

Leonardo Wanderley Lopes¹
 Débora Pontes Cavalcante²
 Priscila Oliveira da Costa¹

Descritores

Voz
 Disfonia
 Percepção Auditiva
 Acústica
 Avaliação

Keywords

Voice
 Dysphonia
 Auditory Perception
 Acoustics
 Evaluation

Endereço para correspondência:

Leonardo Wanderley Lopes
 Departamento de Fonoaudiologia,
 Centro de Ciências da Saúde
 Cidade Universitária - Campus I,
 Castelo Branco, João Pessoa (PB), Brasil,
 CEP: 58051-900.
 E-mail: lwlopes@hotmail.com

Recebido em: 01/10/2013

Aceito em: 16/07/2014

CoDAS 2014;26(5):382-8

Intensidade do desvio vocal: integração de dados perceptivo-auditivos e acústicos em pacientes disfônicos

Severity of voice disorders: integration of perceptual and acoustic data in dysphonic patients

RESUMO

Objetivo: Analisar a correlação entre medidas acústicas e a intensidade do desvio vocal e seu poder discriminatório na detecção da presença de alteração vocal, na classificação da intensidade do desvio e na diferenciação do tipo de voz predominante. **Métodos:** Participaram 186 pacientes disfônicos. O desvio vocal foi avaliado por consenso por três fonoaudiólogos, a partir da emissão da vogal /ε/, utilizando-se uma escala analógico-visual. Foram extraídas a média e desvio padrão (DP) da frequência fundamental (F_0), *jitter*, *shimmer* e *Glottal to Noise Excitation Ratio* (GNE). **Resultados:** A média de F_0 correlacionou-se negativamente com o grau de rugosidade e positivamente com o grau de tensão. O DP de F_0 correlacionou-se positivamente com o grau geral, de rugosidade, de tensão e de instabilidade. Houve correlação positiva entre o *jitter* e *shimmer* e todos os parâmetros perceptivos. Apenas o GNE diferenciou vozes saudáveis de vozes desviadas e classificou o grau de sopro. O *shimmer* separou vozes rugosas de vozes não rugosas. A média de F_0 foi útil para determinar o grau de tensão fonatória e diferenciou vozes rugosas das soprosas e tensas. **Conclusão:** Há correlação entre as medidas acústicas e perceptivo-auditivas. O *shimmer*, o GNE e o DP de F_0 podem ser utilizados na detecção de rugosidade, sopro e tensão, respectivamente. O GNE e a média de F_0 são úteis para classificar o grau de sopro e tensão, respectivamente. A média de F_0 diferenciou vozes rugosas, soprosas e tensas, sendo as vozes rugosas mais graves em relação às outras duas.

ABSTRACT

Purpose: To analyze the correlation between acoustic measures and intensity of vocal deviation, and its discriminatory power in detecting the presence of vocal change, classifying the severity of the deviation and differentiating the type of predominant voice. **Methods:** The sample comprised 186 patients with dysphonia. The vocal deviation from the vowel /ε/ was evaluated by consensus between three speech language pathologists, using a Visual Analog Scale. We extracted the mean and standard deviation (SD) of the fundamental frequency (F_0), jitter, shimmer, and Glottal-to-Noise Excitation Ratio (GNE). **Results:** Mean F_0 was negatively correlated with the degree of roughness and positively with the degree of tension. The F_0 SD was positively correlated with the overall, roughness, tension, and instability grades. The jitter and shimmer were positively correlated with all perceptual parameters. Only the GNE distinguished between healthy and dysphonic voices and rated the degree of breathiness. The shimmer separated rough voices from not rough voices. The mean F_0 was helpful to determine the degree of phonatory tension and to separate rough voices from breathy and strained voices. **Conclusion:** There is a correlation between the acoustic and auditory perceptual measures. The shimmer, the GNE, and F_0 SD can be used to detect roughness, breathing, and strain, respectively. The GNE and mean F_0 are useful to classify the degree of breathiness and strain, respectively. The mean F_0 distinguished between rough, breathy, and strained voices, with rough voices more severe compared to the other two.

Trabalho realizado no Departamento de Fonoaudiologia, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

(1) Departamento de Fonoaudiologia, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

(2) Curso de Fonoaudiologia, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

A voz é multidimensional⁽¹⁾ e sua produção está relacionada aos aspectos anatomofisiológicos, emocionais, orgânicos, ambientais e comportamentais. Portanto, é necessário que a avaliação vocal também siga esse princípio, mapeando os aspectos da produção da voz e correlacionando-os, permitindo, assim, uma visão global da disфонia.

O objetivo da avaliação vocal é analisar a qualidade vocal, identificando se a voz é saudável ou não, diagnosticar a presença de um distúrbio de voz, monitorar a progressão de uma doença ou de uma função, avaliar o prognóstico e identificar se o indivíduo avaliado possui riscos para o desenvolvimento desse distúrbio⁽²⁾.

De modo geral, os estudos na área de avaliação e diagnóstico da voz procuram responder a três questões clínicas essenciais⁽²⁾: qual a capacidade da medida para determinar a presença/ausência de um distúrbio de voz (diagnóstico)? Qual a evidência de que o teste utilizado consegue determinar a natureza (etiologia) de um distúrbio de voz? E qual a capacidade que uma medida possui para determinar a extensão (intensidade) de um distúrbio de voz?

Em uma pesquisa⁽³⁾ realizada com fonoaudiólogos experientes nos Estados Unidos, 100% de 53 entrevistados relataram utilizar medidas perceptivo-auditivas durante a avaliação vocal, seguida por observação da postura e movimentação corporal e pela investigação da dinâmica vocal. Esses métodos de avaliação subjetivos eram significativamente mais utilizados do que a avaliação objetiva por meio das medidas acústicas.

No entanto, em uma revisão sistemática, observou-se que a maioria (60%) dos estudos sobre avaliação de pacientes com distúrbios da voz utilizou medidas acústicas e estava focada na identificação da presença ou ausência desse distúrbio (78%). Poucos estudos (18%) investigaram a habilidade da medida em mensurar a intensidade do distúrbio vocal⁽²⁾. Além disso, a maior parte dos estudos utilizou a imagem laríngea como padrão de referência para definir a presença de um distúrbio de voz.

Diante disso, pode-se refletir que, primeiramente, uma das maiores aplicações de uma medida de avaliação é julgar a efetividade de um tratamento oferecido, que pode incluir desde a ausência de uma doença diagnosticada previamente até a redução da intensidade do desvio vocal. Em segundo lugar, dada a variedade de fatores etiológicos e de manifestações de um distúrbio de voz, nem sempre a imagem laríngea pode ser utilizada como referência para comprovar a ausência/presença de um distúrbio vocal, seja no diagnóstico inicial ou na avaliação pré e pós-intervenção, considerando-se também que um distúrbio de voz pode ser caracterizado por diferentes ajustes do trato vocal, não visualizados na laringoscopia.

Sabe-se que a avaliação perceptivo-auditiva da qualidade vocal apresenta um considerável grau de dificuldade, visto que o julgamento desses parâmetros depende de vários fatores subjetivos, como as referências internas de cada avaliador, escala utilizada e sua respectiva sensibilidade e especificidade, atenção e fadiga do ouvinte, entre outros^(1,4,5). No entanto, esse é o principal instrumento usado pelo fonoaudiólogo na avaliação da qualidade vocal.

Dessa forma, considerando a necessidade de desenvolvimento de mais estudos que investiguem a capacidade das medidas acústicas em determinar a intensidade do desvio vocal, e devido ao número reduzido de pesquisas que tomam como padrão de referência os dados da análise perceptivo-auditiva para identificar a presença ou ausência

de um distúrbio de voz^(2,4-6), o objetivo deste estudo foi analisar a correlação entre medidas acústicas e a intensidade do desvio vocal, bem como o poder discriminatório dessas medidas na detecção da presença de alteração vocal, na classificação da intensidade do desvio e na diferenciação do tipo de voz predominante.

MÉTODOS

Desenho do estudo

Este é um estudo quantitativo, explicativo, de campo e transversal, avaliado e aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal da Paraíba (UFPB), com o parecer de nº 52492/12. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido, autorizando a pesquisa.

Amostra

Participaram desta pesquisa 186 pacientes disfônicos, de ambos os gêneros, com a idade entre 19 e 60 anos, sendo 116 do gênero feminino e 70 do gênero masculino, população atendida no Laboratório de Voz do Departamento de Fonoaudiologia da UFPB, no período compreendido entre agosto de 2012 e março de 2013.

Estavam aptos a participar do estudo sujeitos com idade superior a 18 anos e inferior a 65 anos, apresentando queixa vocal e que tivessem passado por avaliação laringológica, com laudo descrevendo o diagnóstico laríngeo. Foram excluídos deste estudo indivíduos com problemas cognitivos ou neurológicos que impedissem a realização dos procedimentos de gravação. O grupo de pacientes incluiu indivíduos com condição laríngea normal, com lesões benignas de prega vocal (nódulos, cisto, sulco e pólipos), com disфонia por tensão muscular primária e com paralisia unilateral de prega vocal.

Material

Para a gravação das vozes, utilizou-se um *notebook* da marca HP, um microfone *headset* da marca *Logitech* e *software FonoView* versão 4.6h, da CTS Informática. A taxa de amostragem utilizada foi de 44.100 Hz.

Procedimentos

A coleta dos dados foi realizada em um ambiente silencioso, com ruído ambiental inferior a 50 dBNPS, aferido por meio de medidor de nível de pressão sonora digital. O microfone estava situado a uma distância média de 10 cm da comissura labial.

As sessões de gravação foram realizadas no momento da avaliação inicial do paciente, antes da realização da terapia vocal, e tinham duração média de cinco minutos, solicitando-se a emissão da vogal sustentada /e/ em tempo máximo de fonação.

Posteriormente, as vozes foram editadas no *software Sound Forge* versão 10.0, sendo eliminados os dois segundos iniciais e finais da emissão da vogal, devido à maior irregularidade nesses trechos, preservando-se o tempo mínimo de três segundos para cada emissão. A normalização foi realizada no controle “*normalize*” do *Sound Forge*, no modo *peak level*, a fim de obter uma padronização na saída de áudio entre -6 e 6 dB.

Para a análise perceptivo-auditiva da voz, foi utilizada uma escala analógica visual (EAV), com uma métrica de 0 a 100 mm, avaliando-se o grau geral do desvio vocal (GG), o grau de rugosidade (GR), de sopro (GS), de tensão (GT) e de instabilidade (GI). A marcação mais próxima do 0 representa menor alteração e, quanto mais próxima do 100, maiores são as alterações. Essa avaliação foi executada por meio de consenso, com a participação de três fonoaudiólogos especialistas em voz, com experiência em avaliação vocal perceptivo-auditiva.

A sessão de avaliação perceptiva ocorreu em ambiente silencioso. Inicialmente, os juízes foram orientados de que as vozes deveriam ser consideradas saudáveis quando fossem socialmente aceitáveis, produzidas de forma natural, sem esforço, ruído ou condição estável durante a emissão. Eles também foram instruídos de que a rugosidade corresponderia à presença de irregularidade vibratória, a sopro estaria relacionada ao escape de ar audível na emissão, a tensão corresponderia à percepção de esforço vocal e a instabilidade seria identificada pela presença de qualidade vocal, frequência e/ou intensidade flutuante ao longo da emissão. Além disso, os juízes foram treinados com estímulos-âncora, contendo emissões saudáveis e com desvio nos diferentes graus, assim como vozes predominantemente rugosas, soprosas, tensas e instáveis.

Para avaliação, cada emissão da vogal sustentada foi apresentada por três vezes através de caixa de som, em intensidade confortável autorreferida pelos avaliadores. Em seguida, faziam a identificação da presença ou ausência de desvio vocal, do tipo de voz predominante nas vozes desviadas (rugosa, sopro, tensa ou instável) e, por fim, o julgamento da intensidade do desvio (GG, GR, GS, GT e GI).

No final da sessão de avaliação perceptiva, 10% das amostras foram repetidas aleatoriamente, para a análise da confiabilidade da avaliação por consenso dos juízes por meio do Coeficiente Kappa de Cohen. O valor de Kappa foi de 0,80, indicando uma boa concordância entre os avaliadores. A confiabilidade intra-avaliador teve um coeficiente Kappa de 0,79, indicando boa concordância.

Posteriormente, foi realizada uma correspondência da escala numérica (EN) para a EAV, sendo o grau 1 (0–35,5 mm) relacionado às variações normais da qualidade vocal, grau 2 (35,6–50,5 mm) desvio leve a moderado, grau 3 (50,6–90,5 mm) desvio moderado e grau 4 (90,6–100 mm) desvio intenso⁽⁷⁾.

A extração das medidas acústicas foi realizada no *software VoxMetria*, versão 4.7h, da CTS Informática, no módulo análise de qualidade vocal. Para essa avaliação, foram utilizadas as medidas de média e desvio padrão da frequência fundamental (F_0), *jitter*, *shimmer* e *Glottal to Noise Excitation Ratio* (GNE) da vogal sustentada. No *software* utilizado, os valores de referência para os parâmetros de desvio padrão de F_0 , *jitter*, *shimmer* e GNE são 0,2 Hz, 0,6, 6,5 e 0,5%, respectivamente. No caso do desvio padrão de F_0 , *jitter* e *shimmer*, valores maiores que os citados já são considerados alterados. Inversamente, para o GNE, valores menores que 0,5 são categorizados como alterados.

Análise dos dados

Realizou-se a análise estatística descritiva para todas as variáveis analisadas, além de análise estatística inferencial de correlação, com o teste de correlação de Spearman, para correlacionar a intensidade do desvio vocal (saudável, leve, moderado e intenso) e as medidas acústicas.

Os coeficientes de correlação são utilizados para avaliar e quantificar o grau de relacionamento linear entre duas variáveis, observando se as variáveis se modificam conjuntamente e em que grau. O coeficiente de correlação varia de -1 a 1, sendo que os valores negativos indicam que as variáveis se comportam de modo inversamente proporcional, enquanto valores positivos indicam que elas variam proporcionalmente.

Para a classificação dos coeficientes de correlação, adotou-se, nesta pesquisa, que valores de 0,1 a 0,3 representam uma correlação fraca; entre 0,4 e 0,6 indicam correlação moderada; e acima de 0,7 é possível afirmar que o grau de correlação entre as variáveis é forte⁽⁸⁾.

Foi utilizada a ANOVA para comparar a análise das medidas acústicas de acordo com o grau de desvio vocal e tipo de voz predominante, com análise *post hoc*, utilizando-se o teste de Scheffé.

Para a categorização das medidas acústicas quanto ao seu poder discriminatório em relação à presença ou ausência de desvio vocal, foram consideradas úteis as medidas cujas médias apresentavam diferenças entre saudável (grau 1) e desviada (graus 2, 3 e 4).

Quanto à classificação da intensidade do desvio vocal (GG, GR, GS, GT e GI), foram consideradas úteis as medidas acústicas cujas médias apresentavam diferenças entre grau 1 x grau 2, grau 2 x grau 3 e grau 3 x grau 4.

Todas as análises foram realizadas pelo *software Statistical Package for the Social Sciences* (SPSS) versão 2.0 e o nível de significância adotado foi de 0,05.

RESULTADOS

A ocorrência de vozes saudáveis foi de 17,2% (n=32), 55,4% (n=103) eram vozes com desvio leve, 24,2% (n=45) eram vozes com desvio moderado e 3,2% (n=6) correspondiam a vozes com desvio intenso da qualidade vocal (Tabela 1).

Quanto ao tipo de voz predominante, a rugosidade (51,3%, n=80) foi o parâmetro predominante nos sujeitos com presença de desvio da qualidade vocal, seguida pela sopro (28,8%, n=45), tensão (12,8%, n=20) e instabilidade (7,1%, n=11) (Tabela 2).

A média de F_0 teve correlação negativa com o grau de rugosidade ($p=0,019$) e positiva com o grau de tensão ($p=0,03$) (Tabela 3). O desvio padrão de F_0 , por outro lado, apresentou correlação positiva com os valores de GG ($p<0,001$), GR ($p<0,001$), GT ($p<0,001$) e GI ($p<0,001$) (Tabela 3).

Houve correlação positiva entre o *jitter* e GG ($p<0,001$), GR ($p<0,001$), GS ($p=0,014$), GT ($p<0,001$) e GI ($p<0,001$). Quanto ao *shimmer*, a mesma correlação foi encontrada entre os seus valores e o GG ($p<0,001$), o GR ($p<0,001$), o GS ($p=0,027$), o GT ($p<0,001$) e o GI ($p<0,001$) (Tabela 3).

O GNE apresentou correlação negativa com GG ($p<0,001$), GR ($p<0,001$), GS ($p<0,001$) e GI ($p=0,01$) (Tabela 3).

Na comparação entre grupos em relação à intensidade do desvio vocal (GG), constatou-se diferença para as variáveis DP de F_0 ($p<0,001$), *jitter* ($p<0,001$), *shimmer* ($p<0,001$) e GNE ($p<0,001$) (Tabela 4). Porém, na análise *post hoc*, apenas o GNE diferenciou vozes saudáveis (grau 1) de vozes desviadas nos graus 2 ($p=0,090$), 3 ($p<0,001$) e 4 ($p=0,062$).

Comparando-se os grupos em relação à rugosidade, houve diferença em todos os parâmetros (Tabela 4). Na análise *post hoc*, o *shimmer* foi a única medida capaz de separar vozes rugosas de não

Tabela 1. Distribuição dos parâmetros vocais de acordo com o grau geral e os graus de rugosidade, sopro, tensão e instabilidade

Variável	Intensidade do desvio vocal								EAV Média±DP
	VNQV		Leve a moderado		Moderado		Intenso		
	n	%	n	%	n	%	n	%	
GG	32	17,2	103	55,4	45	24,2	6	3,2	44,94±15,79
GR	75	40,3	81	43,5	27	14,5	3	1,6	36,02±20,54
GS	86	46,2	79	42,5	21	11,3	–	–	31,52±17,16
GT	138	74,2	31	16,7	14	7,5	3	1,1	21,36±20,44
GI	125	67,2	49	26,3	11	5,9	1	0,5	25,86±18,88

Legenda: VNQV = variabilidade normal da qualidade vocal; EAV = escala analógico-visual; GG = grau geral; GR = grau de rugosidade; GS = grau de sopro; GT = grau de tensão; GI = grau de instabilidade; DP = desvio padrão

Tabela 2. Tabela descritiva do tipo de voz predominante nos pacientes com emissão vocal desviada

Tipo de voz	n	%
Rugosa	80	51,3
Soprosa	45	28,8
Tensa	20	12,8
Instável	11	7,1
Total	156	100

rugosas, ou seja, capaz de diferenciar o grau 1, do grau 2 ($p<0,05$), 3 ($p<0,001$) e 4 ($p<0,001$).

Na comparação entre grupos, o parâmetro de sopro apresentou diferença entre as médias dos valores de *jitter* ($p=0,004$) e GNE ($p<0,001$) (Tabela 4). A análise *post hoc* demonstrou que o GNE foi útil para classificar o grau de sopro presente e para distinguir vozes soprosas de não soprosas, apresentando diferença entre grau 1 e grau 2 ($p<0,001$), grau 1 e grau 3 ($p<0,001$) e grau 2 e grau 3 ($p=0,006$). Quanto à sopro, o GNE entrou nos critérios de elegibilidade adotados neste estudo, pois não houve nenhum indivíduo com desvio intenso quanto à sopro entre os voluntários pesquisados.

Quanto ao parâmetro de tensão, a comparação entre grupos mostrou haver diferenças entre esse parâmetro e a média de F_0 ($p=0,006$), DP de F_0 ($p<0,001$), *jitter* ($p<0,001$), *shimmer* ($p<0,001$) e GNE ($p=0,032$) (Tabela 4). A análise *post hoc* mostrou que a média de F_0 foi útil para determinar o grau de tensão fonatória, separando vozes sem componente de tensão (grau 1) de vozes desviadas em grau leve ($p=0,039$), vozes com desvio leve de vozes com desvio moderado ($p=0,008$), e vozes com desvio moderado de vozes com desvio intenso ($p=0,002$). O DP F_0 separou vozes saudáveis quanto à tensão de vozes tensas nos graus 2 ($p=0,009$), 3 ($p=0,002$) e 4 ($p<0,001$).

Na comparação entre grupos em relação à instabilidade, houve diferenças entre as médias dos valores de DP de F_0 ($p<0,001$), *jitter* ($p<0,001$), *shimmer* ($p<0,001$) e GNE ($p=0,006$) (Tabela 4). Mas, na análise *post hoc*, nenhuma das medidas cumpriu os critérios estabelecidos para ser útil no diagnóstico ou monitoramento de alterações vocais.

Comparando-se os grupos em relação ao tipo de qualidade vocal predominante nas emissões desviadas, houve diferença entre os valores da média de F_0 ($p<0,001$) e GNE ($p=0,039$) (Tabela 5). Na análise *post hoc*, a média de F_0 separou vozes rugosas de soprosas ($p<0,001$), com maiores valores da média de F_0 em vozes soprosas. Houve diferenciação também entre vozes rugosas e tensas ($p=0,002$), com maiores valores da média de F_0 em vozes tensas do que nas rugosas.

DISCUSSÃO

A associação das avaliações perceptivo-auditivas e acústicas é de fundamental importância para se ter acesso à qualidade vocal, ao grau de deterioração vocal e aos resultados com o tratamento terapêutico e/ou cirúrgico⁽⁹⁾.

As medidas acústicas demonstram-se capazes de identificar desvios na qualidade vocal, esperando-se que elas apresentem uma relação com as anormalidades laríngeas e com a intensidade do desvio vocal⁽⁶⁾.

Nesse contexto, o objetivo do presente estudo foi analisar a correlação entre medidas acústicas e a intensidade do desvio vocal e o poder discriminatório dessas medidas na detecção da presença de alteração vocal, na classificação da intensidade do desvio e na diferenciação do tipo de voz predominante.

Foi encontrada uma correlação fraca entre a média de F_0 e os dados perceptivo-auditivos, demonstrando que os pacientes com

Tabela 3. Correlação entre a intensidade do desvio vocal e os diferentes parâmetros acústicos

Variáveis	Intensidade do desvio vocal									
	GG		GR		GS		GT		GI	
	Correlação	Valor de p	Correlação	Valor de p	Correlação	Valor de p	Correlação	Valor de p	Correlação	Valor de p
Média F_0	0,67	0,37	-0,17	0,019*	0,06	0,46	0,16	0,03*	0,03	0,71
DP F_0	0,48	<0,001*	0,31	<0,001*	-0,01	0,86	0,49	<0,001*	0,37	<0,001*
<i>Jitter</i>	0,66	<0,001*	0,53	<0,001*	0,18	0,014*	0,47	<0,001*	0,46	<0,001*
<i>Shimmer</i>	0,68	<0,001*	0,56	<0,001*	0,16	0,027*	0,49	<0,001*	0,55	<0,001*
GNE	-0,44	<0,001*	-0,39	<0,001*	-0,55	<0,001*	-0,08	0,261	-0,19	0,01*

*Valores significativos ($p<0,05$) – teste de correlação de Spearman

Legenda: GG = grau geral; GR = grau de rugosidade; GS = grau de sopro; GT = grau de tensão; GI = grau de instabilidade; F_0 = frequência fundamental; DP F_0 = desvio padrão da frequência fundamental; GNE = *Glottal to Noise Excitation Ratio*

Tabela 4. Comparação das medidas acústicas em função da intensidade do desvio vocal

Variáveis	VNQV		Leve a moderado		Moderado		Intenso		Valor de p
	Média	DP	Média	DP	Média	DP	Média	DP	
GG									
DP F ₀	5,43	20,01	4,89	9,33	19,13	34,56	74,63	74,05	<0,001*
Jitter	0,193	0,182	0,576	1,356	2,44	3,07	9,53	5,97	<0,001*
Shimmer	4,03	1,65	6,39	3,96	12,95	10,23	30,85	18,06	<0,001*
GNE	0,867	0,127	0,774	0,164	0,58	0,22	0,655	0,169	<0,001*
GR									
Média F ₀	189,74	41,8	178,71	41,53	180,86	50,68	104,51	22,42	0,006*
DP F ₀	5,98	15,62	7,39	21,09	34,06	50,16	21,92	30,29	<0,001*
Jitter	0,46	1,47	0,88	1,98	3,74	3,59	10,07	6,98	<0,001*
Shimmer	4,85	2,84	7,93	7,81	17,91	10,97	26,58	8,88	<0,001*
GNE	0,81	0,17	0,75	0,18	0,54	0,21	0,58	0,14	<0,001*
GS									
Jitter	1,03	2,98	0,99	1,69	3,18	3,81			0,004*
GNE	0,84	0,14	0,68	0,19	0,54	0,19			<0,001*
GT									
Média F ₀	177,44	41,17	202,65	40,79	166,76	47,82	272,2	31,33	0,006*
DP F ₀	5,11	12,74	16,15	35,24	29,56	33,81	127,34	66,74	<0,001*
Jitter	0,73	1,57	1,25	2,59	4,87	5,16	8,98	6,28	<0,001*
Shimmer	6,72	4,52	8,13	6,8	19,61	15,61	35,12	26,11	<0,001*
GNE	0,75	0,18	0,76	0,23	0,59	0,26	0,73	0,19	0,032*
GI									
DP F ₀	6,8	17,45	10,92	27,95	37,58	48,78			<0,001*
Jitter	0,58	1,13	2,06	3,74	4,14	4,87			<0,001*
Shimmer	6,05	4,04	40,43	7,71	22,41	20,74			<0,001*
GNE	0,77	0,18	0,69	0,22	0,62	0,21			0,006*

*Valores significativos (p<0,05) – ANOVA

Legenda: VNQV = variabilidade normal da qualidade vocal; DP = desvio padrão; GG = grau geral; DP F₀ = desvio padrão da frequência fundamental; GNE: *Glottal to Noise Excitation Ratio*; GR = grau de rugosidade; GS = grau de soproidade; GT = grau de tensão; GI = grau de instabilidade

Tabela 5. Comparação das medidas acústicas em função do tipo de voz predominante

Variáveis	Rugosidade		Soproidade		Tensão		Instabilidade		Valor de p
	Média	DP	Média	DP	Média	DP	Média	DP	
Média F ₀	166,13	43,66	198,48	36,43	200,09	41,95	187,17	40,4	<0,001*
GNE	0,73	0,19	0,69	0,2	0,81	0,19	0,81	0,15	0,039*

*Valores significativos (p<0,05) – ANOVA

Legenda: DP = desvio padrão; F₀ = frequência fundamental; GNE = *Glottal to Noise Excitation Ratio*

maior grau de rugosidade apresentaram vozes mais graves, enquanto pacientes com vozes com maior grau de tensão fonatória apresentaram vozes mais agudas.

A correlação negativa entre a média de F₀ e a rugosidade pode ser justificada pelo fato de que a presença de rugosidade está diretamente relacionada à irregularidade vibratória da mucosa das pregas vocais, que pode advir de pequenos edemas, vasodilatações, fadiga, mau uso e abuso vocal⁽¹⁰⁾. Desse modo, lesões como nódulos, pólipos e edemas, que ocasionam aumento de massa nas pregas vocais e, concomitantemente, irregularidade vibratória, podem abaixar a frequência fundamental, tornando a voz mais grave^(11,12).

Quanto à correlação entre a média de F₀ e a tensão fonatória, deve-se considerar que a frequência é determinada, entre outros fatores, pela tensão das pregas vocais, a qual é controlada pelos músculos intrínsecos da laringe, especificamente, o cricoido⁽¹³⁾. Por isso, pacientes com tensão fonatória costumam realizar maior contração da musculatura extrínseca e intrínseca, inclusive com

maior tensão longitudinal das pregas vocais, maior pressão subglótica e maior constrição no trato vocal, gerando um maior número de ciclos glóticos por segundo e, conseqüentemente, uma frequência fundamental mais elevada⁽¹⁴⁾.

Embora, na literatura⁽¹⁵⁾, encontra-se que o grau geral da alteração e a rugosidade são os parâmetros mais relacionados à F₀, no presente estudo, a tensão fonatória e a rugosidade apresentaram-se mais relacionadas à F₀. Resultado semelhante também foi encontrado em pesquisas^(16,17) com crianças, sendo a tensão fonatória, seguida do grau geral e de rugosidade, os parâmetros mais relacionados à F₀.

Nesta pesquisa, o desvio padrão de F₀ apresentou uma moderada correlação positiva com o grau geral do desvio vocal e o grau de tensão fonatória, assim como uma fraca correlação positiva com o grau de rugosidade e grau de instabilidade. As vozes com desvio mais intensos e com maior grau de rugosidade, tensão e instabilidade apresentaram maior valor do desvio padrão de F₀.

Em termos fisiológicos, o desvio padrão de F_0 está diretamente ligado à condição neuromuscular e à regularidade de vibração da mucosa das pregas vocais. Em termos acústicos e perceptivos, tem uma relação direta com a periodicidade do sinal sonoro. Desse modo, considerando-se que alterações na estrutura histológica das pregas vocais interferem nos padrões de vibração glótica, principalmente na onda mucosa, ocasionando desvio na produção vocal, justifica-se a correlação entre os parâmetros perceptivos e o desvio padrão de F_0 (11,15).

Os valores de *jitter* e *shimmer* correlacionaram-se positivamente com todos os parâmetros da análise perceptivo-auditiva, apresentando-se maiores nas vozes com desvios mais intensos.

Alguns estudos (18,19) referem que o *jitter* reflete a intensidade global do desvio vocal, sendo uma medida sensível à presença de desvio da qualidade vocal, o que justifica o fato de que seus valores sejam mais elevados em vozes mais alteradas, considerando-se todos os parâmetros perceptivos.

Pesquisas (6,20-22) que combinam um número de parâmetros acústicos com dados do exame laríngeo sugerem que o *jitter* e *shimmer* podem ser fortes preditores de alteração vocal, podendo detectar alterações vocais sutis, que poderiam passar despercebidas na análise perceptiva ou laríngea.

Estudos (18,19) compararam pacientes pré e pós-terapia vocal, observando-se que existe uma correlação moderada entre os parâmetros acústicos (*jitter*, *shimmer* e proporção harmônico-ruído) e a análise perceptiva. A maior correlação foi entre o grau geral e o *jitter* e *shimmer*.

Na presente pesquisa, a média dos valores de desvio padrão de F_0 , *jitter*, *shimmer* e GNE apresentou diferenças entre os graus de desvio vocal. Porém, na análise *post hoc*, e considerando os critérios de elegibilidade para definir uma medida como útil para categorizar uma voz como saudável ou alterada e para classificar a intensidade do desvio, o GNE foi a única medida útil para diferenciar vozes saudáveis de vozes desviadas. Não houve nenhuma medida útil para classificar a intensidade do desvio vocal.

Quanto ao grau de rugosidade, houve diferenças entre as médias de todas as medidas acústicas estudadas. Mas apenas o *shimmer* foi útil para categorizar vozes rugosas de vozes não rugosas. Não houve uma medida útil para classificar a intensidade da rugosidade presente na emissão vocal.

De modo geral, as medidas de perturbação (*jitter* e *shimmer*) são usadas para descrever a rugosidade encontrada na avaliação perceptiva e a irregularidade vibratória no plano fisiológico, assim como as medidas de ruído são utilizadas como indicadoras de sopro e de fechamento glótico inadequado (6,18,23-27). Um estudo demonstrou que o parâmetro auditivo de rugosidade estava relacionado ao *shimmer* (28).

Em termos do grau de sopro, embora as médias dos valores de *jitter* e GNE tenham se apresentado diferentes, na análise *post hoc*, e conforme os critérios estabelecidos neste trabalho, o GNE foi eficaz para classificar a intensidade de sopro presente na emissão vocal e para diferenciar vozes soprosas de não soprosas.

O GNE mede o ruído adicional no sinal sonoro, independentemente do ruído modulado pelo mecanismo glótico. Indica a origem do sinal vocal, se ele é proveniente da vibração das pregas vocais ou da corrente de ar turbulenta gerada no trato vocal. Ele é capaz de mostrar valores diferentes em ajustes fonatório distintos, em diferentes alterações vocais (23).

O GNE pode ser considerado uma medida mais robusta porque, diferentemente do *jitter* e *shimmer*, seu cálculo não requer uma estimativa anterior da frequência fundamental, o que é uma tarefa difícil na presença de grandes alterações laríngeas e vocais (23-25).

A literatura (23) faz referências à aplicação das medidas de ruído para avaliação e triagem da voz, justamente pelo fato de sempre apresentarem boa correlação com a avaliação perceptiva, como foi encontrado no presente estudo. Pesquisas (23,26) que combinaram vários parâmetros acústicos para a descrição de vozes alteradas demonstraram que o GNE foi a medida mais importante para uma análise independente, possibilitando discriminar vozes normais e desviadas.

O GNE está diretamente relacionado à presença de rugosidade e sopro, que, por sua vez, estão entre os parâmetros universalmente mais confiáveis da avaliação perceptiva. Dessa forma, pelos dados da presente pesquisa, pode-se inferir que o GNE é útil para diagnosticar a presença de alteração vocal, para detectar a presença de sopro e para classificar a intensidade de sopro presente na emissão.

No parâmetro perceptivo de tensão fonatória, houve diferença na média de todas as medidas acústicas. Na análise *post hoc* e com base nos critérios previamente adotados, a média de F_0 foi útil para classificar a intensidade de tensão presente na produção vocal, enquanto o desvio padrão de F_0 separou vozes saudáveis de vozes com tensão fonatória.

O aumento de tensão pode gerar desequilíbrio no sistema e, por consequência, maior dificuldade no controle da emissão. Essa falta de controle na emissão vocal faz com que a frequência fundamental oscile, aumentando os valores do desvio padrão de F_0 .

Uma pesquisa (29) utilizando laringes excisadas investigou as variações de pressão subglótica, comparando análise da dinâmica não linear com análise das medidas de perturbação. Um aumento considerável da pressão subglótica gerou vibração irregular, bifurcação, emissão rugosa e com esforço. Dessa forma, excesso de tensão fonatória, seja com o aumento da pressão subglótica ou por constrições no trato vocal, também pode gerar irregularidade na vibração das pregas vocais, refletidos nos valores de DP de F_0 .

Quanto à instabilidade fonatória, houve diferença nas médias do desvio padrão de F_0 , *jitter*, *shimmer* e GNE. No entanto, na análise *post hoc*, nenhuma das medidas enquadrou-se nos critérios de elegibilidade para classificação ou categorização do desvio vocal.

Por fim, comparando-se as medidas acústicas em relação ao tipo de qualidade vocal predominante nas emissões desviadas, observou-se que as vozes rugosas apresentaram-se diferenciadas das vozes soprosas e tensas, com relação aos valores de F_0 , sendo as vozes rugosas mais graves em relação às outras duas. Fisiologicamente, esse comportamento pode ser justificado pelos seguintes fatores: aumento de contração da musculatura intrínseca e extrínseca nas vozes tensas, ocasionando maior rigidez no sistema e maior pressão longitudinal nas pregas vocais, e diminuição do contato entre as pregas vocais, com menor quantidade de massa em vibração nos casos de fendas glóticas subjacentes às vozes soprosas, ambas as condições gerando uma emissão mais aguda. Por outro lado, o componente de rugosidade está diretamente associado à presença de lesão de massa na borda livre da prega vocal, o que proporciona uma voz mais grave (12).

Na pesquisa e na clínica vocal, o maior desafio é compreender qual recurso fornece a melhor avaliação e a correlação de cada

medida acústica no plano perceptual e fisiológico^(23,26). A partir dos dados levantados nesta pesquisa, observou-se que há correlação entre as medidas acústicas e perceptivo-auditivas na quantificação do desvio vocal. Quanto à utilidade dos parâmetros acústicos na determinação da presença ou ausência de uma alteração vocal, constatou-se que, em termos de diagnóstico, apenas o GNE foi útil para identificar vozes saudáveis e alteradas e para distinguir vozes soprosas de não soprosas, o *shimmer* identificou a presença/ausência de rugosidade e o desvio padrão de F_0 diferenciou vozes com e sem o componente de tensão.

Com relação à finalidade das medidas acústicas em prever a intensidade do desvio vocal, o GNE foi capaz de classificar o grau de sopro da voz, assim como a média de F_0 foi útil para classificar o grau de tensão fonatória.

De forma geral, as medidas que categorizam uma voz em saudável ou alterada, principalmente quanto ao grau geral, como o GNE na presente pesquisa, podem ser usadas em procedimentos de triagem e de diagnóstico de distúrbios da voz, enquanto as medidas capazes de classificar a intensidade do desvio podem ser úteis para o monitoramento do distúrbio de voz ao longo da terapia vocal.

CONCLUSÃO

Há correlação entre as medidas acústicas e perceptivo-auditivas. O GNE é útil para diagnosticar vozes saudáveis e alteradas, para identificar a presença e classificar o grau de sopro na emissão. O *shimmer* pode ser utilizado para diagnosticar a presença de rugosidade na voz. A média de F_0 é útil para classificar a intensidade de tensão fonatória e o desvio padrão de F_0 pode ser utilizado para diagnosticar a presença de tensão na emissão vocal. A média de F_0 diferenciou vozes rugosas, soprosas e tensas, sendo as vozes rugosas mais graves em relação às outras duas.

**LWL, DPC e POC contribuíram na construção e desenvolvimento deste trabalho. LWL realizou a coleta de dados, análise dos resultados e revisão final do manuscrito; DPC realizou a coleta de dados e análise dos resultados; POC realizou a tabulação dos dados, análise estatística e revisão do manuscrito.*

REFERÊNCIAS

- 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. *Eur Arch Otorhinolaryngol*. 2001;258(2):77-82.
- Roy N, Barkmeier-Kraemer J, Eadie T, Sivasankar MP, Mehta D, Paul D, et al. Evidence-based clinical voice assessment: a systematic review. *American Journal of Speech-Language Pathology* [cited 2013 Set 13]. Available from: <http://ajslp.asha.org/papbyrecent.dtl>
- Behrman A. Common practices of voice therapists in the assessment of patients. *J Voice*. 2005;19(3):454-69.
- Maryn Y, Roy N, De Bodt M, Cauwenberge PV, Corthals P. Acoustic measurement of overall voice quality: a meta-analysis. *J Acoust Soc Am*. 2009;126(5):2619-34.
- Martens JW, Versnel H, Dejonckere PH. The effect of visible speech in the perceptual rating of pathological voices. *Arch Otolaryngol Head Neck Surg*. 2007;133(2):178-85.
- Brockmann M, Drinnan MJ, Storck C, Carding P. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *J Voice*. 2011;25(1):44-53.
- Yamasaki R, Leão S, Madazio G, Padovani M, Azevedo R, Behlau M. Correspondência entre escala analógico visual e a escala numérica na avaliação perceptivo-auditiva de vozes. In: 16º Congresso Brasileiro de Fonoaudiologia; 2008 Set. 24-27; Campos do Jordão: Sociedade Brasileira de Fonoaudiologia; 2008.
- Dancey C, Reidy J. Estatística sem matemática para psicologia: usando SPSS para Windows. Porto Alegre: Artmed; 2006.
- Uloza V, Verikas A, Bacauskiene M, Gelzini A, Pribisiene R, Kaseta M, et al. Categorizing normal and pathological voices: automated and perceptual categorization. *J Voice*. 2011;25(6):700-8.
- McAllister A, Sederholm E, Sundberg J. Perceptual and acoustic analysis of vocal registers in 10-year-old children. *Logoped Phoniatr Vocol*. 2000;25(2):63-71.
- Jotz GP, Cervantes O, Settani FAP, Angelis EC. Acoustic measures for the detection of hoarseness in children. *Int Arch Otorhinolaryngol*. 2006;10(1):14-20.
- Pontes PAL, Vieira VP, Gonçalves MIR, Pontes AAL. Características das vozes roucas, ásperas e normais: análise acústica espectrográfica comparativa. *Rev Bras Otorrinolaringol*. 2002;68(2):182-8.
- Lind K, Wermke K. Development of the vocal fundamental frequency of spontaneous cries during the first 3 months. *Int J Pediatr Otorhinolaryngol*. 2002;64(2):97-104.
- Van Houte E, Van Lierde K, Claeys S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J Voice*. 2011;25(2):202-7.
- McAllister A, Sederholm E, Ternström S, Sundberg J. Perturbation and hoarseness: a pilot study of six children's voices. *J Voice*. 1996;10(3):252-61.
- Lee EK, Son YI. Muscle tension dysphonia in children: voice characteristics and outcome of voice therapy. *Int J Pediatr Otorhinolaryngol*. 2005;69(7):911-7.
- Lopes LW, Lima ILB, Almeida LNA, Cavalcante DP, Almeida AAF. Severity of voice disorders in children: correlations between perceptual and acoustic data. *J Voice*. 2012;26(6):819.e7-12.
- 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.
- Ma EPM, Yu EML. Multiparametric evaluation of dysphonic severity. *J Voice*. 2005;20(3):380-90.
- Zhang Y, Jiang JJ. Acoustic analysis of sustained and running voices from patients with laryngeal pathologies. *J Voice*. 2008;22(1):1-9.
- Ortega J, Cassinello N, Dorcatto D, et al. Computerized acoustic voice analysis and subjective scaled evaluation of the voice can avoid the need for laryngoscopy after thyroid surgery. *Surgery*. 2009;145(3):265-71.
- Gelzini A, Verikas A, Bacauskiene M. Automated speech analysis applied to laryngeal disease categorization. *Comput Methods Programs Biomed*. 2008;91(1):36-47.
- Godino-Llorente JJ, Osma-Ruiz V, Sáenz-Lechón N, Vilda-Gómez P, Blanco-Velasco M, Cruz-Roldán F. The effectiveness of the glottal to noise excitation ratio for the screening of voice disorders. *J Voice*. 2010;24(1):47-56.
- Madazio G, Leão S, Behlau M. The phonatory deviation diagram: a novel objective measurement of vocal function. *Folia Phoniatr Logop*. 2011;63(6):305-11.
- Fröhlich M, Michaelis D, Strube HW, Kruse E. Acoustic voice analysis by means of the hoarseness diagram. *J Speech Lang Hear Res*. 2000;43(3):706-20.
- Michaelis D, Fröhlich M, Strube HV. Selection and combination of acoustic features for the description of pathologic voices. *J Acoust Soc Am*. 1998;103(3):1628-39.
- Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice*. 2006;20(4):527-44.
- 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.
- Jiang JJ, Zhang Y, Ford CN. Nonlinear dynamics of phonations in excised larynx experiments. *J Acoust Soc Am*. 2003;114(4 Pt 1):2198-205.