Influence of clinical context in characterization of severity of vocal deviation

Influência da escuta contextualizada na percepção da intensidade do desvio vocal

Flávia Pereira da Costa[1], Rosiane Yamasaki[1], Mara Behlau[1]

ABSTRACT

Purpose: We verify if the clinical context interferes in the assessment of vocal deviation, considering the overall degree of severity. **Methods:** We selected 22 voice recordings of 12 women and 10 men diagnosed with organic or functional dysphonia, aged between 25 and 75 years old, pre-and post-therapy. The vocal sample was analyzed by two SLP voice specialists, one of which (SLP-1) was the patients' therapist and conducted a contextualized clinical assessment. On the other hand, the second voice specialist (SLP-2), did not know any of the patients and conducted the assessment only by listening to the recordings. The speech material used was the sustained vowel /ɛ/ and continuous speech (number counting 1 to 10). The overall degree of vocal deviation should be marked on a visual analog scale of 100 mm. **Results:** In the sustained vowel the SLP-1 produced an average of 53.8 on pre-therapy evaluation (range 17 to 100), while the SLP-2 produced an average of 62.8 (range 32 to 100). In the post-therapy assessments, the average was 22.8 for SLP-1 (range 7 to 47), and 51.9 for SLP-2 (range 28 to 92). To the continuous speech the post-therapy assessments was the only with significant difference, the SLP-1 produced an average of 18.41 (range 5 to 55) while the SLP-2 produced an average of 43.55 (range 18 to 80). **Conclusion:** The sustained vowel suffers more influences of demographic data and diagnostic vocal than continuous speech.

Keywords: Voice; Voice quality; Voice disorders; Speech therapy; Demographic data

RESUMO

Objetivo: Verificar se a escuta clínica contextualizada influencia a avaliação perceptivo-auditiva da intensidade do desvio vocal. **Métodos:** Foram selecionados 22 registros vocais de 12 mulheres e dez homens, faixa etária de 25 a 75 anos, pré e pós fonoterapia. A amostra vocal foi analisada por duas fonoaudiólogas especialistas em voz. A avaliadora 1 era a terapeuta dos pacientes e realizou a escuta clínica contextualizada e a avaliadora 2 não conhecia nenhum dos pacientes e realizou a escuta do sinal sonoro. O material de fala foi a emissão da vogal sustentada "é" e a contagem de números de um a dez. A estratégia utilizada foi a de marcar o grau geral de desvio vocal em uma escala analógica visual de 100 mm. **Resultados:** Na vogal sustentada, no momento pré-fonoterapia, a avaliadora 1 produziu a média de 53,8 pontos (faixa de 17 a 100), enquanto a avaliadora 2 produziu a média de 62,8 (faixa de 32 a 100). No momento pós fonoterapia, a média da avaliadora 1 foi de 22,8 pontos (faixa de 7 a 47) e a da avaliadora 2 foi de 51,9 pontos (faixa de 28 a 92), para vogal sustentada. Para a fala encadeada, o momento pós fonoterapia foi o único que se encontrou diferença, sendo que a avaliadora 1 teve média de 18,41 pontos (faixa de 5 a 55) e a avaliadora 2, média de 43,55 pontos (faixa de 18 a 80). **Conclusão:** A vogal sustentada sofreu mais influência do conhecimento de dados demográficos e de diagnóstico vocal do que a fala encadeada.

Descritores: Voz; Qualidade da voz; Distúrbios da voz; Fonoterapia; Dados demográficos

INTRODUCTION

Perceptual voice assessment is the main tool in speech-language pathology clinics and sovereign regarding the others methods. It allows the voice quality characterization and also quantifying the perceived deviation[1]. Through these data it is possible to infer about the vocal health and anatomic-functional of larynx and vocal tract of the speaker, it will contribute to directing speech-language pathologist clinical logic. It is a fast, economic, non-invasive, and robust assessment[2-4].

Voice quality of dysphonic subjects involves multidimensional characteristics that may be partially identified, if not fully, through perceptual assessment. Although it is considered the gold-standard by many researchers, the perceptual assessment has been the target of many critics due its subjective nature. However, studies show that the big variability among perceptual judgments are more related to inadequate experimental procedures[5]. To make this analysis more reliable and robust it is necessary to control important factors that may interfere in its results, as the experience and training of the evaluator, the vocal deviation degree, the stimuli type, the assessment scale, the vocal parameters selected, and the instruction provided prior assessment. A carefully control of these factors increases analysis reliability and reduces the subjectivity degree involved[1,5,6-12].

Recently, as in clinical and scientific environments, the perceptual answering scale most used is the analogue visual one. Although the numeric seems to be easier to be used it ends to concentrate diverse results in the same assessment degree[5,13] and the punctuation in the analogue visual scale is more proper to distinguish slight changes in vocal quality. The analogue-visual scale is used in self-evaluation scales and in CAPE-V (Consensus Auditory Perceptual Evaluation of voice – ASHA 2003), the protocol developed in consensus by voice experts, considering modern trends to human perception measurements[14].

Analogue visual scale correlation with numeric scale allowed the acquaintance of cut values which made possible to distinguish normal from deviated voices. In a study done in Finland[15], the found value was 35 mm while in Brazil[16] the value obtained was 35.5 mm. The similarity of values found in both studies performed in different countries showed this evaluation is robust and reliable.

Besides the interference factors already described the perceptual evaluation may also be influenced by context in which it is performed. The knowledge of patients' clinical conditions by the evaluator, as demographic data, gender, age, physician diagnose[17], dysphonia history, and clinic assessment, as well as the hearing situation, which may be clinical or scientific[18], may interfere in the result of perceptual evaluation. Clinical voice assessment usually is made in the presence of the patient, together with the knowledge of his data, diagnose, voice health condition, and his clinic progress. Yet the blind assessment usually is performed by the sound register without access to patients' basic data. Both assessment situations, contextualized and blind are very often used in speech-language practice, as for clinical or scientific means, however it is necessary to know if these conditions influence the result of perceptual evaluation.

The purpose of this research was to verify whether the contextualized clinic assessment, characterized by the knowledge of the patients' clinic history by the evaluator, influences the perceived vocal deviation in perceptual evaluation.

METHODS

This study was performed in Centro de Estudos da Voz (CEV) with approval of Ethics in Research of *Instituto de Ciências Biomédicas* of *Universidade de São Paulo* (ICB-USP), under the protocol number 1026.

Voice sample

It was selected from CEV voices files, 22 dysphonic patients' voice samples, 12 women and 10 men with either organic or functional dysphonia diagnose, age ranging from 25 to 75 years old, either low or high vocal demand. The sample was taken from the speech-language pathologist 1 voice data that performed the patients' treatment and it consisted of sustained vowel /ε/ and continuous speech (counting numbers from one to ten), pre and post therapy.

Vocal sample of pre and post therapy moments were randomly presented to both evaluators with 10% of repetition to internal reliability. Therefore, each evaluator performed the perceptual assessment of 48 sustained emissions – vowel /ε/ - and 48 continuous speech emissions – counting numbers from one to ten. The sustained vowels and continuous speech were assessed in two different sections with a seven days interval between then. Speech-language pathologist 1, clearly knew her patients and their progress in clinical treatment.

Evaluators

Perceptual assessment was performed by two evaluators, both speech-language pathologists and voice specialists. The speech-language pathologist 1, 30 years of experience, performed the contextualized assessment and was the therapist of the patients. Therefore, she knew the laryngeal diagnose, the dysphonia history, and performed the vocal rehabilitation of all patients. The speech-language pathologist 2, 15 years of experience, performed the blind assessment and did not have access to any patient information. Both evaluators had broad experience in dysphonic voices assessment and in the use of analogue visual scale.

Assessment protocol

The perceptual assessment of general vocal deviation degree (G) was performed using the visual analogue scale (VAS) of 100 mm, commonly used by the evaluators. As each millimeter corresponds to a vocal deviation degree this scale allows 100 graduation possibilities. According to previous studies[19], there are three cut points determining four distribution ranges at VAS: 35.5 mm (0.702 sensitivity and 1.000 specificity); 50.5 mm (0.769 sensitivity and 1.000 specificity), and 90.5 mm (0.962 mm sensitivity and 0.953 specificity).

Thus the scores from 0 to 35.5 correspond to normal variability of vocal quality; from 35.6 mm to 50.5, slight to moderate deviation; from 50.6 mm to 90.5 mm, moderate deviations; and from 90.6 to 100 mm severe deviations.

Statistical analysis

Data were computed and analyzed as follow: comparison of assessments from both speech-language pathologists pre and post therapy as for sustained vowel as continuous speech; agreement degree between evaluators; differences between assessments according to gender, age, clinical diagnose, and vocal demand. To statistical analysis it was adopted the significance level of 5% (0.05). The non-parametric tests of Wilcoxon and equality of two proportions were used.

RESULTS

The results to sustained vowels showed a low degree of agreement inter-evaluators for both pre and post therapy moments. Yet the agreement to continuous speech was low only at

post-therapy. Evaluators had internal reliability to both tasks.

Speech-language pathologist 2, the one performing blind assessment, considered the highest values of deviation to both moments at sustained vowel task as showed by the mean punctuation data between evaluators, identifying difference between assessments (Table 1).

To continuous speech the difference between evaluators was found only to post-therapy moment in which the highest values of deviation were scored by speech-language pathologist 2 (Table 2).

Gender variables, physician diagnoses, and vocal demand did not showed differences at result, indicating they do not interfere at assessment. But, the sample was not widely enough to precisely state this.

DISCUSSION

Perceptual evaluation is the main tool of voice assessment in dysphonic patients and the also very used in scientific researches. Therefore it may be used for both scientific and clinical means and, many times, performed by clinic speech-language pathologists. Although it is a subjective assessment, the proper control of interference factors of this evaluation, as the listener experience, vocal deviation degree, vocal parameters, stimuli type, and evaluator training this analysis may become robust and reliable[3,5]. Besides these factors, it is not know whether the contextualized evaluation, when the speech-language pathologist know from the complaint, the larynx diagnose trough clinic progress in therapy, is an interference factor to perceptual evaluation.

At the present study the main interferences factors in perceptual evaluation were proper controlled. Thus it was

Table 1. Perceptual voice assessment values of sustained vowel /ɛ/ by evaluators 1 and 2 pre and post therapy

	Sustained vowel	Mean	Median	Standard-deviation	Minimum	Maximum	CI	p-value
Pre	EV 1	53.82	41.5	28.09	17	100	11.74	0.001*
	EV 2	62.82	57	24.77	32	100	10.35	
Post	EV 1	22.82	21.5	11.33	7	47	4.73	<0.001*
	EV 2	51.95	48	14.36	38	92	6	

*Significant values (p<0.05) – Wilcoxon Test
Note: EV = Evaluation; CI = Mean confidence interval

Table 2. Perceptual voice assessment of continuous speech by evaluators 1 and 2 pre and post therapy

	Continuous speech	mean	Median	Standard-deviation	minimum	Maximum	CI	p-value
Pre	EV 1	60.32	58	25.16	26	100	10.51	0.334
	EV 2	57.77	49.5	25.71	25	100	10.74	
Post	EV 1	18.41	14	14.04	5	55	5.87	<0.001*
	EV 2	43.55	15.01	15.01	18	80	6.27	

*Significant values (p<0.05) – Wilcoxon Test
Note: EV = Evaluation; CI = Mean confidence interval

optioned to perform the analysis of vocal deviation general degree (g) from the Japanese GRBAS[20] scale, since it is a robust and reliable parameter[3,11], measured by the analogue visual scale since it has more reliability inter-evaluators[13] and it is more sensitive to small differences in punctuation than numeric scale[2].

Evaluation results showed differences mainly in post--therapy (Tables 1 and 2), pointing out that probably the previous knowledge of speech-language pathologist 1 influenced them, since she was the responsible for the voice therapy and followed all the patients clinic progress, from evaluation to the end of rehabilitation. This hypothesis shows the subjectivity involved in perceptual evaluation[21] and reinforce the need to define better the purpose of this evaluation in order to better chose the proper evaluation condition.

The differences in the post-therapy results in which the speech-language pathologist 2 scored more punctuations than speech-language pathologist 1, as for sustained vowel as for continuous speech (Tables 1 and 2), demonstrated that following the patient and his therapy direct affects the assessment, mainly post-therapy, by context and human value in seeking changes in vocal behavior. During the vocal rehabilitation process, therapy involves change in muscle patterns, healthier vocal behavior review, and replacement of inadequate habits, besides the process of adaption to vocal image. To follow all these variables certainly produces an impact in the clinician perception, even keeping the health distance patients-therapist.

Evaluators may develop internal reference patterns according to the assessment models[22,23]. Therefore inter-evaluators differences, mainly in professionals with distinct experience level, may be expected. However, in this study, both speech--language pathologists had clinical and scientific experience in dysphonic voices assessment and broad hearing training.

The study that identified the cut values to analogue visual scale in Brazil[19] had higher sensitivity and specificity levels, using vocal deviation general degree (G) for its analysis, showing the instrument to be good to perceptual evaluation of voice general deviation degree. This scale is composed by three cut values and four distribution ranges[15,19]. These founds differ from other protocols that also uses the analogue visual scale in their concept, as the CAPE-V, which have different distribution ranges[14].

The use of different vocal tasks to perform the perceptual evaluation is a bias that have been explored in the last years and researches show that scores of vocal deviation general degree have been higher to sustained vowels compared to continuous speech tasks[10,12]. In these researches it was observed higher vocal deviations in sustained vowel evaluation than continuous speech, pointing out that both tasks kinds must be used to clinic judgment in perceptual evaluation of dysphonia severity[24], agreeing with the results of the present study.

In the continuous speech production the emission is really close to natural patients' speech, because there is more interaction between source and filter (articulation, speech rate, and rhythm), while in sustained vowel there is only important information about source and filter. In some way, the data in continuous speech made the judgment between the two speech-language pathologists closer in pre-therapy. It is hard to raise hypothesis about the observed differences in sustained vowel and, perhaps, in this sound type, the personal preferences and internal reference system may had higher value in perceptual judgment.

Trough the data it was possible to observe that only in post-therapy there was a higher distance between analyses, which may be direct consequence of contextualized assessment. It is worthy to point out that the assessments between the evaluators in pre-therapy were very close even to a 100 possibilities scale. This shows that clinic experience of the evaluators in the dysphonic patients analysis strongly favors the reliability inter-evaluators.

Carefully observing the results and relating them with the deviation degree found in analogue visual scale[18], it was verified once again the interference of task type during the perceptual evaluation. It is possible to infer that speech--language pathologist 1, knowing the data and the patients therapy moments, indentified vocal changes, while speech--language pathologist 2, performing blind assessment, did not perceive vocal changes between the therapy moments for sustained vowel (Table 1), but perceived a slight change in voices during continuous speech task (Table 2).

Differences between clinical vocal evaluations and scientific environment evaluations have been reported, pointing out that clinical bias, as the place of assessment and clinical knowledge may affect the results[18]. According with the study, in clinical environment speech-language pathologists has access to medical history throughout anamneses prior to performing vocal assessment. Yet in scientific environment the perceptual evaluation is usually performed without the context, clinical information may be hold and voice samples random to posterior assessment[18]. In this scope we may infer that prior knowledge changes the assessments results and may increase or decrease the severity of vocal deviation.

When comparing the perceptual evaluation without larynx diagnoses knowledge and after this be revealed to evaluators, it was verifies an increase in vocal deviation severity scores, once they had the diagnoses[17]. Therefore it is recommended that assessment is performed in consistent situations, as pre and post therapy, in order to assure its validity. The clinical context knowledge is already including diagnoses and also other data, as gender, age, and therapeutic progress, must be another bias to be avoided in perceptual evaluation.

Stimuli presented for the first time, without any patient context and after several weeks, with complementary data, in private, whether voice was pre or post therapy, showed that perceptual evaluation was more dependent of context

information (pre/post treatment) than just the hearing of sound sign and that only blind tests may offer reliable results in vocal assessment[25],evidencing that way as is performed, the importance to delimit the assessment way to avoid a bias.

Experiment design is crucial to know exactly which are being evaluated. Surely, the patients' therapist is not an impartial evaluator. Scientific researches involving perceptual evaluation need to have a precise and clear experimental design. Therefore, if the purpose of the research is to evaluate vocal sample in clinic point of view, the evaluators need to have access to patients' clinical information. On the other hand, if the purpose of the research is exclusively to analyze the degree of deviation of sound sign, none information need to be offered to evaluators and the therapist that made the patients' following, which voices are part of the sample, must be excluded from evaluation.

The need to standardize the procedures to clinical voice evaluation is necessary[18,25], in the attempt to elucidate potential addictions in voice perceptual evaluation.

Summing up, the research results showed that the clinic context significantly influences the severity of vocal deviation evaluation in different vocal tasks, mainly in post-therapy. But, the limitation of the study is the low number of vocal sample and the low number of evaluators. In future studies it is suggested the increase of vocal samples as well as the use of more evaluators for both analysis, that is, contextualized and blind assessments.

CONCLUSION

Vocal assessment with previous clinic context impacts on dysphonia perception even to skilled speech-language pathologists. Sustained vowel suffers higher variability between evaluators than continuous speech.

REFERENCES

1. Kreiman J, Gerratt BR, Kempster GB, Erma A, Berke GS. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. J Speech Hear Res. 1993;36(1):21-40.

2. Wuyts FL, De Bodt MS, Heyning PHV. Is the reliability of a visual analog scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. J Voice. 1999;13(4):508-17.

3. Oates J. Auditory-perceptual evaluation of disordered voice quality. Folia Phoniatr Logop. 2009;61(1):49-56.

4. Carding PN, Awilson J, Mackenzie K, Deary IJ. Measuring voice outcomes: state of the science review. J Laringol Otol. 2009;123(8):823-9.

5. Patel S, Shrivastav R. Perception of dysphonic vocal quality: some thoughts and research update. ASHA SID-3 Newsletter - Voice Voice Disorders. 2007;17(2):3-6.

6. Behlau M, Madazio G, Feijó D, Pontes P. Avaliação de voz. In: Behlau M. Voz: o livro do especialista. Vol. 1, Avaliação de voz. Rio de Janeiro: Revinter; 2001. p. 85-246.

7. Speyer R, Wieneke GH, Dejonckere PH. Documentation of progress in voice therapy: perceptual, acoustic, and laryngostroboscopic findings pretherapy and posttherapy. J Voice. 2004;18(3):325-40.

8. Bele IV. Reliability in perceptual analysis of voice quality. J Voice. 2005;19(4):555-73.

9. Eadie TL, Doyle, PC. Direct magnitude estimation and interval scaling of pleasantness and severity in dysphonic and normal speakers. J Acoust Soc Am. 2002;112(6):3014-21.

10. Zraick RI, Wendel K, Smith-Olinde L. The effect of speaking task on perceptual judgment of the severity of dysphonic voice. J Voice. 2005;19(4):574-81.

11. Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. J Voice. 2006;20(4):527-44.

12. Wolfe V, Cornell R, Fitch J. Sentence/vowel correlation in the evaluation of dysphonia. J Voice. 1995;9(3):297-303.

13. Kreiman J, Gerratt BR, Ito M. When and why listeners disagree in voice quality assessment tasks. J Acoust Soc Am. 2007;122(4):2354-64.

14. Behlau M. Consensus Auditory- Perceptual Evaluation of Voice (CAPE-V), ASHA 2003. Refletindo sobre o novo. Rev Soc Bras Fonoaudiol. 2004;9(3):187-9.

15. Simberg S, Sala E, Laine A, Rönnemaa AM. A fast and easy screening method for voice disorders among teacher students. Logoped Phoniatr Vocol. 2001;26(1):10-16.

16. Yamasaki R, Leão SHS, Madazio G, Padovani M, Azevedo R. Análise perceptivo-auditiva de vozes normais e alteradas: escala analógica visual. In: Anais do 15. Congresso Brasileiro de Fonoaudiologia; 7. Congresso Internacional de Fonoaudiologia; 16-20 out 2007; Gramado, Brasil. São Paulo: Sociedade Brasileira de Fonoaudiologia; 2007. p. 16-20.

17. Eadie T, Sroka A, Wright DR, Merati A. Does knowledge of medical diagnosis bias auditory-perceptual judgments of dysphonia? J Voice. 2011;25(4):420-9.

18. Solomon NP, Helou LB, Stojadinovic A. Clinical versus laboratory ratings of voice using the CAPE-V. J Voice. 2011;25(1):e7-14.

19. Yamasaki R, Leão SHS, Madazio G, Padovani M, Azevedo R, Behlau M. Correspondência entre Escala Analógico-Visual e a Escala Numérica na avaliação perceptivo-auditiva de vozes. In: Anais do 16. Congresso Brasileiro de Fonoaudiologia; 24 a 27 set 2008; Campos de Jordão, Brasil. São Paulo: Sociedade Brasileira de Fonoaudiologia; 2008. p. 24-7.

20. Hirano M. Clinical examination of voice. New York: Springer; 1981. (Disorders of human communication, 5),

21. Chan KM, Yiu EM. The effect of anchors and training on the reliability of perceptual voice evaluation. J Speech Lang Hear Res. 2002;45(1):111-26.

22. Eadie TL, Kapsner M, Rosenzweig J, Waugh P, Hillel A, Merati A. The role of experience on judgments of dysphonia. J Voice. 2010;24(5):564-73.

23. Hakkesteegt MM, Brocaar MP, Wieringa MH, Feenstra L. The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. J Voice. 2008;22(2):138-45.

24. Maryn Y, Roy N. Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. J Soc Bras Fonoaudiol. 2012;24(2):107-12.

25. Ghio A, Révis J, Merienne S, Giovanni A. Top-down mechanisms in dysphonia perception: the need for blind tests. J Voice. 2013;27(4):481-5.