the ATSHA.

Although the perceptual voice parameters of the ATSHA obtained moderate to high correlation with majority of the acoustic measures in the present study, the findings for the frequency was completely different. Besides, there was no significant correlation between pitch and the acoustic features of intensity and jitter. To interpret this part of findings, let us look at the clinical voice data of our participants. As noted before, the participants were patients with different types of functional (with and without vocal mass lesions) and neurological voice disorders.

Although we tried to select our patients from mild, moderate, and severe cases based on the results of clinical assessments gathered by both ENT and SLP in the diagnosis of voice disorders, our patients generally obtained moderate overall severity of dysphonia according to the ATSHA [21]. Out of 6 perceptual parameters of the ATHSA, the least problem perceived was the pitch parameter which rated a bit more than normal level in all participants during both tasks (Mean score: 17.92 and 17.64 in sustained vowels and sentence reading, respectively).

However, the roughness was rated as moderately abnormal. Moreover, values of the acoustic features in our patients are another issue which should be noted. Although different instruments used for the acoustic analysis in the literature makes the interpretation of the acoustic data a little challenging, comparing the acoustic measures of our patients with suggested normative data for *Praat* measurements [28, 29] showed that fundamental frequency, intensity, jitter, and shimmer values were in normal range and only the mean value of HNR was abnormal (<20). Although we expected significant correlation between frequency and its perceptual correlate, pitch, there was no relation between perceptual voice assessment by the ATSHA and frequency according to the correlation analyses. Apparently, this finding can be expected due to lack of noticeable pitch problems and abnormal frequency values recorded in our patients based on both perceptual and acoustic voice assessment. It is likely that if our patients had severe abnormality in the pitch, we might find remarkable relation between perceptual parameter of the ATSHA and acoustic measure of frequency.

The authors suppose that ATSHA can predict frequency of the voice in condition that the subjects demonstrate mild, moderate or severe problems in pitch. Further investigations are recommended to evaluate construct validity of the ATSHA scale in patients with voice disorders who have high or low pitch as well as acoustic features which are out of normal range appropriate for demographic characteristics of speakers.

This study indicate that the voice tasks used during auditory-perceptual assessment have no effect on the scores recorded by the experienced SLP. As mentioned before, mean comparison and Pearson correlation were applied to investigate the effect of tasks. Based on the mean comparison analysis, the experienced SLP rated all vocal parameters of the ATSHA equal in vowels and reading; only the roughness and strain rated significantly more severe during vowels compared to sentence reading. However, the effect of tasks on all vocal parameters was not supported by Pearson correlation analysis.

According to the literature, results related to the effect of tasks on perceptual evaluation of the voice in patients with dysphonia are inconsistent. No difference was reported by de Krom (1994) and Revis et al. (1999) in perceptual assessments based on the tasks [30, 31], while Wolfe et al. (1995) and Zraick et al. (2005) reported significant difference for vowels versus connected speech [32, 33]. Zraick et al. (2005) hypothesized that a significant difference would be expected in ratings of overall severity during the sustained vowel versus connected speech. The authors found that the listeners perceived the severity of dysphonia as more severe in the sustained mood; however, they reported no remarkable difference on perceptual judgment of the severity of dysphonic voice during oral reading and the sustained vowel /a/ [33].

Apparently, there are several factors influencing the results of auditory-perceptual voice assessment regarding the tasks such as type of voice (normal or dysphonic), type of perceptual vocal parameter and experience of the rater. Undoubtedly, variety of tasks are needed to perform voice assessment but the current study indicate that when an experienced rater who are familiar with the ATSHA rates several parameters of the voice in patients with dysphonia, the effect of tasks used is not remarkable.

The present study is the first one demonstrated that the Persian version of CAPE-V (called the ATSHA) is a valid voice scale to perceptually estimate intensity, jitter, shimmer and HNR in patients with voice disorders. However, this study was not able to provide enough evidence for construct validity of the ATSHA based on the measurements of frequency. Further studies are recommended to investigate construct validity of the ATSHA in patients with pitch deviations. The authors concluded that vowel prolongation and sentence reading has no remarkable effect on the perceptual ratings in dysphonia. Future research can study the effect of conversation speech sample (another voice task used during assessment by the ATSHA) on the perceptual voice evaluation by the ATSHA.

Ethical Considerations

Compliance with ethical guidelines

All participants signed the written informed consent form before study initiation. The study protocol was approved by the Ethics Committee of School of Rehabilitation of Tehran University of Medical Sciences (protocol number 914320080).

Funding

This study was part of a research supported by School of Rehabilitation, Tehran University of Medical Sciences (Grant Number: 91-04-2-20080).

Authors contributions

The authors contributions is as follows: Designing the project, writing the article, and interpreting the results: Seyyedeh Maryam Khoddami; and Data gathering and data analysis: Nazila Salary Majd.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgements

The authors appreciate Dr. Michael Drinnan for his useful comments on designing the project. The authors also thank Dr. Nader Fallahian for his valuable comments on voice recording and analyses.

References

[1] Hirano M. Objective evaluation of the human voice: clinical aspects. Folia Phoniatr (Basel). 1989; 41(2-3):89-144. [DOI:10.1159/000265950]

[2] Hartl DM, Hans S, Crevier Buchman L, Laccourreye O, Vaissiere J, Brasnu D. Dysphonia: Current methods of evaluation. Annales d'Otolaryngologie et de Chirurgie Cervico-faciale. 2005; 122(4):163-72. [DOI:10.1016/S0003-438X(05)82344-8]

[3] Sataloff RT. Professional voice. The science and art of clinical care. San Diego: Singular; 1997. [PMCID]

[4] Carding P. Evaluating voice therapy: Measuring the effectiveness of treatment. London: Whurr Publishers; 2000.

[5] Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, et al. 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 Laryngological Society (ELS). European Archives of Oto-Rhino-Laryngology. 2001; 258(2):77-82. [DOI:10.1007/s004050000299] [PMID]

[6] Klein S, Piccirillo JF, Painter C. Comparative contrast of voice measurements. Otolaryngology–Head and Neck Surgery. 2000; 123(3):164–9. [DOI:10.1067/mhn.2000.107682] [PMID]

[7] Behrman A. Common practices of voice therapists in the evaluation of patients. Journal of Voice. 2005; 19(3):4540-69. [DOI:10.1016/j.jvoice.2004.08.004] [PMID]

[8] Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen N, Wilson JA. The reliability of three perceptual evaluation scales for dysphonia. European Archives of Oto-Rhino-Laryngology. 2004; 261(8):429-34. [DOI:10.1007/s00405-003-0707-7] [PMID]

[9] Uloza V, Saferis V, Uloziene I. Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonomicrosurgery. Journal of Voice. 2005; 19(1):138-45. [DOI:10.1016/j.jvoice.2004.01.009] [PMID]

[10] Oates J. Auditory-perceptual evaluation of disordered voice quality pros, cons and future directions. Folia Phoniatrica et Logopaedica. 2009; 61(1):49-56. [DOI:10.1159/000200768] [PMID]

[11] Kreiman J, Gerratt BR, Kempster GB, Erma A, Berke GS. Perceptual evaluation of voice quality: Review, tutorial and a framework for future research. Journal of Speech, Language, and Hearing Research. 1993; 36(1):21-40. [DOI:10.1044/jshr.3601.21] [PMID]

[12] Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. Journal of Speech, Language, and Hearing Research. 1996; 39(2):311-21. [DOI:10.1044/jshr.3902.311] [PMID]

[13] Selby JC, Gilbert HR, Lerman JW. Perceptual and acoustic evaluation of individuals with laryngopharyngeal reflux pre- and post-treatment. Journal of Voice. 2002; 17(4):557-70. [DOI:10.1067/S0892-1997(03)00017-1]

[14] Bhuta T, Patrick L, Garnett JD. Perceptual evaluation of voice quality and its correlation with acoustic measurements. Journal of Voice. 2003; 18(3):299–304. [DOI:10.1016/j.jvoice.2003.12.004] [PMID]

[15] Lopes LW, Lima I, Almeida L, Cavalcante D, and Almeida A. Severity of voice disorders in children: Correlations between perceptual and acoustic data. Journal of Voice. 2012; 26(6):819.e7-e12. [DOI:10.1016/j.jvoice.2012.05.008] [PMID]

[16] Ma EP, Yiu EM. Multipara metric evaluation of dysphonic severity. Journal of Voice. 2006; 20(3):380-90. [DOI:10.1016/j.jvoice.2005.04.007] [PMID]

[17] De Bodt MS, Wuyts FL, Van de Heyning PH, Croux C. Test–retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. Journal of Voice. 1997; 11(1):74-80. [DOI:10.1016/S0892-1997(97)80026-4]

[18] Kempster, GB Gerratt BR, Abbott KV, Barkmeier-Kraemer J, Hillman RE. Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. American Journal of Speech-Language Pathology. 2009; 18(2):124-32. [DOI:10.1044/1058-0360(2008/08-0017)]

[19] Karnell MP, Melton SD, Childes JM, Coleman TC, Dailey SA, Hoffman HT. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. Journal of Voice. 2007; 21(5):576-90. [DOI:10.1016/j.jvoice.2006.05.001] [PMID]

[20] Zraick RI, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, et al. Establishing the validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). American Journal of Speech-Language Pathology. 2011; 20(1):14-22. [DOI:10.1044/1058-0360(2010/09-0105)]

[21] Salary Majd N, Khoddami SM, Drinnan M, Kamali M, Amiri-Shavaki Y, Fallahian N. [Validity and rater reliability of Persian version of the Consensus Auditory Perceptual Evaluation of Voice (Persian)]. Audiology. 2014; 23(3):65-74.

[22] Mozzanica F, Ginocchio D, Borghi E, Bachmann C, Schindler A. Reliability and validity of the Italian version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Folia Phoniatrica et Logopaedica. 2013; 65(5):257-65. [DOI:10.1159/000356479] [PMID]

[23] Nú-ez-Batalla F, Morato-Galán M, García-López I, Ávila-Menéndez A. Validation of the Spanish adaptation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Acta Otorrinolaringológica Española. 2015; 66(5):249-57. [DOI:10.1016/j.otoeng.2015.08.001]

[24] de Almeida SCC. Validity and reliability of the 2nd European Portuguese version of the "Consensus Auditory-Perceptual Evaluation of Voice" (II EP CAPE-V) [PhD. dissertation]. Setúbal: Institute of Setúbal; 2016.

[25] Özcebe E, Aydinli FE, Tiğrak TK, İncebay Ö, Yilmaz T. Reliability and validity of the Turkish version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Journal of Voice. 2017; (In press). [DOI:10.1016/j.jvoice.2017.11.013]

[26] Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. Clinical Linguistics & Phonetics. 2010; 24(9):742-58. [DOI:10.3109/02699206.2010.492446] [PMID]

[27] Vaz Freitas S, Melo Pestana P, Almeida V, Ferreira A. Integrating voice evaluation: Correlation between acoustic and audio-perceptual measures. Journal of Voice. 2015; 29(3):390.e1-7. [DOI:10.1016/j.jvoice.2014.08.007] [PMID]

[28] Boersma P and Weenink D. *Praat*: Doing phonetics by computer (Version 5.1.17). Amsterdam: University of Amsterdam; 2009.

[29] Maryn Y, Corthals P, De Bodt M, Van Cauwenberge P, and Deliyski D. Perturbation measures of voice: A comparative study between Multi-Dimensional Voice Program and *Praat*. Folia Phoniatrica et Logopaedica. 2009; 61(4):217-26. [DOI:10.1159/000227999] [PMID]

[30] De Krom G. Consistency and reliability of voice quality ratings for different types of speech fragments. Journal of Speech, Language, and Hearing Research. 1994; 37(5):985-1000. [DOI:10.1044/jshr.3705.985] [PMID]

[31] Revis J, Giovanni A, Wuyts F, Triglia JM. Comparison of different voice samples for perceptual analysis. Folia Phoniatrica et Logopaedica. 1999; 51(3):108-16. [DOI:10.1159/000021485] [PMID]

[32] Wolfe VI, Cornell R, Fitch J. Sentence/vowel correlation in the evaluation of dysphonia. Journal of Voice. 1995; 9(3):297–303. [DOI:10.1016/S0892-1997(05)80237-1]

[33] Zraick RI, Wendel K, Smith-Olinde L. The effect of speaking task on perceptual judgment of the severity of dysphonic voice. Journal of Voice. 2005; 19(4):574-81. [DOI:10.1016/j.jvoice.2004.08.009] [PMID]