

# Análise espectrográfica da voz: efeito do treinamento visual na confiabilidade da avaliação

## *Spectrographic analysis of the voice: effect of visual training on the reliability of evaluation*

Amanda Freitas Valentim<sup>1</sup>, Marcela Guimarães Côrtes<sup>2</sup>, Ana Cristina Côrtes Gama<sup>3</sup>

### RESUMO

**Objetivo:** Verificar se o treinamento visual melhora a concordância inter e intra-avaliadores da interpretação da análise acústica espectrográfica. **Métodos:** Trata-se de um estudo experimental longitudinal que utilizou espectrogramas coletados em um banco de dados de consultório particular, referentes a vozes de homens e mulheres analisadas pelo programa Gram 5.0 em escala de 60 dB, em *display scroll*, paleta BW, com análise de frequências em escala linear, banda estreita, FFT 1024 e resolução 5.4 Hz. Vinte sujeitos, estudantes de Fonoaudiologia, avaliaram 25 espectrogramas, sendo cinco destes repetidos para avaliar a concordância intra-sujeitos, antes e após um treinamento de análise de espectrogramas. Os parâmetros analisados foram: forma do traçado espectrográfico, grau de escurecimento dos harmônicos, estabilidade do traçado espectrográfico, presença de ruído, presença de sub-harmônicos e definição dos harmônicos. O treinamento teve a duração total de quatro horas. Utilizou-se a estatística Kappa para avaliar a concordância intrassujeitos e o teste Qui-quadrado de ajuste para a concordância intersujeitos. **Resultados:** Em relação à concordância intra-sujeitos, os aspectos escurecimento do traçado, estabilidade do traçado, presença de sub-harmônicos, e definição de harmônicos obtiveram melhora desta e os demais aspectos mantiveram concordância quase perfeita depois do treinamento. Quanto à concordância intersujeitos, os aspectos de estabilidade do traçado e presença de sub-harmônicos obtiveram melhora, e o aspecto presença de ruído obteve piora após o treinamento. **Conclusão:** O treinamento visual melhorou a confiabilidade inter e intra-avaliadores na interpretação de espectrogramas.

**Descritores:** Voz; Acústica da fala; Espectrografia do som; Variações dependentes do observador; Capacitação

### INTRODUÇÃO

Atualmente, a avaliação fonoaudiológica da voz pode ser feita por meio da avaliação perceptivo-auditiva, ou da análise acústica, que são medidas ou gráficos gerados por programas de computador.

Pode-se dizer que a avaliação acústica é uma análise objetiva, pois quantifica o sinal sonoro. Porém, o ser humano participa em maior ou menor grau do processo de avaliação

e da compreensão do resultado observado, o que indica que não existe uma avaliação puramente objetiva<sup>(1)</sup>.

Medidas acústicas da voz têm recebido atenção considerável por serem de natureza não invasiva e relativamente aplicáveis na clínica e em pesquisas. Porém, essas medidas devem ser utilizadas juntamente com a avaliação perceptivo-auditiva para serem consideradas clinicamente significativas<sup>(2)</sup>.

A análise acústica tem a espectrografia como uma de suas principais ferramentas. O espectrograma pode ser definido como um gráfico que mostra a intensidade por meio do escurecimento ou coloração do traçado, as faixas de frequência no eixo vertical e o tempo no eixo horizontal. Sua representação mostra estrias horizontais, denominadas harmônicos. O espectrograma demonstra visualmente as características acústicas da emissão, porém essas informações exigem interpretação por parte do avaliador<sup>(3)</sup>.

Não existe correlação direta entre dados auditivos e espectrográficos, entretanto, um estudo encontrou que os parâmetros perceptivo-auditivos de grau geral da disфонia, rouquidão/asperidade, sopro e instabilidade se relacionaram com traçado irregular dos harmônicos nos espectrogramas em mais de 66% dos casos avaliados; *loudness* fraca se relacionou com grau

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

(1) Pós-graduanda (Mestrado) do Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

(2) Pós-graduanda (Mestrado) do Programa de Pós-Graduação em Ciências da Saúde - Saúde da Criança e do Adolescente da Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

(3) Doutora, Professora Adjunta do Departamento de Fonoaudiologia da Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

**Endereço para correspondência:** Amanda Freitas Valentim. R. David Campista, 207/501, Floresta, Belo Horizonte (MG), Brasil, CEP: 30150-090. E-mail: amandafvalentim@gmail.com

**Recebido em:** 1/6/2009; **Aceito em:** 24/9/2009

de escurecimento dos harmônicos fraco em 87,5% dos casos; quebras de sonoridade na voz, com falhas na continuidade do traçado em 62,5%; disфония e instabilidade vocal, com ruído entre os harmônicos em 97,4%; disфония, rouquidão/aspereza com diminuição da concentração de energia nas altas frequências em 48,7%; disфония, rouquidão/aspereza e sub-harmônicos em 79,5%. Isso mostra que esses aspectos são importantes parâmetros a serem analisados em um espectrograma<sup>(4)</sup>.

Outra pesquisa, realizada com vozes disártricas, encontrou traçado espectrográfico instável em todos os casos e, na maioria deles, ausência de harmônicos superiores e presença de ruído entre harmônicos. Esses achados são compatíveis com as vozes rouco-soprosas, com alteração de ressonância e incoordenação da musculatura laríngea encontradas nesses sujeitos. Apesar disso, não houve equivalência estatisticamente significativa entre os parâmetros auditivos e acústicos, reforçando a ideia de que estas avaliações devem ser complementares<sup>(5)</sup>.

A avaliação espectrográfica da voz apresenta subjetividade da análise, portanto, tal julgamento perceptivo-visual pode sofrer influências da experiência do avaliador.

Assim como a avaliação espectrográfica da voz, a videolaringoscopia, exame que possibilita visualizar as estruturas laríngeas e o padrão vibratório das pregas vocais, necessita de uma análise perceptivo-visual, sofrendo, portanto, a influência de fatores relacionados à instrumentação, treino e habilidade dos avaliadores, que farão a interpretação do exame<sup>(6)</sup>.

Na literatura, estudos que analisaram o efeito do treinamento visual na concordância intra e interavaliadores na interpretação de avaliações por imagem, como espectrogramas de fala e exames radiológicos, demonstraram diferentes resultados nos momentos pré e pós treinamento, fato que pode ser justificado principalmente pelo grau de complexidade da imagem e, conseqüentemente, da análise perceptivo-visual<sup>(7-11)</sup>.

Não se encontrou na literatura nenhum trabalho que estudasse o efeito do treinamento visual na avaliação espectrográfica vocal, portanto, essa pesquisa pretende analisar o quanto o treinamento visual interfere na concordância intra e inter examinadores na avaliação da análise espectrográfica da voz.

## MÉTODOS

Trata-se de um estudo experimental longitudinal, que utilizou espectrogramas coletados em um banco de dados de consultório particular, com a autorização do profissional responsável. Os registros pertencem a homens e mulheres disfônicos ou sem alteração de voz, com mais de 18 anos de idade. O material de voz colhido foi a emissão sustentada da vogal /a/.

A gravação das vozes arquivadas foi realizada diretamente em computador, equipado com microfone profissional, do tipo condensado, estéreo, omnidirecional, sensibilidade de -20 dB, da marca EQUITEK E-100, ligado a uma fonte de eletricidade (*Phanton Power*) de uma mesa de som marca MACKIE 1202 VLZ – 12 canais. Os indivíduos estavam em pé, com o microfone situado a 10 cm da boca e com ângulo de captação direcional de 90°. O microfone estava deslocado do corpo da unidade de gravação para evitar captação de ruído

do maquinário. Foi utilizada fita digital marca Panasonic® e as gravações foram realizadas em ambiente silente.

A gravação da vogal sustentada foi transferida para o PC IBM Aptiva E30P, processador AMD – K6 – 2 / 500 MHz, memória de 128 mega bytes RAM, espaço de disco 8,4 Giga bytes, placa de som *Crystal SoundFusion*.

As emissões foram analisadas acusticamente pelo programa Gram 5.0, o qual gera a imagem espectrográfica do som. As vozes foram gravadas em arquivo mono, *sample rate* de 11k e resolução de 16 bit e as espectrografias, geradas em escala de 60 dB em *display scroll* e paleta BW. A análise das frequências foi realizada em escala linear, banda estreita, FFT 1024 e resolução 5.4 Hz. Os espectrogramas foram colocados em slides do *software* PowerPoint para serem apresentados aos participantes.

Foram selecionados para o estudo, 20 estudantes que estavam cursando 3º a 5º períodos de Fonoaudiologia de uma universidade. Estes foram escolhidos devido ao fato de não terem participado de nenhum tipo de treinamento em análise espectrográfica. A amostra foi composta por uma pessoa do sexo masculino e 19 do sexo feminino, com idades variando de 19 a 35 anos. Todos os participantes assinaram o termo de consentimento livre e esclarecido.

Para análise da concordância intra e inter avaliadores, utilizou-se de 25 espectrogramas, dos quais cinco foram repetidos. Foi realizada uma breve explicação sobre os parâmetros espectrográficos a serem analisados, a partir de exemplos de cada item e então os participantes realizaram uma primeira avaliação. Depois do treinamento visual, os sujeitos refizeram a análise dos mesmos 25 espectrogramas. Portanto, os resultados são de avaliação anterior e posterior ao treinamento, com um intervalo de uma semana entre as duas avaliações. As vozes não foram apresentadas em momento algum, sendo a análise puramente visual.

Para o treinamento visual, utilizou-se de aproximadamente 100 espectrogramas que foram exibidos e classificados. O treinamento foi realizado em dois encontros, com duração total de quatro horas.

Foi utilizado o Protocolo de Análise Visual/Espectrográfica, elaborado pelas pesquisadoras, em que o avaliador deveria classificar os seguintes parâmetros (Figuras 1 a 3):

- Forma do traçado espectrográfico (regular, irregular ou ausente): o traçado foi considerado regular quando as linhas dos harmônicos eram bem definidas, nítidas; irregular quando não era possível definir os harmônicos em alguma parte do traçado; ausente quando não havia traçado.
- Grau de escurecimento dos harmônicos (forte, normal, fraco ou não é possível avaliar): o grau de escurecimento foi considerado forte quando os harmônicos apresentavam-se em tons de preto e cinza escuro e podiam ser vistos ao longo de todo o espectro; fraco quando os harmônicos apresentavam-se em tons de cinza claro e se apagavam em algumas regiões do espectro; normal quando apresentava coloração intermediária. Nos casos de traçado ausente, não é possível avaliar o grau de escurecimento.
- Estabilidade do traçado espectrográfico (estabilidade, instabilidade ou não é possível avaliar): considerou-se que o traçado tinha instabilidade quando apresentava oscila-

ções evidentes ou quebras; estabilidade quando o traçado mantinha-se contínuo e constante, com ondulações mínimas. Nos casos de traçado ausente, não é possível avaliar a estabilidade do traçado.

- Presença de ruído (presente ou ausente): quando o traçado apresentava, entre os harmônicos, o fundo sombreado ou pontilhado, deixando-o mais escuro, considerou-se que o ruído estava presente. Quando não havia esse sombreado e podia-se ver o fundo branco, considerou-se ruído ausente.
- Presença de sub-harmônicos (presente ou ausente): considerou-se que o traçado apresentava sub-harmônicos quando, em alguma parte do espectrograma, verificava-se a presença de traçado entre dois harmônicos consecutivos. Os sub-harmônicos foram considerados ausentes quando isto não acontecia.
- Definição dos harmônicos (o avaliador deve escrever até qual frequência os harmônicos encontram-se definidos): foi considerado que os harmônicos estavam definidos até a frequência mais alta onde estes ainda eram nítidos.

Essa pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais (UFMG) sob o parecer ETIC 675/07.

Para análise dos dados, foi utilizada a estatística Kappa para avaliar a concordância intraexaminadores, que foi classificada de acordo com o seguinte critério<sup>(12)</sup>: quase perfeita - Kappa entre 0,80 e 1,00; boa - Kappa entre 0,60 e 0,80; moderada - Kappa entre 0,40 e 0,60; regular - Kappa entre 0,20 e 0,40; discreta - Kappa entre zero e 0,20; pobre - Kappa entre -1 e zero.

Utilizou-se, também, o teste Qui-quadrado de ajuste para avaliar a concordância interexaminadores, e o índice de significância adotado foi de  $p < 0,05$ .

## RESULTADOS

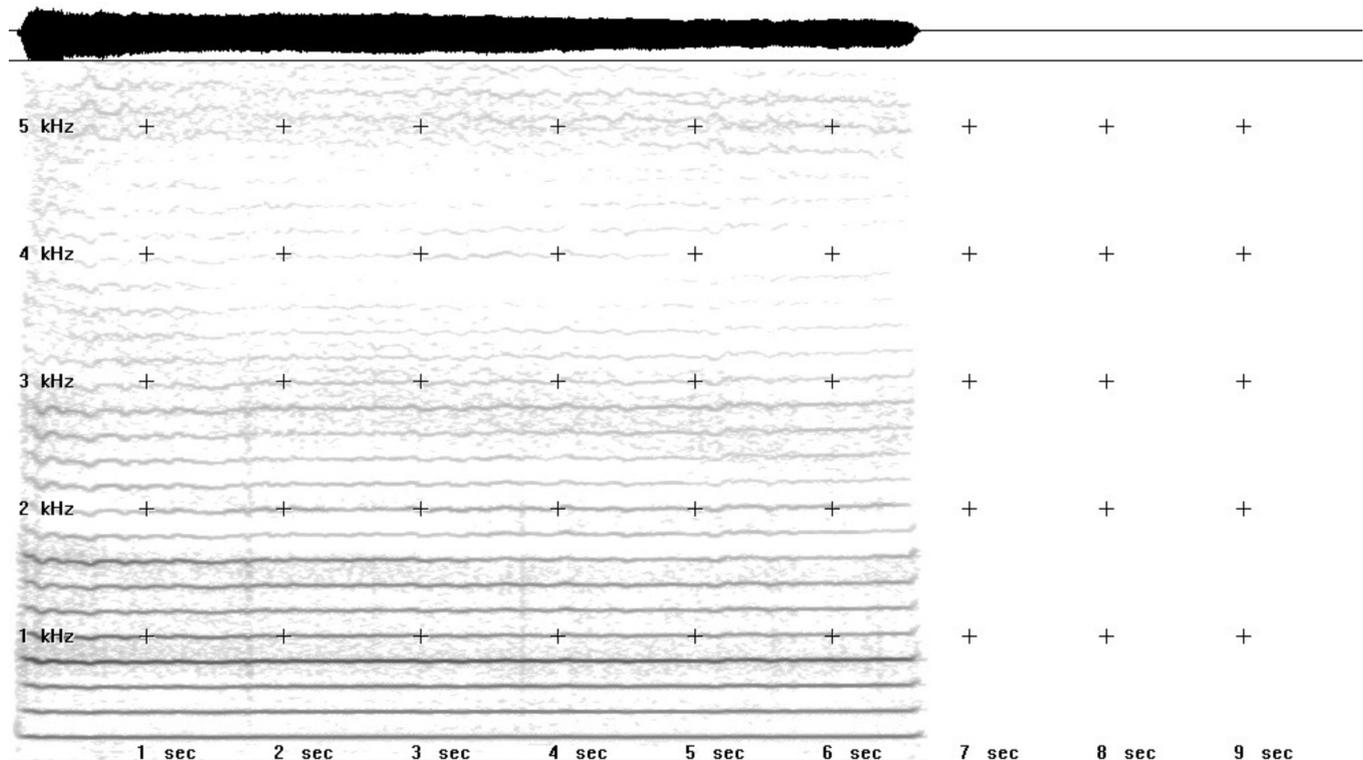
A Tabela 1 mostra os valores de concordância intra-avaliadores individuais e gerais, por parâmetro espectrográfico, antes e depois do treinamento. Apenas dois aspectos apresentavam concordância quase perfeita e um, concordância boa antes do treinamento. Todos os parâmetros apresentaram melhora ou mantiveram o mesmo nível de concordância depois do treinamento, sendo que cinco dos seis aspectos apresentaram concordância quase perfeita ou boa.

A Tabela 2 apresenta os valores de concordância interavaliadores por parâmetro espectrográfico, antes e depois do treinamento. Antes do treinamento, apenas um parâmetro apresentava valor de p estatisticamente significante e dois, menores que 0,10. Todos os parâmetros apresentaram diminuição no valor de p após o treinamento, com exceção de presença de ruído, sendo que dois foram significantes e todos os valores de p foram menores que 0,10.

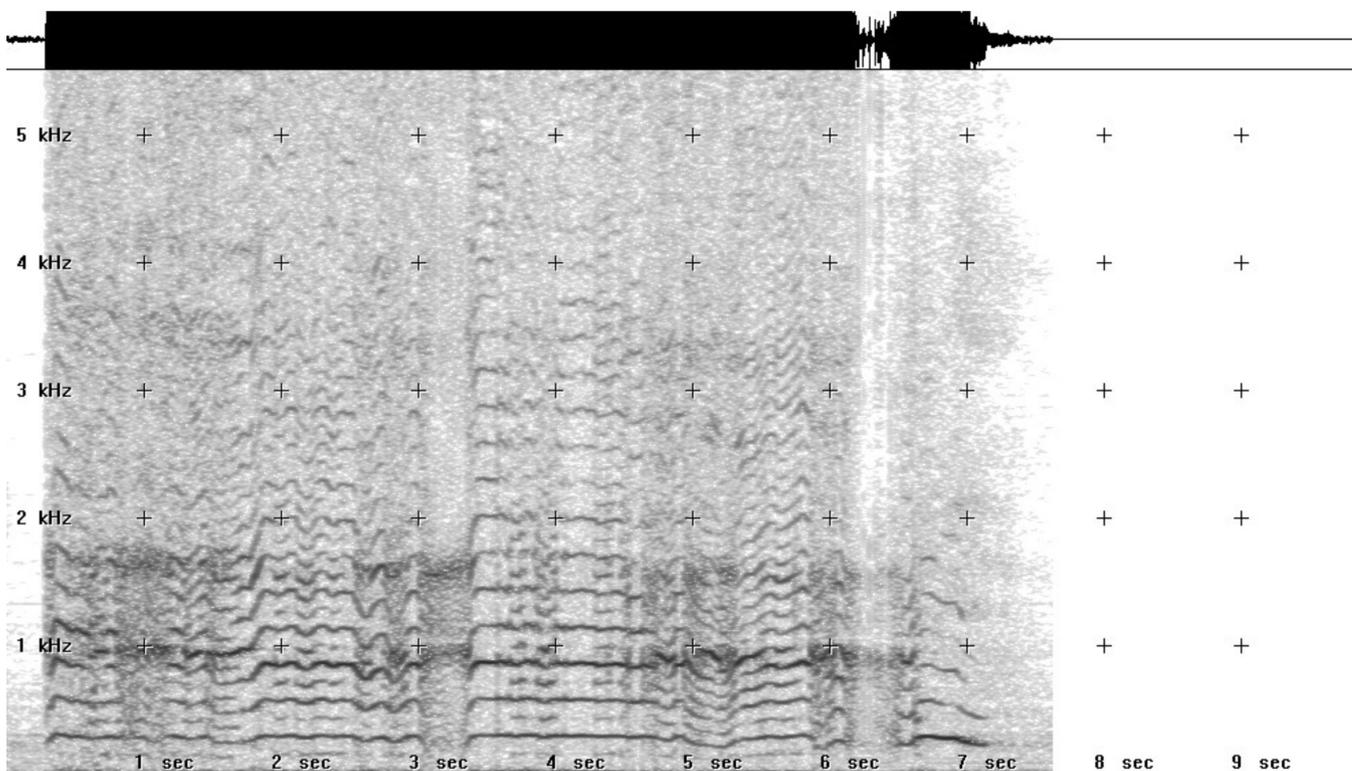
## DISCUSSÃO

Atualmente, o padrão-ouro utilizado para avaliação de voz é a avaliação perceptivo-auditiva, que é subjetiva e depende da experiência do avaliador<sup>(13)</sup>.

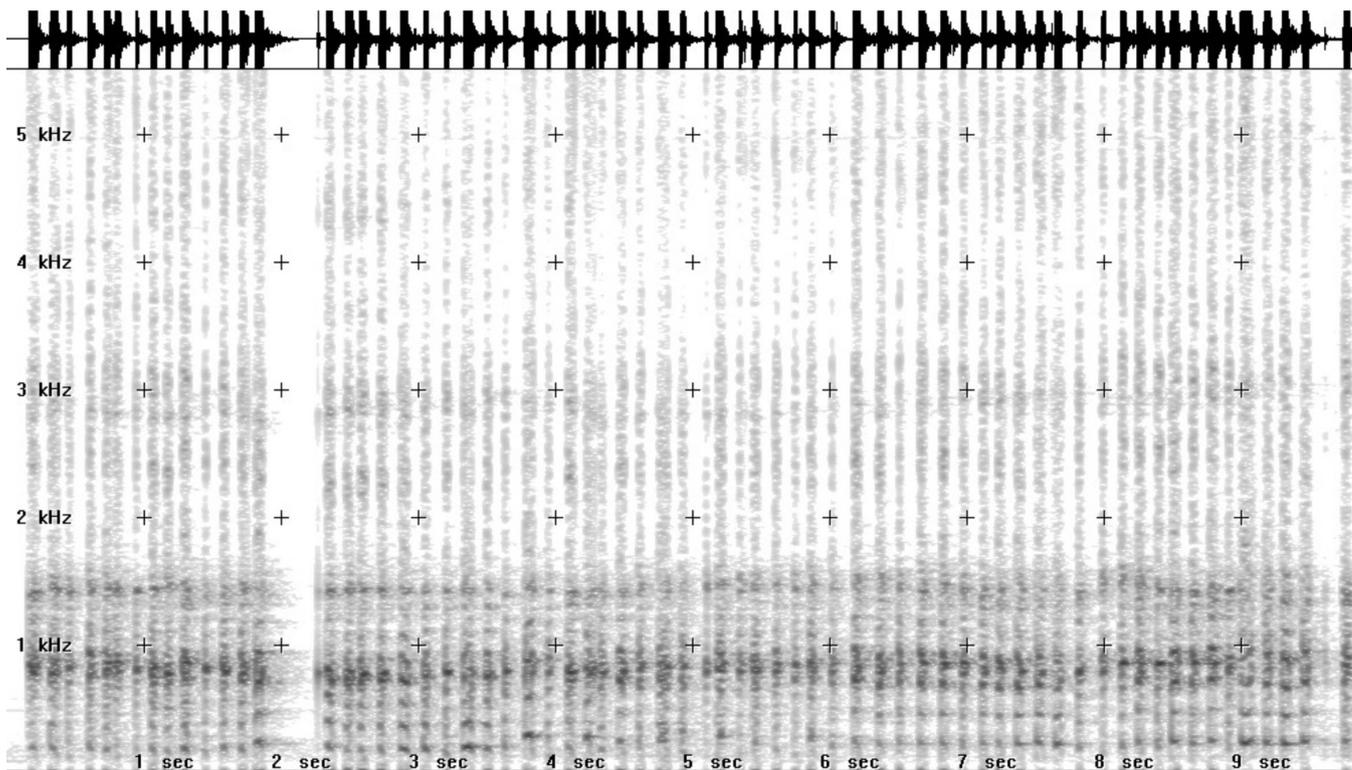
Diante da subjetividade da avaliação perceptivo-auditiva, cada vez mais, têm-se utilizado ferramentas complementares para avaliação da voz, como a análise acústica. Um dos tipos de análise acústica são as medidas a curto prazo, que são valores gerados por programas de computador, sendo dessa forma mais objetivos. Porém, como cada programa utiliza uma maneira diferente para calcular os parâmetros acústicos, os valores



**Figura 1.** Espectrograma com traçado regular, intensidade fraca, estável, com ruído ausente, sub-harmônicos ausentes e harmônicos definidos até 3 kHz



**Figura 2.** Espectrograma com traçado irregular, intensidade normal, instável, com presença de ruído, presença de sub-harmônicos e harmônicos definidos até 2 kHz



**Figura 3.** Espectrograma com traçado ausente, no qual não é possível avaliar estabilidade e grau de intensidade, com presença de ruído, sub-harmônicos ausentes e harmônicos definidos até 0 kHz

de normalidade diferem entre eles. Além disso, esses valores também variam conforme os instrumentos de gravação, ruído ambiental, gênero e idade do falante<sup>(14,15)</sup> o que mostra que a qualidade do equipamento utilizado no registro das vozes, o

tipo de programa e as características anatomofuncionais da laringe podem influenciar nos resultados destas medidas a curto prazo.

A espectrografia vocal é uma outra forma de análise acústi-

**Tabela 1.** Valores individuais e gerais para análise da concordância intra-avaliadores antes e depois do treinamento visual, para cada parâmetro espectrográfico

Sujeitos	Forma		Intensidade		Estabilidade		Ruído		Sub-harmônicos		Definição	
	A	D	A	D	A	D	A	D	A	D	A	D
1	1,00	1,00	0,41	0,62	0,00	1,00	1,00	0,55	0,62	1,00	-0,11	1,00
2	0,55	1,00	1,00	0,55	1,00	1,00	0,55	1,00	-0,25	1,00	0,55	1,00
3	1,00	1,00	0,55	0,62	0,55	1,00	0,55	0,55	1,00	1,00	0,00	0,62
4	1,00	1,00	0,00	-0,11	0,55	0,29	0,62	1,00	0,29	1,00	-0,25	0,00
5	0,62	0,55	0,38	0,55	1,00	0,00	1,00	1,00	1,00	1,00	-0,25	0,29
6	0,33	1,00	0,06	0,17	0,67	1,00	1,00	1,00	0,55	1,00	-0,25	1,00
7	1,00	0,55	0,55	0,00	0,00	0,29	1,00	0,55	0,55	1,00	-0,36	1,00
8	1,00	1,00	0,00	1,00	0,17	1,00	1,00	0,55	1,00	1,00	0,55	-0,15
9	1,00	1,00	1,00	0,06	0,55	1,00	1,00	1,00	1,00	0,62	-0,33	0,00
10	1,00	1,00	0,64	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,44	0,00
11	1,00	1,00	-0,43	0,33	-0,36	1,00	1,00	1,00	1,00	1,00	0,29	1,00
12	1,00	1,00	0,55	0,29	0,55	0,55	1,00	1,00	0,00	0,55	0,33	0,00
13	1,00	0,00	0,44	1,00	1,00	0,55	1,00	1,00	1,00	0,62	0,00	0,17
14	1,00	1,00	0,00	0,67	0,55	1,00	1,00	0,55	0,62	1,00	0,17	0,00
15	0,00	1,00	0,64	1,00	1,00	0,62	1,00	1,00	0,62	0,62	-0,36	1,00
16	1,00	0,55	-0,25	0,55	1,00	0,55	1,00	0,55	1,00	1,00	0,62	0,50
17	0,29	1,00	0,06	0,62	0,17	0,62	1,00	1,00	0,29	1,00	0,55	1,00
18	0,44	1,00	-0,15	1,00	-0,25	0,55	1,00	1,00	1,00	0,62	0,38	1,00
19	0,55	0,00	0,29	1,00	0,29	0,55	1,00	1,00	1,00	1,00	0,55	0,00
20	1,00	1,00	0,64	0,67	1,00	1,00	1,00	1,00	1,00	1,00	0,62	1,00
Geral	0,80	0,85	0,43	0,62	0,59	0,75	0,94	0,87	0,69	0,90	0,35	0,56
Classif.*	a	a	c	b	c	b	a	a	b	a	d	c

Estatística Kappa

**Legenda:** A = antes do treinamento; D = depois do treinamento; \*Classif. = classificação da concordância<sup>(12)</sup>; a = quase perfeita; b = boa; c = moderada; d = regular; e = discreta; f = pobre

**Tabela 2.** Valores da concordância interavaliadores antes e após o treinamento visual, para cada parâmetro espectrográfico

Parâmetro espectrográfico	Concordância antes	Concordância depois
Forma do traçado	0,110	0,088
Intensidade do traçado	0,129	0,096
Estabilidade do traçado	0,060	0,045*
Presença de ruído	0,005*	0,085
Presença de sub-harmônicos	0,127	0,048*
Definição de harmônicos	0,259	0,060

Teste Qui-quadrado de ajuste

\* Valores significativos ( $p < 0,05$ )

ca, sendo representada por um gráfico tridimensional, resultado da relação entre frequência, intensidade e tempo. Apesar da avaliação espectrográfica da voz ser considerada um exame objetivo, existe a subjetividade da análise visual do traçado espectrográfico pelo avaliador.

Esta pesquisa analisou a concordância das respostas inter e intra-avaliadores, antes e depois de um treinamento visual com espectrogramas. Não foram encontrados na literatura, estudos que avaliassem o efeito do treinamento visual na confiabilidade da interpretação de espectrogramas de voz.

Para o presente estudo, realizou-se um cálculo amostral para o teste Kappa, por meio do *software* Minitab for Windows®, considerando-se uma população de 125 sujeitos. Verificou-se que, para se obter um poder de 90% no teste, seria necessária uma amostra de 44 sujeitos. Dessa forma, todos os 125 estudantes de primeiro a quinto períodos do curso de Fonoaudiologia da UFMG foram convidados a fazer parte do estudo, porém apenas 20 se dispuseram a participar. Observou-se que muitos se interessaram, mas apresentaram dificuldade em disponibilizar duas horas durante dois dias de sua vida acadêmica, já que o curso de Fonoaudiologia nesta universidade ocorre em período integral.

Ao contrário do verificado por um estudo<sup>(16)</sup>, no qual 35 profissionais e graduandos de Odontologia avaliaram perda óssea em 25 radiografias periodontais antes e depois de um treinamento, não houve evasão de participantes ao longo do estudo, o que indica que aqueles que se dispuseram a participar e iniciaram o treinamento, consideraram este relevante.

Na avaliação realizada antes do treinamento, no presente estudo, verificou-se concordância intra-avaliadores (Tabela 1) quase perfeita para os parâmetros de forma do traçado espectrográfico e presença de ruído e concordância boa para o parâmetro de presença de sub-harmônicos, o que indica que estes são aspectos visuais fáceis de serem avaliados,

sendo necessária apenas uma breve explicação dos critérios de análise de tais parâmetros. Resultado semelhante foi encontrado em uma pesquisa que verificou confiabilidade ótima intrassujeitos de três avaliadores na identificação de pontos cefalométricos em 12 tomografias computadorizadas *cone beam*<sup>(17)</sup>. O mesmo não ocorre com a avaliação visual da presença de displasia em lâminas patológicas do esôfago de Barrett, já que um estudo com dois sujeitos avaliando 229 lâminas desse tecido, encontrou concordância pobre ou fraca intra-avaliadores<sup>(18)</sup>. Uma pesquisa que analisou a identificação de nódulos pulmonares em 25 tomografias computadorizadas avaliadas por quatro radiologistas torácicos, também não obteve boa concordância, encontrando grande variabilidade intra-avaliador<sup>(19)</sup>. Os diferentes resultados podem ser explicados, provavelmente, pela complexidade da imagem do exame, que pode tornar a avaliação mais ou menos complexa. Um estudo piloto semelhante a esta pesquisa, conduzido pelas autoras, realizou um treinamento visual de quatro horas com três avaliadores com pouca experiência, utilizando espectrogramas de voz. Verificou-se que a concordância intrassujeitos já era boa inicialmente e passou a quase perfeita depois do treinamento<sup>(20)</sup>. Tais resultados convalidam os achados desta pesquisa, demonstrando que o nível de complexidade da imagem do traçado espectrográfico da voz é pequeno e que, portanto, sua análise visual é uma tarefa relativamente fácil.

Após o treinamento, verificou-se que os aspectos de intensidade do traçado e estabilidade do traçado, obtiveram melhora no nível de concordância intrassujeitos (Tabela 1) de moderada para boa; o aspecto presença de sub-harmônicos, de boa para quase perfeita e definição de harmônicos, de regular para moderada. Os demais aspectos mantiveram concordância quase perfeita. Tais resultados sugerem que, apesar da análise visual do traçado espectrográfico não ser uma tarefa de difícil execução, o treinamento visual melhora a performance do avaliador na análise de tais parâmetros.

A concordância intersujeitos (Tabela 2) nesta pesquisa, antes do treinamento, obteve valor estatisticamente significativo apenas para o aspecto presença de ruído, o que indica uma confiabilidade ruim para a maioria dos parâmetros. Tais achados foram semelhantes ao obtido no estudo de lâminas patológicas do esôfago de Barrett<sup>(18)</sup> e diferente do encontrado na identificação de pontos cefalométricos em tomografias computadorizadas *cone beam*<sup>(17)</sup>. Tais discordâncias podem ser justificadas pela diferença metodológica e de complexidade da imagem dos três estudos. Além disso, sugerem que a concordância interavaliador é menor que a concordância intra-avaliador na avaliação perceptivo-visual quando os mesmos não possuem treinamento. Esta diferença demonstra a importância de uma explicação mais detalhada dos critérios de avaliação, possibilitando estabelecer-se um consenso sobre os conceitos e graus de variação dos parâmetros visuais analisados.

Quanto à concordância interavaliadores (Tabela 2) após o treinamento, foi observado que todos os aspectos, exceto presença de ruído, apresentaram diminuição nos valores de *p*, indicando melhora na concordância. O estudo piloto, semelhante a esta pesquisa, também encontrou melhora na

confiabilidade em geral, após o treinamento<sup>(20)</sup>. No presente estudo, os parâmetros de estabilidade do traçado e presença de sub-harmônicos, passaram a apresentar concordância ótima. Os parâmetros de forma do traçado, intensidade do traçado e definição de harmônicos, também apresentaram melhora na concordância, porém, tais parâmetros não foram considerados significantes, possivelmente devido ao tamanho da amostra. Tais resultados indicam a importância do treinamento na confiabilidade da análise perceptivo-visual do traçado espectrográfico da voz, principalmente no que tange à concordância interavaliador.

Assim como outros tipos de treinamento visual descritos com oito examinadores avaliando espectrogramas de fala<sup>(8)</sup>; com médicos avaliando imunocintilografia<sup>(9)</sup>; com cinco docentes de Odontologia utilizando exame visual e radiográfico de amálgamas dentárias<sup>(10)</sup>; com três profissionais e graduandos de Odontologia utilizando radiografias dentárias para se analisar cáries<sup>(11)</sup> e com radiografias periodontais para se verificar perda óssea<sup>(16)</sup>, o treinamento visual realizado neste estudo melhorou a concordância da avaliação e deve ser considerado, portanto, um aspecto importante destas análises perceptivo-visuais.

O conhecimento adquirido em um treinamento pode ser comparado ao adquirido com a experiência profissional. Dessa forma, os resultados obtidos também foram semelhantes aos do estudo que analisou imagens mamográficas com 22 avaliadores experientes e inexperientes<sup>(21)</sup>, e a pesquisa que estudou a interpretação de tomografia computadorizada craniofacial por três estudantes de Odontologia de diferentes períodos<sup>(22)</sup>. Ambos os estudos observaram que sujeitos com mais experiência apresentam avaliações mais concordantes e confiáveis.

Verificou-se que a avaliação do parâmetro presença de ruído após o treinamento (Tabela 2) obteve aumento no valor de *p* (passando de 0,005 para 0,085), o que indica que a concordância desse aspecto piorou com o treinamento. Uma possível explicação seria a de que vozes normais podem apresentar certa quantidade de ruído e os avaliadores classificaram-nas como ruído ausente. Porém, a fronteira entre o normal e o alterado é muito tênue. Dessa forma, a concordância quanto à análise de presença de ruído pode ter piorado devido ao fato de os avaliadores estarem mais atentos e terem apresentado mais dúvidas em relação à quantidade de ruído que pode ser considerada normal ou alterada.

Durante a realização das avaliações desta pesquisa, o tempo para analisar cada espectrograma foi livre, sendo que as autoras esperaram que todos os avaliadores terminassem a análise de cada um, para então passar ao próximo. Observou-se que, antes do treinamento, a média de tempo utilizado para avaliar cada espectrograma foi de um minuto e meio e, após, passou a um minuto. Isso indica que, provavelmente, o treinamento contribuiu para que a análise ficasse mais fácil e, dessa forma, mais rápida.

Neste trabalho, as avaliações foram puramente visuais. Outro estudo<sup>(23)</sup>, realizado com seis examinadores experientes avaliando 70 vozes disfônicas, verificou que a concordância interavaliadores aumentou com a adição dos espectrogramas à informação auditiva, o que mostra a importância da utilização de ambas as formas de avaliação de voz, que trazem informações complementares. Dessa forma, sugere-se que

pesquisas futuras investiguem a influência da apresentação da voz na avaliação espectrográfica, para verificar se o inverso também ocorre.

Este estudo pesquisou a influência do treinamento visual apenas na avaliação de examinadores inexperientes. São necessários outros trabalhos para verificar se o mesmo ocorre com examinadores experientes.

Como não foram encontradas outras pesquisas semelhantes, que pudessem nortear o tempo de treinamento, optou-se por realizar um de quatro horas. Assim, seria interessante realizar o mesmo tipo de estudo com diferentes tempos de treinamento, avaliando se um maior ou menor tempo de treinamento poderia influenciar os resultados encontrados.

## CONCLUSÃO

O treinamento visual melhorou a concordância intra-avaliadores para os parâmetros de escurecimento do traçado, estabilidade do traçado, presença de sub-harmônicos, definição de harmônicos e manteve uma concordância quase perfeita nos aspectos de forma do traçado e presença de ruído. Na análise interavaliadores, o treinamento visual melhorou a concordância para os parâmetros de estabilidade do traçado e presença de sub-harmônicos e piorou a concordância para o parâmetro de presença de ruído.

O treinamento visual melhorou a confiabilidade inter e intra-avaliadores na interpretação de espectrogramas.

## ABSTRACT

**Purpose:** To verify whether visual training improves the reliability inter and intra-observer of spectrograms interpretation. **Methods:** This was a longitudinal experimental study that used spectrograms collected in a database of a private office, regarding the voices of men and women analyzed by the program Gram 5.0 in 60 dB scale, display scroll, palette BW, with frequency analysis performed in linear scale, narrow band, FFT size 1024, and frequency resolution 5.4 Hz. Twenty subjects, Speech-Language Pathology students, evaluated 25 spectrograms, of which five were repeated to evaluate intra-observer reliability, before and after a visual training of spectrograms analysis. The analyzed parameters were: spectrogram regularity, harmonic darkening, spectrogram stability, presence of noise components, harmonics definition and presence of sub-harmonics. The training session lasted four hours. Kappa statistic was used to evaluate intra-observer reliability, and the adjusted Chi-square test, to evaluate inter-observer reliability. **Results:** Intra-observer reliability of harmonic darkening, spectrogram stability, harmonic definition and presence of sub-harmonics improved, and the other aspects kept almost perfect agreement after the training. Inter-observer reliability of spectrogram stability and presence of sub-harmonics improved, and presence of noise components got worse after training. **Conclusion:** Visual training improved inter and intra-observers reliability in spectrograms interpretation.

**Keywords:** Voice; Speech acoustics; Sound spectrography; Observer variation; Training

## REFERÊNCIAS

- Behlau M, organizadora. Voz: o livro do especialista. Rio de Janeiro: Revinter; c2001.
- Eadie TL, Doyle PC. Classification of dysphonic voice: acoustic and auditory-perceptual measures. *J Voice*. 2005;19(1):1-14.
- 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.
- Drumond LB, Gama ACC. Correlação entre dados espectrográficos e perceptivo-auditivos de vozes disfônicas. *Fono Atual*. 2006;8(35):49-58.
- Ortiz KZ, Carrillo L. Comparação entre as análises auditiva e acústica nas disartrias. *Rev Soc Bras Fonoaudiol*. 2008;13(4):325-31.
- Patel RR. Visual perceptions in laryngeal imaging. *Perspect Voice Voice Disord*. 2007;17:7-10.
- Pisoni DB, Greene BG, Carrell TD. Identification of visual displays of speech: comparisons of naive and trained observers. *J Acoust Soc Am*. 1983;73(Suppl 1):S102.
- Greene BG, Pisoni DB, Carrell TD. Recognition of speech spectrograms. *J Acoust Soc Am*. 1984;76(1):32-43.
- Rubinstein M, VanDaele P, Wegener W, Guardia M. Adherence to imaging protocol and impact of training on the interpretation of CEA-Scan® (arcitumomab) imaging for colorectal cancer. *Rev Med Nucl Alasbimn J*. 2004;6(25): Article N° AJ25-5. [cited 2010 July 28]. Available from: <http://www2.alasbimnjournal.cl/alasbimn/CDA/imprime/0,1208,PRT%253D11064,00.html>.
- Fernandes ET, Ferreira e Ferreira E. Substitution of amalgam restorations: participative training to standardize criteria. *Braz Oral Res*. 2004;18(3):247-52.
- Carmona GP, Devito KL, Pontual MLA, Haiter Neto F. Influência da experiência profissional no diagnóstico radiográfico de cáries. *Cienc Odontol Bras*. 2006;9(1):87-92.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-74.
- Bele IV. Reliability in perceptual analysis of voice quality. *J Voice*. 2005;19(4):555-73.
- Deliyski DD, Shaw HS, Evans MK, Vesselinov R. Regression tree approach to studying factors influencing acoustic voice analysis. *Folia Phoniatr Logop*. 2006;58(4):274-88.
- Deliyski DD, Evans MK, Shaw HS. Influence of data acquisition environment on accuracy of acoustic voice quality measurements. *J Voice*. 2005;19(2):176-86.
- Lanning SK, Best AM, Temple HJ, Richards PS, Carey A, McCauley LK. Accuracy and consistency of radiographic interpretation among clinical instructors in conjunction with a training program. *J Dent Educ*. 2006;70(5):545-57.
- de Oliveira AE, Cevidanes LH, Phillips C, Motta A, Burke B, Tyndall D. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2009;107(2):256-65.

18. Lopes CV, Pereira-Lima JC, Hartmann AA, Tonelotto E, Salgado K. Displasia no esôfago de Barrett - concordância intra e interobservador no diagnóstico histopatológico. *Arq Gastroenterol.* 2004;41(2):79-83.
19. Armato SG 3rd, Roberts RY, Kocherginsky M, Aberle DR, Kazerooni EA, Macmahon H, et al. Assessment of radiologist performance in the detection of lung nodules: dependence on the definition of "truth". *Acad Radiol.* 2009;16(1):28-38.
20. Côrtes MG, Gama ACC. Análise visual de parâmetros espectrográficos pré e pós-fonoterapia para disfonias. *Rev Soc Bras Fonoaudiol.* 2010;15(2):243-9.
21. Ângelo MF, Schiabel H. Uma ferramenta para treinamento na avaliação de imagens mamográficas via internet. *Radiol Bras.* 2002;35(5):259-65.
22. Gaia BF, Perella A, de Cara AC, Antunes JL, Cavalcanti MG. CT interpretation of craniofacial anomalies: a comparative analysis by undergraduate dental students. *Braz Oral Res.* 2005;19(1):58-62.
23. 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.