

Priscila Campos Martins¹
Thalita Evaristo Couto¹
Ana Cristina Côrtes Gama²

Descritores

Fonoaudiologia
Escala Visual Analógica
Voz
Disfonia
Distúrbios da Voz

Keywords

Speech, Language and Hearing Sciences
Visual Analog Scale
Voice
Dysphonia
Voice Disorders

Endereço para correspondência:

Priscila Campos Martins
Avenida Professor Alfredo Balena, 190,
Sala 251, Belo Horizonte (MG), Brasil,
CEP: 30130-100.
E-mail: priscila.fonoaudiologia@gmail.com

Recebido em: 05/09/2014

Aceito em: 30/01/2015

Avaliação perceptivo-auditiva do grau de desvio vocal: correlação entre escala visual analógica e escala numérica

Auditory-perceptual evaluation of the degree of vocal deviation: correlation between the Visual Analogue Scale and Numerical Scale

RESUMO

Objetivo: Analisar a concordância intra e interavaliadores para as escalas visual analógica e numérica na tarefa de vogal sustentada e determinar pontos de corte numéricos da escala visual analógica correspondentes aos graus da escala numérica. **Métodos:** Foram selecionadas 205 amostras da tarefa da vogal /a/ sustentada de modo habitual. Seis especialistas em voz avaliaram o grau geral de desvio vocal pela escala visual analógica e, após dois dias, pela escala numérica. Foram comparados os resultados encontrados nas avaliações pelas duas escalas e analisadas as concordâncias intra e interavaliadores, a correlação entre as escalas e a estimativa dos pontos de corte utilizando-se respectivamente os coeficientes de correlação intraclass e de concordância Kappa, o coeficiente de Spearman e a Análise de Variância, os valores de sensibilidade e especificidade. **Resultados:** Houve elevada correlação entre as escalas. Os valores de corte numéricos encontrados para a escala visual analógica correspondentes aos graus da escala numérica foram: neutro (grau zero) – 0 a 34 mm, leve (grau um) – 34,1 a 51 mm, moderado (grau dois) – 51,1 a 63,5 mm, intenso (grau três) – 63,6 a 77,5 mm e extremo (grau quatro) – acima de 77,5 mm. **Conclusão:** As escalas visual analógica e numérica apresentaram uma alta correlação, sendo observada maior concordância intra e interavaliador na escala visual analógica. Foram encontrados valores de corte numéricos para a escala visual analógica. Essa correlação possibilita a comparação entre os resultados encontrados na avaliação do grau geral de desvio vocal pelas duas escalas, que são amplamente utilizadas em pesquisas e na rotina clínica fonoaudiológica.

ABSTRACT

Purpose: To analyze the intra- and inter-rater agreement for visual analog scale and numerical scale in task of sustained vowel and to determine numerical cutoff points to visual analog scale corresponding to the degrees of the numeric scale. **Methods:** We selected 205 samples of the usual task of the sustained vowel /a/. Six voice specialists rated the overall degree of vocal deviation, first by visual analog scale and, after two days, by the numeric scale. The results obtained by both scales were compared and the intra- and inter-rater agreement, the correlation between the scales, and the estimated cutoff points using the intraclass correlation and concordance Kappa coefficients, the Spearman coefficient, and analysis of variance, and the values of sensitivity and specificity were analyzed. **Results:** A strong correlation was observed between the scales. The following numerical cutoff values were found for visual analog scale corresponding to the numerical scale: neutral (degree zero) – 0 to 34 mm; mild (degree one) – 34.1 to 51 mm; moderate (degree two) – 51.1 to 63.5 mm; intense (degree three) – 63.6 to 77.5 mm; and extreme (degree four) – above 77.5 mm. **Conclusion:** The visual analog scale and numerical scale showed a strong correlation, being observed the greater intra- and inter-rater agreement in visual analog scale. Numerical cutoff values for visual analog scale were found. This correlation enables the comparison between the results found in the evaluation of the overall degree of vocal deviation by both scales, which are widely used in research and in the clinical speech therapy routine.

Trabalho realizado no Curso de Fonoaudiologia, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG).

(1) Curso de Fonoaudiologia, Faculdade de Medicina, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

(2) Departamento de Fonoaudiologia, Faculdade de Medicina, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

Fontes de financiamento: Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

Apesar de não existir um consenso quanto ao conceito de voz normal e voz alterada^(1,2), uma emissão considerada de boa qualidade pelos ouvintes e produzida sem desconforto pelo falante caracteriza uma voz normal, enquanto uma emissão considerada como ruidosa pelos ouvintes e produzida com certo desconforto pelo falante caracteriza uma disfonia⁽¹⁾.

A avaliação fonoaudiológica da voz é um instrumento eficaz para descrever o perfil vocal do indivíduo, caracterizar a qualidade vocal e quantificar o desvio vocal⁽³⁾. As formas de avaliação mais preconizadas na rotina clínica fonoaudiológica são as análises acústica e a perceptivo-auditiva da voz. A análise acústica quantifica o sinal sonoro, o que torna a análise vocal mais objetiva. A análise perceptivo-auditiva é uma avaliação subjetiva baseada na impressão auditiva que o avaliador tem sobre a emissão vocal do indivíduo, permitindo a inferência de dados anatomofisiológicos importantes e fornecendo informações sobre os aspectos psicossociais da voz⁽⁴⁾.

Apesar das críticas feitas à subjetividade e à imprecisão da terminologia envolvida nesse procedimento, a avaliação perceptivo-auditiva é tradicional na rotina clínica fonoaudiológica, considerada padrão ouro para a análise da qualidade vocal⁽³⁾. Para reduzir potencialmente a variabilidade e as inconsistências da análise perceptivo-auditiva, foram desenvolvidas diversas escalas de avaliação⁽⁴⁾, sendo que as escalas mais utilizadas são a *Consensus Auditory Perceptual Evaluation – Voice (CAPE-V)*^(5,6), cuja mensuração da gravidade do distúrbio vocal é realizada por meio de uma escala visual analógica (EVA), e a escala GRBAS^(7,8), em que a mensuração é realizada por meio de uma escala numérica (EN).

A EVA CAPE-V e a EN GRBAS são amplamente utilizadas para avaliação perceptivo-auditiva da voz tanto na pesquisa quanto na prática clínica fonoaudiológica. Na clínica fonoaudiológica, assim como em pesquisas, é interessante a comparação do resultado da avaliação vocal por meio dessas duas escalas, que utilizam tanto a tarefa de vogal sustentada quanto de fala automática e encadeada. Porém, as distintas formas de mensuração dessas duas escalas inviabilizam essa comparação. Sendo assim, efetuar a comparação entre a EVA e a EN, assim como analisar a correlação entre elas, determinando-se pontos de corte numéricos específicos, possibilitará a comparação entre essas duas escalas.

A EVA é composta por uma reta, geralmente com comprimento de 100 mm, em que os ouvintes devem marcar o ponto que corresponde à extensão da variação de uma dada característica, como o grau de desvio da voz, por exemplo^(4,9). A EVA é ancorada por dois descritores verbais para representar os extremos, sendo comumente ancorado por “ausência de dor ou alteração” (em 0 mm na reta) e “máximo de dor ou alteração” (em 100 mm na reta). Não são recomendados números ou descritores verbais em pontos intermediários da reta para evitar a aglomeração de pontuação em torno de um valor numérico preferido^(9,10). Amplamente utilizada na área da saúde por diferentes profissionais, a EVA é utilizada principalmente como instrumento de autopercepção do paciente quanto à dor ou ao incômodo^(4,9,11), mas também é usada para mensurar a gravidade de uma doença

ou alteração^(6,10). A EN é composta por um número específico de pontos equidistantes, como, por exemplo, uma escala de quatro pontos para indicar o grau de desvio da voz, onde zero equivale a uma voz sem desvio e três, a um desvio extremo da voz^(4,9). A EN é utilizada principalmente para mensurar a gravidade de uma doença ou alteração^(6,10). Uma pesquisa⁽¹⁰⁾ que comparou a EVA com a EN na avaliação da gravidade da rinite alérgica, apontou que a EVA representa melhor a gravidade do problema, uma vez que, por meio dela, o paciente apresenta melhor sua percepção da alteração. Além disso, os médicos e pacientes que participaram da pesquisa relataram que a utilização da EVA é mais rápida e fácil que a da EN.

A análise perceptivo-auditiva da voz é influenciada também por fatores como o tempo de treinamento e a experiência prévia do avaliador, bem como as orientações recebidas, o grau de desvio da qualidade vocal e a tarefa de fala utilizada^(1,3,12). As tarefas de vogal sustentada e fala encadeada são amplamente utilizadas na avaliação perceptivo-auditiva da voz. O comportamento vocal se diferencia substancialmente nessas duas tarefas de fala, levando a diferenças na avaliação perceptivo-auditiva do tipo e grau de desvio da voz⁽¹³⁾.

Outras pesquisas^(2,14) definiram, com grande similaridade entre elas, o ponto de corte numérico para a variabilidade normal da voz para a EVA correspondente ao grau zero da EN na avaliação perceptivo-auditiva de vozes a partir da tarefa de fala encadeada. Não foi encontrado na literatura nenhum estudo que houvesse analisado tal correspondência na tarefa de vogal sustentada. A primeira pesquisa supracitada⁽¹⁴⁾ foi realizada utilizando uma EVA genérica e a EN GRBAS. Outras pesquisas^(2,15) correlacionaram uma EVA a uma EN, ambas genéricas. Somente a terceira pesquisa supracitada⁽²⁾, além de encontrar o ponto de corte para a variabilidade normal da voz, definiu os valores de corte indicativos dos diferentes graus de desvio vocal para a EVA a partir dos valores de uma EN de quatro pontos. Na presente pesquisa, optou-se pela análise das EVAs e ENs genéricas, como nas pesquisas anteriores, uma vez que inclui as escalas CAPE-V e GRBAS, mas não se limita a elas. Além disso, a GRBAS é uma escala de quatro pontos e, no presente estudo, optou-se por uma EN de cinco pontos, conforme sugerido no estudo realizado anteriormente⁽²⁾ e na Classificação Internacional da Funcionalidade (CIF)⁽¹⁶⁾. Na CIF, recomenda-se que a mensuração de um problema (uma deficiência, limitação, restrição ou barreira) em uma escala genérica seja realizada com cinco pontos, sendo zero=neutro, um=problema leve, dois=problema moderado, três=problema intenso e quatro=problema extremo.

Os objetivos do presente estudo são analisar a concordância intra e interavaliadores para a EVA e a EN da tarefa de vogal sustentada e determinar os pontos de corte numéricos da EVA correspondentes aos graus de desvio vocal neutro (grau zero), leve (grau um), moderado (grau dois), intenso (grau três) e extremo (grau quatro).

MÉTODOS

A presente pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da instituição, sob o parecer de número ETIC

0276.0.203.000-10. Trata-se de um estudo observacional analítico transversal, de natureza quantitativa.

Foram selecionadas 205 amostras de 381 vozes arquivadas em banco de dados do consultório de um dos autores e do Ambulatório de Fonoaudiologia do Hospital das Clínicas da Universidade Federal de Minas Gerais. O banco era composto por amostras vocais de homens e mulheres com idades acima de 18 anos, a partir da tarefa da vogal /a/ sustentada de forma habitual por aproximadamente cinco segundos, vozes neutras e alteradas de graus variados, com diagnóstico de disfonia. Todos os participantes envolvidos concordaram em participar da pesquisa e assinaram o Termo de Consentimento Livre e Esclarecido. Cada autor escutou 381 vozes separadamente, usando o fone de ouvido modelo *Multilaser Vibe Headphone* estéreo, e as classificou de acordo com o grau geral da disfonia em: sem desvio ou com graus de desvio leve, moderado, intenso e extremo. Foram selecionadas as vozes que apresentaram concordância de pelo menos dois avaliadores, totalizando 205 amostras compostas por vozes neutras e alteradas de graus variados.

O tamanho da amostra foi determinado a partir do índice Kappa proposto por *Fleiss*, com poder estatístico de 80% e nível de significância de 5%, definindo 205 vozes para seis avaliadores. Para analisar a concordância intra-avaliadores, repetiu-se 20% das vozes aleatoriamente, totalizando 226 vozes.

Foram selecionados seis avaliadores, fonoaudiólogos especialistas em voz, com idades entre 24 e 36 anos, com mais de um ano de experiência na área.

Para avaliação, foram utilizadas uma EVA de 100 mm e uma EN de cinco pontos. O parâmetro avaliado em cada voz foi o grau geral de desvio vocal (G), que corresponde à intensidade da alteração da voz, podendo variar de “sem alteração” ao grau mais intenso de alteração. Cada avaliador recebeu dois envelopes, o primeiro contendo um CD-ROM com as vozes gravadas aleatoriamente e o protocolo de EVA, e o segundo envelope contendo outro CD-ROM com as mesmas vozes gravadas, porém com nomeação e ordem diferentes, bem como o protocolo de EN. Além disso, cada avaliador recebeu um fone de ouvido modelo *Multilaser Vibe Headphone* estéreo.

As tarefas de análise foram realizadas de forma independente e em duas etapas: avaliação das amostras de vogal sustentada por meio da EVA e avaliação dessas mesmas vozes por meio da EN. O intervalo entre as duas etapas compreendeu dois dias, para reduzir a possibilidade de memorização. Na EVA, o avaliador escutou a voz e marcou na reta de 100 mm a medida que considerava de desvio para aquela voz, sendo 0 mm nenhum desvio e 100 mm nível máximo de desvio. Na EN, o avaliador mensurou o G em uma escala de cinco pontos: zero=neutro, um=leve, dois=moderado, três=intenso e quatro=extremo. Os avaliadores podiam escutar as vozes mais de uma vez quando necessário para se certificarem de suas respostas.

Após cada avaliador realizar as tarefas solicitadas, foi realizada a análise estatística por meio dos *softwares* Microsoft® Office Excel 2007 e STATA (Stata Corporation, College Station, Texas) versão 12.0. Em todas as análises, foi considerado nível de significância de 5%. Para analisar a concordância intra e interavaliadores da EVA, foi utilizado o coeficiente de

correlação intraclassa (ICC) e, para a EN, foi utilizado o coeficiente de concordância Kappa. Foi realizada síntese numérica dos dados, sendo que, para a EVA, foram analisadas médias, quartis, mínimo, máximo e desvio padrão, além de gráficos de dispersão e *box plot*; para a EN, foram analisadas frequência e proporções. A correlação entre as escalas foi avaliada por meio do coeficiente de Spearman. Para a comparação das médias, foi utilizada a Análise de Variância (ANOVA). Para estimar os pontos de corte da correspondência entre a EVA e a EN, foram verificadas a sensibilidade e a especificidade das escalas por meio dos valores de verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN), sendo a fórmula utilizada para sensibilidade $VP/(VP+FN)$ e, para especificidade, $VN/(VN+FP)$. Os valores de VP, VN, FP e FN foram encontrados no cruzamento da escala categoria (EN) com os quintis da escala contínua (EVA).

RESULTADOS

Observou-se que a EVA apresentou maiores valores de concordância intra e interavaliador quando comparada com a EN (Tabelas 1 e 2).

Tabela 1. Valores da concordância inter e intra-avaliador na análise pela escala visual analógica

		Escala visual analógica							
		Concordância interavaliador				Concordância intra-avaliador			
E	ICC	ICC 95%	IC	A	E	ICC	ICC 95%	IC	
I	0,496	0,307	0,645		1	I	0,788	0,547	0,909
						M	0,882	0,707	0,952
					2	I	0,835	0,637	0,929
						M	0,910	0,778	0,963
					3	I	0,873	0,714	0,947
						M	0,932	0,833	0,973
M	0,855	0,726	0,916		4	I	0,873	0,681	0,949
						M	0,932	0,810	0,974
					5	I	0,926	0,826	0,969
						M	0,961	0,905	0,984
					6	I	0,738	0,418	0,889
						M	0,849	0,589	0,941

Na análise estatística, foi utilizado o coeficiente de correlação intraclassa (ICC). **Legenda:** E = escala; IC = intervalo de confiança; A = avaliador; I = individual; M = média

Tabela 2. Valores da concordância inter e intra-avaliador na análise pela Escala Numérica

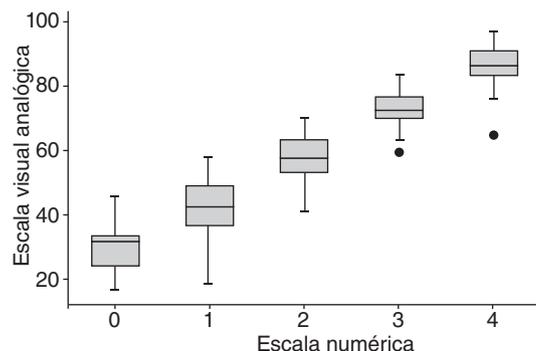
		Escala Numérica					
		Concordância interavaliador			Concordância intra-avaliador		
R	Kappa	Valor de p	A	C (%)	Kappa	Valor de p	
0	0,075	<0,001	1	66,67	0,521	<0,001	
1	0,144	<0,001	2	42,86	0,270	0,007	
2	0,008	0,075	3	80,95	0,700	<0,001	
3	0,051	0,059	4	76,19	0,629	<0,001	
4	0,272	<0,001	5	57,14	0,382	0,001	
Geral	0,114	<0,001	6	66,67	0,538	<0,001	

Na análise estatística, foi utilizado o coeficiente de concordância Kappa. **Legenda:** R = respostas; A = avaliador; C = concordância

Houve alta correlação entre a EVA e a EN, verificada por meio do coeficiente de correlação de Spearman (0,95). Observou-se também que a distribuição entre as duas escalas foi, de forma geral, simétrica, apresentando-se mais assimétrica nos graus zero, três e quatro da EN. Houve menor dispersão no grau três da EN e presença de *outliers* nos graus três e quatro, indicando que a análise de um dos avaliadores para aquele grau apresentou valores discrepantes em relação à dos outros avaliadores (Figura 1).

Ao realizar análise da EVA pelos graus da EN por meio do teste ANOVA, observou-se diferença entre as médias ($p < 0,001$). Esse resultado mostra que o valor médio encontrado na EVA para cada grau se diferencia de um grau para outro na EN.

Os pontos de corte para a EVA equivalentes aos graus da EN definidos por meio da síntese numérica e avaliação da sensibilidade e especificidade dos pontos encontrados estão especificados na Tabela 3 e na Figura 2.



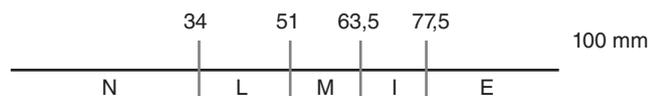
Foi realizada síntese numérica dos dados da correlação entre as escalas visual analógica e numérica, assim como o gráfico box plot para representar essa correlação.

Figura 1. Correlação entre a escala visual analógica e a escala numérica na tarefa de vogal sustentada

Tabela 3. Pontos de corte para a escala visual analógica equivalentes aos graus da escala numérica na tarefa de vogal sustentada

Grau	Valores de corte (mm)	Sensibilidade	Especificidade	Eficiência
0	34	0,76	0,93	0,89
1	51	0,62	0,90	0,83
2	63,5	0,66	0,93	0,89
3	77,5	0,70	0,94	0,89
4	Acima de 77,5	0,90	0,94	0,93

Foi realizada uma síntese numérica dos dados para encontrar os valores de corte e a avaliação da sensibilidade e especificidade dos pontos encontrados.



Legenda: N = neutro; L = leve; M = moderado; I = intenso; E = extremo.

Figura 2. Régua de graduação na escala visual analógica, com base nos respectivos valores de corte, de acordo com a análise perceptivo-auditiva.

DISCUSSÃO

O termo voz normal é controverso e a própria natureza da manifestação vocal humana oferece limitações para estabelecer um critério que seja amplamente aceitável e permita atingir consenso^(2,15). Essa dificuldade é abordada pela literatura, que ressalta a importância de uma normatização, apesar de todas as ressalvas que possam ser feitas sobre o que é voz normal e suas alterações⁽¹⁾.

Estudos anteriores sobre avaliações perceptivo-auditivas da voz têm mostrado que a confiabilidade dessa análise pode ser aumentada com eliminação de fatores que influenciam a variabilidade do avaliador, usando, por exemplo, tarefas correspondentes, âncoras sintéticas externas, estímulos de vogal sustentada e classificações unidimensionais, além dos protocolos validados e utilizados em larga escala⁽¹⁷⁾. No presente estudo, os principais fatores de interferência na análise perceptivo-auditiva foram devidamente controlados. Como avaliadores, foram selecionados fonoaudiólogos especialistas em voz que tinham acima de um ano de experiência em avaliação e tratamento dos distúrbios vocais. Optou-se por realizar a análise do grau geral do desvio (G), por ser considerado, segundo a literatura, um parâmetro robusto e confiável⁽³⁾. Em pesquisa anterior⁽¹⁸⁾, que tinha como objetivo caracterizar a confiabilidade intra e interavaliadores na avaliação de todos os parâmetros da GRBAS utilizando a EN, o G, juntamente com os parâmetros rugosidade e sopro, foi o que apresentou maior concordância interavaliador. Foi utilizada para avaliação na presente pesquisa uma EN de cinco pontos, conforme proposto pela CIF⁽¹⁶⁾ para quantificar um defeito, uma incapacidade, problemas ou barreiras. Estudo anterior⁽²⁾ sobre a correlação entre a EVA e a EN na avaliação perceptivo-auditiva do G também sugeriu tal proposta a fim de possibilitar uma melhor definição da variabilidade normal das vozes e uma separação dos desvios leves, assim como uma categorização mais precisa dos desvios intensos^(2,15). Na presente pesquisa, optou-se pela tarefa de vogal /a/ sustentada produzida de forma habitual, uma vez que, segundo a literatura, os resultados encontrados na avaliação perceptivo-auditiva podem se diferenciar segundo o estímulo apresentado, influenciando também a concordância interavaliador^(11,13,19). A vogal sustentada é mais fácil de se obter e padronizar, uma vez que não é afetada pela articulação dos sons da fala e é caracterizada por uma configuração relativamente estática da laringe e dos músculos supralaríngeos com o mínimo de variação durante sua produção⁽²⁰⁾. Essa estabilidade e consistência na produção da vogal sustentada colabora para uma alta concordância interavaliadores^(18,20). Uma pesquisa⁽²¹⁾ correlacionou a tarefa de fala à confiabilidade encontrada na avaliação perceptivo-auditiva da voz, e a tarefa que apresentou maior concordância interavaliadores foi a emissão da vogal /a/ sustentada de forma habitual, para o parâmetro G.

Observou-se no presente estudo uma concordância interavaliadores maior na avaliação pela EVA (0,855) do que na EN (0,114) (Tabelas 1 e 2), resultados concordantes com um estudo que correlacionou as duas escalas na avaliação perceptivo-auditiva do grau geral de desvio vocal^(2,15), observando valores de concordância interavaliadores discretamente maiores para a

EVA (0,849) do que para a EN (0,821). A elevada concordância interavaliadores permite estabelecer o limite entre a variabilidade normal da qualidade vocal e o grau geral do desvio vocal. Outros estudos que analisaram também o parâmetro G, comparando a EVA com a EN, apontaram que a EVA apresenta maior confiabilidade interavaliadores e é mais sensível a pequenas diferenças de marcação que a EN⁽⁴⁾. Uma pesquisa⁽²²⁾ que comparou a avaliação do G e dos parâmetros rugosidade, sopro, astenia e tensão de 14 vozes disfônicas na tarefa de vogal sustentada, realizada por 29 avaliadores, usando a EN da escala GRBAS e uma EVA, verificou que houve maior concordância interavaliadores na avaliação pela EN. O resultado encontrado se difere do presente estudo, o que pode estar relacionado aos parâmetros avaliados, uma vez que se avaliou apenas o grau geral de desvio vocal. Pode também estar relacionado ao tamanho da amostra assim como ao número de avaliadores.

A concordância intra-avaliador no presente estudo também foi maior na avaliação pela EVA (0,849 a 0,961) do que na EN (0,270 a 0,700) (Tabelas 1 e 2), o que mostra que os ouvintes foram mais consistentes em suas avaliações por meio da EVA e corrobora resultados de estudos anteriores^(2,15).

Quanto à correlação entre as duas escalas, na presente pesquisa, foi observada uma alta correlação entre a EVA e a EN, verificada por meio do coeficiente de correlação de Spearman (0,95) para a tarefa de vogal sustentada. Esse achado corrobora outras pesquisas^(2,4,14), apesar de terem sido realizadas apenas com a tarefa de fala encadeada. Todos os achados comprovam que existe uma correlação entre as escalas, permitindo comparações entre seus resultados.

Em um estudo desenvolvido na Finlândia, que tinha como objetivo definir a diferenciação entre variações normais da qualidade vocal e alterações vocais por análise perceptivo-auditiva, dois fonoaudiólogos avaliaram o parâmetro G de 226 amostras por meio da escala numérica GRBAS e por meio de uma EVA⁽¹⁴⁾. Após realizarem uma comparação entre a avaliação por meio das duas escalas, os autores concluíram que, em uma EVA de 100 mm, 34 mm seria o ponto de corte para variação normal da qualidade da voz, sendo que vozes marcadas acima desse ponto representariam uma voz alterada⁽¹⁴⁾. Esse estudo foi reproduzido no Brasil, avaliando o mesmo parâmetro, porém as avaliações foram realizadas por uma única avaliadora⁽¹⁵⁾. O resultado encontrado foi muito próximo ao encontrado no estudo finlandês⁽¹⁴⁾, sendo o valor de corte obtido de 34,5 mm, demonstrando que essa forma de análise parece ser bastante robusta e não sofre influências de aspectos culturais da qualidade vocal⁽¹⁵⁾. Uma segunda reprodução do estudo foi realizada no Brasil com quatro avaliadores e 211 amostras vocais, com o objetivo de definir valores de corte dos diferentes graus de desvio vocal na EVA, de acordo com os valores de uma EN de quatro pontos⁽²⁾. O valor de corte encontrado para variação normal da qualidade da voz foi de 35,5 mm; para grau de desvio vocal de leve a moderado foi de 35,6 a 50,5 mm; para o grau moderado a intenso, foi de 50,6 a 90,5 mm; e para o grau intenso, foi acima de 90,5 mm. Esse foi o único estudo realizado anteriormente que comparou uma EVA com uma EN de quatro pontos na avaliação perceptivo-auditiva do grau geral de desvio vocal e determinou valores de corte para a EVA correspondentes

aos graus da EN. Todos esses estudos foram realizados com amostras vocais da tarefa de fala encadeada^(2,14,23).

Na presente pesquisa, foram encontrados os seguintes valores de corte numéricos: neutro (grau zero) — 0 a 34 mm, leve (grau um) — 34,1 a 51 mm, moderado (grau dois) — 51,1 a 63,5 mm, intenso (grau três) — 63,6 a 77,5 mm, extremo (grau quatro) — acima de 77,5 mm (Tabela 3 e Figura 1). Os resultados encontrados nesta pesquisa mostram semelhanças com os achados da literatura^(2,14,23) e que o tipo de tarefa de fala parece não influenciar a correlação entre as escalas EVA e EN para os graus neutro e leve de desvio. Os pontos próximos a 34 mm e a 50 mm são apresentados como valores de corte de variação da qualidade vocal neutra e desvio leve respectivamente para a fala encadeada de acordo com a literatura^(2,14,23) e para a vogal sustentada conforme o achado desta pesquisa. Os valores correspondentes aos graus de desvio moderado e intenso entre as escalas EVA e EN referidos pela literatura⁽²⁾ são diferentes dos resultados desta pesquisa (Tabela 3). Tal discrepância pode ser justificada pela tarefa de fala, já que a tarefa de vogal sustentada tende a ser avaliada com maior desvio quando comparada à fala encadeada⁽²⁴⁾. Esta pesquisa analisou também o grau de desvio extremo seguindo a orientação da CIF⁽¹⁶⁾; tal correspondência não foi avaliada pela literatura.

Os resultados encontrados permitem uma comparação entre a avaliação perceptivo-auditiva do grau geral de desvio vocal pela EVA e a avaliação pela EN. Sugere-se que novos estudos sejam realizados comparando a EVA com a EN na avaliação perceptivo-auditiva de outros parâmetros vocais utilizando a tarefa de vogal sustentada e de fala encadeada.

CONCLUSÃO

Verificou-se que a EVA e a EN apresentaram uma alta correlação, sendo observada uma maior concordância intra e interavaliador na EVA. Foram encontrados valores de corte numéricos de 34 mm para grau zero (neutro), 51 mm para grau um (leve), 63,5 mm para grau dois (moderado), 77,5 mm para grau três (intenso) e acima de 77,5 mm para grau quatro (extremo) da EVA na tarefa de vogal sustentada. Essa correlação possibilita a comparação entre os resultados encontrados na avaliação do grau geral de desvio vocal pela EVA e pela EN, escalas amplamente utilizadas em pesquisas e na rotina clínica fonoaudiológica.

**PCM, TEC e ACCG conceberam e planejaram o projeto, assim como analisaram e interpretaram os dados.*

REFERÊNCIAS

1. Behlau M, organizadora. Voz: o livro do especialista. vol. 1. Rio de Janeiro: Revinter; 2001.
2. Yamasaki R, Leão SHS, Madazio G, Padovani M, Azevedo R. Correspondência entre Escala Analógico-Visual e a Escala Numérica na Avaliação Perceptivo-Auditiva de Vozes. In: XVI Congresso Brasileiro de Fonoaudiologia e VIII Congresso Internacional de Fonoaudiologia; 2008 set 24-27; Campos do Jordão.

3. Oates J. Auditory-perceptual evaluation of disordered voice quality: pros, cons and future directions. *Folia Phoniatr Logop.* 2009;61(1):49-56.
4. Nemr K, Simões-Zenari M, Cordeiro GF, Tsuji D, Ogawa AI, Ubrig MT, et al. GRBAS and Cape-V Scales: high reliability and consensus when applied at different times. *J Voice.* 2012;26(6):812e17-22.
5. American Speech-Language-Hearing Association. Consensus auditory-perceptual evaluation of voice (CAPE-V). Rockville: ASHA Special Interest Division 3, Voice and Voice Disorders; 2002.
6. Zraick RI, Kempster GB, Connor NP, Thipeault S, Klaben B, Bursac Z, et al. Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *Am J Speech Lang Pathol.* 2011;20:14-22.
7. Hirano M. Clinical examination of voice. New York: Springer-Verlag; 1981.
8. Silva RSA, Simões-Zenari M, Nemr NK. Impacto de treinamento auditivo na avaliação perceptivo-auditiva da voz realizada por estudantes de Fonoaudiologia. *J Soc Bras Fonoaudiol.* 2012;24(1):19-25.
9. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res.* 2011;63(S11):S240-52.
10. Rouve S, Didier A, Demoly P, Jankowsky R, Klossek JM, Annesi-Maesano I. Numeric score and visual analog scale in assessing seasonal allergic rhinitis severity. *Rhinology.* 2010;48(3):285-91.
11. Pires MDE, Oliveira G, Behlau M. Aplicação do protocolo de participação em atividades vocais - PPAV em duas diferentes escalas de resposta. *J Soc Bras Fonoaudiol.* 2011;23(3):297-300.
12. Chan K, Li M, Law T, Yiu E. Effects of immediate feedback on learning auditory perceptual voice quality evaluation. *Int J Speech Lang Pathol.* 2012;14(4):363-9.
13. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *J Commun Disord.* 2010;43(3):161-74.
14. Simberg S, Laine A, Sala E, Rönnemaa AM. Prevalence of voice disorders among future teachers. *J Voice.* 2000;14(2):231-5.
15. Yamasaki R, Leão SHS, Madazio G, Padovani M, Azevedo R. Análise perceptivo-auditiva de vozes normais e alteradas: escala analógica visual. In: XV Congresso Brasileiro de Fonoaudiologia e VII Congresso Internacional de Fonoaudiologia; 2007 out 16-20; Gramado.
16. Organização Mundial da Saúde. CIF: Classificação Internacional de Funcionalidade, Incapacidade e Saúde. São Paulo: EDUSP; 2003.
17. Iwarsson J, Petersen NR. Effects of consensus training on the reliability of auditory perceptual ratings of voice quality. *J Voice.* 2011;25(3):304-12.
18. Freitas SV, Pestana PM, Almeida V, Ferreira A. Audio-perceptual evaluation of Portuguese voice disorders - an inter and intra-judge reliability study. *J Voice.* 2014;28(2):210-5.
19. Costa FP, Yamasaki R, Behlau M. Influência da escuta contextualizada na percepção da intensidade do desvio vocal. *Audiol Commun Res.* 2014;19(1):69-74.
20. Law T, Kim JH, Lee KY, Tang EC, Lam JH, van Hasselt AC, et al. Comparison of rater's reliability on perceptual evaluation of different types of voice sample. *J Voice.* 2012;26(5):666.e13-21.
21. Lu FL, Matteson S. Speech tasks and inter-rater reliability in perceptual voice evaluation. *J Voice.* 2014;28(6):725-32.
22. Wuyts FL, De Bodt MS, Van de Heyning PH. 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.
23. Behlau M, Oliveira G, Madazio G, Yamasaki R. Speech Language Pathology and the voice specialist in Brazil: an overview. In: Yui E., editor. *International perspectives on voice disorders.* Bristol: Multilingual Matters; 2013. p. 30-41.
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.