



RESENHA

**POR UMA ANÁLISE CONSENSUAL E INTERDISCIPLINAR
DA QUALIDADE DE VOZ BASEADA EM PARÂMETROS
AUDITIVOS, ACÚSTICOS E FISIOLÓGICOS**

*Izabel Cristina Viola**

*Análise da qualidade vocal de um grupo de indivíduos disfônicos:
uma abordagem interpretativa e integrada de dados
de natureza acústica, eletroglotográfica e perceptiva.*

Zuleica Antonia de Camargo

Pontifícia Universidade Católica de São Paulo, tese de doutorado, Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem, sob a orientação da Profa. Dra. Sandra Madureira, 2002.

O estudo da voz requer a integração de muitas áreas, as quais, ao abordá-la, nem sempre conseguem sair do campo da subjetividade ou da análise isolada

* Fonoaudióloga, doutoranda em Linguística Aplicada e Estudos da Linguagem – PUC-SP; Faculdade de Fonoaudiologia da PUC-SP.

de dados. Este tem sido um dos maiores desafios dos fonoaudiólogos clínicos – integrar e interpretar os dados que observam na avaliação e durante o processo terapêutico, com teorias (e normas) estudadas.

Esta tese vem nos falar desses vínculos, de rigor metodológico, alertar para normas e medidas-padrão e, mais uma vez, lembrar que o aparelho fonador é adaptável e dotado de plasticidade. Apresenta a proposta de focar a qualidade vocal por meio de dois modelos teóricos com o objetivo de investigar os correlatos acústicos do sinal da fala e a sua correspondência nas esferas auditiva e fisiológica, de um grupo de indivíduos portadores de disфония, por paralisia de prega vocal. As teorias escolhidas foram a Teoria Acústica da Produção da Fala (modelo fonte-filtro para a produção das vogais), idealizado por Fant (1970), e Modelo Fonético de Descrição da Qualidade Vocal de Laver (1980, apud Camargo, 2002).

A autora caracteriza seu trabalho da seguinte maneira: “(...) uma abordagem interpretativa e integrativa, por considerar a esfera acústica como intermediadora entre a perceptiva e a fisiológica e por se pautar pela correlação de dados obtidos por vários instrumentais advindos dessas três esferas” (p. 231).

Salienta que é necessário entender melhor os mecanismos de sincronização e combinação de efeitos glóticos e supraglóticos, para avançar na apropriação do universo sonoro que permeia a qualidade vocal (p. 6), romper com a clássica dicotomia estabelecida entre a normalidade e a alteração vocal na dimensão da produção e percepção do sinal sonoro, para observar as conseqüências mais amplas dessa ruptura (p. 41).

Compreender a produção vocal neste contínuo, observando as transformações que o sinal sonoro sofre ao longo do percurso, sem dúvida facilita a identificação dos ajustes usados (compensatórios ou não), a escolha da abordagem terapêutica, o prognóstico e a qualidade de vida do cliente.

O campo do diagnóstico e tratamento das disfonias tem progredido rapidamente com as novas tecnologias disponíveis na área das Ciências da Fala e da Medicina. Entretanto, as informações obtidas nessas áreas têm de se combinar com a avaliação perceptivo-auditiva, devido à capacidade integradora desta. Entretanto, a tradicional avaliação de qualidade vocal é bastante discutida pela subjetividade e disparidade entre julgamentos. Com o desenvolvimento dos estudos

acústicos e da sua aplicabilidade aos dados clínicos, o grande desafio passou a ser correlacionar e interpretar os achados, mesmo “quando nem sempre há uma correlação entre ambos” (p. 3) e quando estudos com rigor metodológico na dimensão da patologia são escassos.

Nesse sentido, a autora escolhe a Teoria Acústica da Produção da Fala, pois ela permite

(...) a possível identificação de mobilizações de vários segmentos do aparelho fonador (...) desde o controle e modulação do fluxo de ar pelas pregas vocais, possibilitando a especial condição de decomposição do sinal de saída emanado pela boca do falante disfônico, sem a necessidade de registro a partir de meios invasivos ao trato vocal, os quais poderiam comprometer a qualidade do sinal registrado. (p. 28)

Tradicionalmente, a Fonoaudiologia enfoca a produção sonora e a qualidade vocal parcialmente independentes, ou seja, ora considerando somente características supraglóticas, ora glóticas e usando títulos subjetivos e discutíveis para denominá-la, como branca, velada, pastosa, molhada, etc... O modelo de Laver (1980, apud Camargo, 2002) de Descrição Fonética da Qualidade Vocal considera a qualidade vocal (característica individual do falante) como resultado da anatomia do aparelho fonador e dos ajustes musculares realizados pelo falante na atribuição de sentido.

A qualidade vocal é entendida como veículo de informações de características físicas psíquicas e sociais do falante e passa a ser estudada do ponto de vista de *settings* (ajustes) laríngeos e supralaríngeos, isto é, de traços recorrentes na posição do trato vocal conseqüente a um ajuste muscular de longa duração.

Assim, a descrição fonética da qualidade vocal em indivíduos disfônicos contribui para o entendimento dos planos alterados e preservados, de acordo com os princípios de interdependência e compatibilidade entre os ajustes, possibilitando a investigação das várias e refinadas compensações que podem ocorrer graças à plasticidade vocal do trato e afastando a comparação ou referência à normalidade, usada tradicionalmente na situação de disfonia.

Embora a autora não tenha tido a intenção de validar uma adaptação do protocolo para uso clínico, o modelo resultante tem aplicações para tal fim, já que foi discutido com parâmetros da área fonoaudiológica e para o português

brasileiro (p. 226). Este é um grande mérito do trabalho, que deve ser reconhecido pela comunidade científica, usando-o inclusive para o treinamento auditivo na formação de fonoaudiólogos.

Como a própria autora reconhece nas suas considerações finais, a adoção do Roteiro de Avaliação da Qualidade Vocal, com motivação fonética possibilitou

(...) transcender as limitações impostas por procedimentos de análise acústica pautados exclusivamente em estimativas numéricas e pelas descrições de ordem perceptiva essencialmente impressionísticas, além de pouco consistentes para a aplicação em diferentes contextos da prática clínica fonoaudiológica. (p. 231)

Chamam a atenção a leitura esmiuçada da literatura (a bibliografia ocupa 23 páginas, com 249 referências) e o rigor do método usado para o corpo do trabalho, incluindo vários estudos-piloto, realizados para validar o método de análise e assegurar a confiabilidade dos resultados. Assim, o trabalho principal incorpora vários outros, que certamente poderão ser publicados de forma independente.

Tal rigor é justificado pela autora devido ao fato de o sinal vocal alterado ser característica estrutural aperiódico, exigindo cuidado no delineamento e na condução dos procedimentos de coleta e análise de dados, começando pela garantia da qualidade das amostras coletadas (p. 197).

A descrição do método para a coleta das amostras é esmiuçada no Capítulo 5, em detalhes, o que nos permite uma visão abrangente dos critérios de composição do grupo estudado, da seleção do indivíduo de referência, do ambiente e dos instrumentos usados na gravação, do material e dos procedimentos adotados em todas as etapas da coleta das amostras: acústica e eletroglotográfica (realizadas simultaneamente), otorrinolaringológica e análise perceptiva feita por juízes. Com o mesmo cuidado, a autora trata, no Capítulo 6, dos métodos de edição, tratamento e análise dos dados. A metodologia da pesquisa ocupa quarenta 40 páginas (isto é, 17% do corpo da tese), dividida nos capítulos 5 e 6.

Isso se contrapõe veementemente ao que tem ocorrido na Fonoaudiologia: a utilização desenfreada de *softwares* destinados à inspeção de dados clínicos sem a preocupação com as condições de coleta e o rigor da análise do material e dessas condições sendo elaborados estudos normativos.

Foram estudados cinco sujeitos disfônicos devido à paralisia e fixação unilateral de prega vocal, em posição paramediana. O grupo foi criteriosamente selecionado, de forma a contemplar o equilíbrio entre as idades (54, 51 e 53), sexo (feminino), duração variada da sintomatologia, permitindo ainda a comparação tanto com um sujeito presbifônico (72 anos) e com o sujeito-referência (52 anos) sem alterações vocais.

A análise acústica envolveu os seguintes procedimentos: observação de padrões espectrais; com a extração de medidas de longo e curto termo e espectrografia de banda larga. No âmbito da espectrografia de banda larga, foram processados os estímulos eleitos para extração das medidas de tempo de início de sonorização (VOT).

As amostras de fala foram coletadas na fala encadeada (texto) e com a vogal [a], em padrão entrecortado e sustentado, em três padrões de frequência e intensidade: habitual, máxima e mínima, com e sem apoio manual nos eletrodos do eletroglotógrafo fixado no pescoço. A digitalização dos dados foi realizada em duas frequências de amostragem (22050 e 44100Hz) a fim de contemplar diferentes análises (p. 104).

Vale ressaltar que, diferentemente da rotina clínica de avaliação vocal, a vogal [a] foi selecionada para análise acústica a partir das pesquisas de Vieira et alii (1996a, b, 1997, apud Camargo, 2002), que demonstraram maior confiabilidade na extração das medidas de perturbação de frequência (*jitter*).

Por outro lado, as medidas de longo termo foram realizadas com texto lido (em três repetições) submetido à remoção das pausas marcadas por silêncio na emissão. A análise de texto neste tipo de medida mostrou-se importante na fidedignidade da fala espontânea (p. 134). Todos os dados acústicos foram submetidos a tratamento estatístico.

Os dados da fisiologia laríngea foram provenientes da eletroglotografia (conformação da onda Lx e medidas correspondentes) e do exame laringológico, particularmente a telaringovideostroboscopia e a vídeoquimografia, realizados no mesmo dia, após a coleta da amostra de fala.

Na eletroglotografia, a autora encontrou qualidade superior na emissão da vogal entrecortada e na emissão encadeada (a arara), quando comparadas com as emissões sustentadas. Para caracterizar e especificar a forma e a geometria da onda e irregularidades, foram extraídas as medidas de f_0 , coeficiente de contato, perturbação em frequência e amplitude, índice de velocidade e VOT. Estas serviram de contraponto às medidas de VOT extraídas por meio da análise acústica.

A análise perceptiva auditiva envolveu dois grupos diferentes de juízes, sendo um para a avaliação da qualidade vocal com motivação fonética (composto de quatro doutores experientes no modelo VPAS) e o outro (composto por fonoaudiólogos alunos de curso de especialização), para o julgamento de inteligibilidade de sílabas plosivas surdas e sonoras, iniciadas por consoantes, a partir de amostras de fala editadas e gravadas em CD-ROM e de roteiro impresso.

A dicotomia entre normal e patológico realmente foi deixada de lado, cedendo lugar à descrição de padrões adaptativos e, em diversas ocasiões, os dados do indivíduo de referência partilharam algumas semelhanças com aqueles dos sujeitos, como, por exemplo, quanto à ocorrência de ataques vocais bruscos e instabilidades da estrutura harmônica das vogais sustentadas.

A contribuição do enfoque fonético da qualidade vocal foi promissora, uma vez que revelou aspectos que não seriam salientados numa proposta de avaliação tradicional, tais como: ajustes de laringe (na dimensão vertical), lábios, mandíbula, língua e faringe (p. 206).

Os achados indicaram relevância dos aspectos acústicos de longo termo em correspondência àqueles de curto termo (estrutura harmônica, formântica e tempo de início de sonorização – VOT) e com o referencial do modelo fonético de qualidade vocal. (p. 202)

As características de fonação dos indivíduos apontadas pelos juízes no modelo fonético de qualidade vocal foram correlacionadas com as medidas de longo termo, que, por meio de seus traçados, permite, inclusive, o resgate das emissões

de um falante. Há padrões recorrentes na emissão de um mesmo falante, apesar das instabilidades e irregularidades, o que também confirma a individualidade nos ajustes compensatórios escolhidos por cada um deles (p. 205).

As medidas de espectrogramas de banda estreita, com os valores no eixo de frequência em 11KHz e 22 kHz, mostram a presença de ruído no sinal acústico de indivíduos que apresentam registro de harmônicos (p. 205) com a presença de subarmônicos (descritos também como bifurcações, característicos da voz de qualidade áspera, confirmados ainda nas medidas de curto termo). O interessante é observar a gradiência dessa tendência de substituição dos formantes por ruído, em função do maior comprometimento ou maior grau da paralisia dos sujeitos estudados, percebidas na análise perceptiva e fisiológica.

Na análise eletroglotográfica, não foram evidenciadas mudanças no traçado de onda Lx em função da pressão digital sobre os eletrodos durante as várias etapas do registro, conforme proposto por Coton e Conture (1990, apud Camargo, 2002, p. 215).

Foi encontrada concentração de ar não-sonorizado ao início e/ou final das emissões, por volta de 1500Hz, justamente na zona espectral de ganho de amplitude, por provável influência de cavidades subglóticas. Afirma a autora que

(...) tais achados denotam o avanço à construção do conhecimento sobre os mecanismos compensatórios e que poderá, inclusive, vir a causar impacto clínico, especialmente por considerar um sistema que gera muitas polêmicas em termos do campo da reabilitação das disfonias: a esfera respiratória e a modalidade de enfoque da dinâmica respiratória. (p. 215)

A autora aponta com propriedade que a formalização das medidas depende da captura do sinal, sua prévia inspeção e aplicação de procedimentos cuidadosamente delineados para tal finalidade, de forma a garantir que a ausência de registro ou a impossibilidade de extração de valores seja um dado confiável. Lembra justamente que o parâmetro a ser aferido é aquele que sofre maior influência do nível de irregularidade, que representa uma alusão ao quão paradoxais podem se tornar os resultados das medidas acústicas de perturbação. Conclui que os valores de *jitter* e *shimmer* mostraram-se aumentados em proporção à indicação de manifestações de alterações de mecanismos laríngeos, com base em estudo de Vieira (1997, p. 216).

Houve consenso entre as medidas de VOT extraídas pelas técnicas de eletroglotografia e análise acústica, especialmente nas condições de avaliação semi-automática (p. 210). O VOT, por não contemplar somente aspectos da dinâmica laríngea, destacou-se como indicativo na mobilização dos articuladores e ressoadores e como um índice promissor para a compreensão das limitações impostas pela disфонia. A autora salienta que ele poderia refletir a esfera de coordenação de eventos subglóticos (esfera respiratória), glóticos e supraglóticos. Note que o referido índice não tem sido explorado no campo de estudos da voz (p. 218).

Foram encontradas evidências mostrando que os maiores desvios de valores de VOT para plosivos sonoros ocorreram nos casos em que se detectou diminuição das condições vibratórias (ou praticamente ausência), evidências essas confirmadas pela fisiologia laríngea (p. 220).

As medidas de VOT apresentaram correlação com a inteligibilidade de fala (sílabas), além de revelarem que ajustes de longo termo da qualidade vocal influenciam padrões de emissão no plano segmental. Podemos exemplificar que o fechamento de mandíbula associado à retração do corpo de língua ou às tendências laríngeas de escape de ar, voz áspera e constrição faríngea dificultaram a percepção da sonoridade. Por outro lado, a presença de qualidade de voz denotando laringe alta, hiperfunção e aspereza não promoveu erros no julgamento dos contrastes surdo-sonoro (p. 220).

Segundo a autora, as dificuldades (expressas no VOT) parecem residir na sincronização entre os gestos glótico e oral, revelando a dificuldade de iniciar o processo de sonorização de modo preciso.

Vale registrar a identificação de valores de VOT zero, pouco usual no português, e para situação de sons surdos, na falante com maior alteração de ordem vibratória.

Este é um trabalho de leitura obrigatória para todos os fonoaudiólogos que trabalham ou pretendem trabalhar com *Análise Acústica e/ou Voz*. Os objetivos de interpretar e integrar os dados de natureza diversa são plenamente atendidos na leitura, que, às vezes, requer o ir e vir, o ler e reler, o parar e voltar, o pensar e o falar, próprios de uma leitura densa e rica. Dada a extensão e complexidade do trabalho, nenhuma resenha, por maior que seja, consegue de fato apreender toda sua importância.