 DOI: 10.47595/cjsiurj.v5i1

ANÁLISE DOS PROTOCOLOS VPAS E VPAS SIMPLIFICADO NA PERÍCIA DE COMPARAÇÃO DE LOCUTORES

ANALYSIS OF VPAS AND SIMPLIFIED VPAS IN SPEAKER COMPARISON FORENSICS

Renata Christina Vieira[1]

Tiago Rosa Pereira[2]

Resumo: Este artigo discute a utilização de dois protocolos de análise vocal, o VPAS (Vocal Profile Analysis Scheme) e o VPAS Simplificado, na comparação de falantes, especialmente em contextos forenses. No Brasil, peritos criminais em fonética forense são instruídos a utilizar o método combinado de análises perceptiva e acústica. O estudo foca na aplicabilidade do roteiro VPAS para identificação de falantes, baseado no Modelo Fonético de Descrição da Qualidade Vocal, bem como na comparação desse método com uma versão simplificada. O estudo usou amostras de fala de 10 indivíduos e aplicou ajustes fonatórios nos dois roteiros. Os resultados revelaram desafios na simplificação do VPAS, pois a redução dos ajustes e a ausência de graduação dificultaram a distinção entre características vocais distintas. Além disso, algumas simplificações ignoraram ajustes comuns na prática clínica. Os autores concluem que o VPAS simplificado pode comprometer a precisão na identificação de falantes, já que a falta de detalhamento e graduação limita a distinção entre diferentes características vocais. Eles destacam a importância de considerar a diversidade e complexidade da qualidade vocal ao simplificar os protocolos de análise para aplicação forense.

Palavras-chave: Percepção auditiva; Percepção de fala; Voz; Qualidade de voz.

[1] Docente de Pós graduação em Documentoscopia para Justiça – Faculdade Instituto Universitário do Rio de Janeiro – FIURJ – Rio de Janeiro/RJ. Docente da Pós-graduação em Voz – Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO – Rio de Janeiro/RJ. Assistente técnica em Fonética Forense. Fonoaudióloga Forense, especialização em Voz pelo CEFAC, mestrado em Linguística pela UERJ e doutorado em Linguística Aplicada e Estudos da Linguagem pela PUC-SP, MBA pela FGV em Gestão Estratégica, Aperfeiçoamento em Neurociências, Linguagem e Cognição pelo CEFAC, Formação em Coach Vocal pelo CEV e curso em Facial Action Coding System (FACS) pelo F-M Group International. E-mail: renatachristinavieira@gmail.com.

[2] Fonoaudiólogo graduado pela Universidade Federal do Rio de Janeiro (UFRJ), Especialista em Promoção da Saúde e Desenvolvimento Social pela Escola Nacional de Saúde Pública Sergio Arouca (ENSP/Fiocruz), Mestrado em Pesquisa Clínica em Doenças Infecciosas pelo Instituto Nacional de Infectologia Evandro Chagas (INI/Fiocruz) e Doutorado em Ensino em Biociências e Saúde pelo Instituto Oswaldo Cruz (IOC/Fiocruz). Professor de Oratória, Coordenador e Docente da Especialização em Voz da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). E-mail: dr.tiagorosap@gmail.com.

Abstract: *This article delves into the use of two vocal analysis protocols, VPAS and Simplified VPAS, in speaker comparison, particularly in forensic contexts. In Brazil, criminal experts in forensic phonetics are trained to use a combined method of perceptual and acoustic analyses. The study focuses on the applicability of the VPAS script for speaker identification based on the Phonetic Model of Vocal Quality Description and compares this method with a simplified version. The study used speech samples from 10 individuals and applied phonatory settings both scripts. The results revealed challenges in simplifying the VPAS, as reducing adjustments and lacking grading made it difficult to distinguish between distinct vocal characteristics. Furthermore, some simplifications disregarded common adjustments in clinical practice. The authors conclude that the Simplified VPAS may compromise accuracy in speaker identification since the lack of detail and grading limits the distinction between different vocal traits. They emphasize the importance of considering vocal quality's diversity and complexity when simplifying analysis protocols for forensic application.*

Keywords: *Auditory perception; Speech perception; Voice, Voice Quality.*

Introdução

No contexto forense, a padronização de técnicas entre peritos em identificação de falantes é uma busca cada vez mais frequente, visto que não existe consenso sobre os procedimentos metodológicos a serem adotados na área (GOLD, FRENCH, 2011). Devido à alta complexidade desse tipo de perícia, os métodos mais utilizados internacionalmente são a combinação das análises perceptiva e acústica. No Brasil, a Capacitação Nacional para Peritos Criminais em Fonética Forense também instrui os peritos a utilizar o método combinado (TONACO, SILVA, 2016), adicionando a sua investigação a aplicabilidade de um protocolo de avaliação perceptiva.

O objetivo geral deste estudo foi o de verificar a aplicabilidade do *Vocal Profile Analysis Scheme* (VPAS) no reconhecimento de falantes, a partir da reflexão acerca dos resultados de um experimento perceptivo baseado no Modelo Fonético de Descrição da Qualidade Vocal (LAVER, 1980) e de dinâmica vocal.

Os protocolos de análise perceptiva são comumente aplicados na clínica fonoaudiológica, visto que esse tipo de investigação é considerado padrão ouro para avaliação da voz (DEJONCKERE ET AL, 2001). A aplicabilidade de avaliações perceptivas para uso forense tem sido um assunto pouco abordado na literatura. Para o uso da análise perceptiva na perícia de identificação de falantes, consideramos primordial a escolha de um instrumento capaz de descrever vozes da maneira mais minuciosa possível.

De acordo com Eriksson (2012)[3], o protocolo VPAS é utilizado com fins forenses na Alemanha, no Brasil e no Reino Unido. Segundo o autor, na Alemanha, a polícia (*Bundeskriminalamt (BKA)*) emprega uma variante simplificada do protocolo associada ao protocolo GRBAS - *Grade, Roughness*, *Breathiness*, *Asteny* and *Strain* – (Isshiki et al, 1966; Hirano, 1981).

O GRBAS é utilizado na avaliação da fonte glótica (nível das pregas vocais), enquanto o VPAS é empregado para analisar os ajustes supraglóticos (acima das pregas vocais). Acreditamos que a integração dessas duas avaliações pode ser motivada por diversos fatores possíveis. Dificuldades no treinamento auditivo para usar o VPAS, especialmente entre profissionais linguistas, podem fazer com que os peritos prefiram utilizar o GRBAS.

A maior familiaridade da polícia alemã com o GRBAS pode proporcionar uma sensação de segurança ao aplicá-lo. No entanto, a natureza exclusivamente fonatória do GRBAS não engloba completamente a avaliação perceptiva no contexto forense, justificando assim o uso do VPAS como um complemento na análise dos ajustes supraglóticos. Além disso, é possível que a polícia alemã esteja passando por um período de adaptação e transição entre protocolos.

No Reino Unido, Foulkes, French e colaboradores (2015) exploraram métodos de caracterização de vozes. No estudo, além de empregarem algumas medidas acústicas, como a Mel-Frequency Cepstral Coefficients (MFCC) e distribuições de formantes de longo termo (LTFD), os pesquisadores investigaram a análise perceptiva para fins forenses por meio do roteiro VPAS. Os autores chegaram à conclusão de que, embora as medidas acústicas fossem semelhantes entre si, a análise perceptiva oferecia informações complementares, extremamente relevantes para a caracterização vocal. Para isso, estão utilizando uma versão adaptada do VPAS.

No Brasil, a Capacitação Nacional para Peritos Criminais em Fonética Forense orienta os peritos a empregar o método combinado. Segundo Tonaco e Silva (2016), alguns peritos da Polícia Federal passaram a adotar o VPAS (ANEXO 1) na perícia de Identificação de Falantes.

San Segundo e Mompean (2017) desenvolveram um protocolo perceptivo simplificado para avaliar a qualidade vocal, baseado no VPAS, visando "reduzir problemas típicos associados à multidimensionalidade da qualidade da voz e permitir uma quantificação fácil da similaridade dos falantes" (2017). No estudo, os pesquisadores avaliaram as vozes de 24 falantes masculinos (12 pares de gêmeos monozigóticos) falantes do espanhol Peninsular

[3] Anders Eriksson apresentou o panorama atual sobre o uso do protocolo VPA no contexto forense, no Brasil e no mundo, em sua palestra intitulada "*Vocal Profile Analysis, part 1: what kind of scales are there?"* no *Workshop on Vocal Profile Analysis*, evento que ocorreu em abril de 2016, na Unicamp.

Padrão, usando uma versão simplificada do VPAS (ANEXO 2). A partir das classificações perceptivas, foi calculado um índice de semelhança entre os pares de gêmeos e os pares não-gêmeos. Conforme os autores, os resultados indicam que, em média, os pares de gêmeos apresentam maior semelhança entre si do que os pares não-gêmeos.

O objetivo deste artigo é refletir acerca do uso do VPAS simplificado e apresentar a comparação dos dois protocolos de análise perceptiva em relação aos aspectos fonatórios, a fim de analisarmos o uso destas ferramentas na tarefa de comparação de locutores. Neste estudo, a escolha de focar exclusivamente nos aspectos fonatórios ao comparar protocolos perceptivos foi fundamentada na necessidade de uma análise detalhada e comparativa dos elementos fonatórios presentes nesses protocolos. Tal abordagem visa aprofundar a compreensão de como diferentes protocolos avaliam e interpretam os aspectos fonatórios da fala humana, preenchendo uma lacuna específica na literatura especializada e demonstrando o quanto cada protocolo pode ou não contribuir na perícia de comparação de locutores.

Considerações sobre os protocolos VPAS e VPAS simplificado

O pioneiro modelo fonético descritivo da qualidade vocal foi desenvolvido por Laver (1980), representando um marco nos estudos fonéticos ao introduzir uma unidade analítica para descrever as qualidades vocais, alterando significativamente a perspectiva das análises impressionistas anteriormente utilizadas (VIEIRA, 2018)

Com base nos fundamentos teóricos desse modelo fonético para a descrição das qualidades vocais, foi proposto, aplicado e aprimorado ao longo dos anos um roteiro de avaliação perceptiva para analisar tais qualidades, o VPAS. Esta evolução culminou na versão de Laver e Mackenzie-Beck em 2007, uma ferramenta que proporciona meios para avaliar, de forma perceptiva, os ajustes na qualidade e na dinâmica vocal. Esse roteiro foi posteriormente traduzido e adaptado para o português por Camargo e Madureira em 2008002E

O roteiro VPAS foi elaborado com base no modelo fonético de qualidade vocal, cuja unidade de análise é o *setting* ou ajuste. O ajuste pode ser compreendido como a tendência do falante em manter uma postura fonatória ou articulatória por um período específico. Todos os ajustes são propostos e descritos em relação a um ajuste de referência, chamado de ajuste neutro. Esse roteiro viabiliza uma avaliação perceptiva da qualidade vocal por meio das seguintes categorias de ajustes: trato vocal, fonatórios, tensão muscular e aspectos prosódicos.

O VPAS utiliza uma escala numérica de 1 a 6 para avaliar os graus de ajustes. O grau "1" está relacionado a uma discreta diferença em relação ao ajuste neutro, o "2" indica uma leve

diferença, enquanto o "3" representa uma diferença moderada. Estes graus de 1 a 3 referem-se a ajustes compatíveis com uma fala normal. Já os graus "4", "5" e "6" estão associados a alterações vocais, sendo o "4" uma diferença notável em relação ao ajuste neutro, o "5" uma diferença marcante e o "6" uma diferença extrema.

A aplicação do roteiro é dividida em duas etapas. Na primeira, a amostra de fala é ouvida para identificar os ajustes não neutros em quatro categorias: trato vocal, ajustes fonatórios, tensão muscular e elementos de dinâmica vocal. A segunda etapa é subdividida em duas partes. Na primeira, o juiz reouve a amostra, focando em segmentos específicos para identificar como os ajustes não neutros são produzidos. Na segunda parte, cada ajuste não neutro é classificado com base no desvio da neutralidade, utilizando a escala de 1 a 6.

É importante notar como os graus são identificados em diferentes categorias. Nas categorias de trato vocal e tensão muscular, o juiz escolhe entre os graus 1 e 3 para ajustes moderados e entre 4 e 6 para ajustes severos. Na categoria fonatória, a gradação é relacionada ao ajuste neutro: graus de 1 a 3 quando o ajuste modal predomina e de 4 a 6 quando outro ajuste associado prevalece em relação ao ajuste modal.

Devido à sua amplitude e capacidade de proporcionar uma descrição vocal precisa e detalhada, o Modelo Fonético de Descrição de Qualidade Vocal faz do roteiro VPAS um instrumento frequente em pesquisas de variados propósitos. Conforme Camargo e Madureira (2008, p. 94):

> a aplicabilidade do VPAS em nosso meio pode significar um importante avanço nas atividades clínicas e científicas na área de voz, cobrindo as necessidades daqueles que procuram por reabilitação e assessoria vocal, que estudam as variantes regionais e o reconhecimento de falantes.

Mackenzie-Beck (2005) explora várias aplicações do roteiro, incluindo a definição de padrões de normalidade, estudo da qualidade vocal na interação mãe-filho, análise em diferentes populações, correlações entre qualidade vocal e emoções, investigações sobre a interpretação de intenções com base nas características vocais, ensino de línguas estrangeiras e teatro, fonética forense e aplicações clínicas. Além disso, a autora destaca possíveis temas para futuras pesquisas, como dados normativos em diferentes idiomas, características associadas a sotaques, diferenças entre gêneros e faixas etárias, estudo de aspectos tonais e comunicação não verbal, e pesquisas mais detalhadas visando aplicações clínicas.

San Segundo e Mompean (2017) defendem que o VPAS simplificado é um protocolo confiável para caracterizar perceptivamente a qualidade vocal, sendo uma ferramenta útil para

avaliar a similaridade entre falantes. Nesta versão, os autores reduziram o número de configurações do roteiro original e eliminaram as graduações, acreditando que essa versão simplificada tem potencial em áreas como a fonética forense, pesquisa vocal e prática profissional.

Eles resumem as simplificações feitas no VPAS:

- Redução de 36 para 22 configurações;
- 10 tipos principais de ajustes com 22 configurações possíveis dentro desses tipos;
- Ausência de graduação, apenas uma classificação binária (neutro/não neutro);
- Ausência de marcação para ajustes intermitentes;
- Possibilidade de incluir descrições holísticas.

No roteiro do VPAS simplificado (ANEXO 2), em relação aos ajustes fonatórios, o juiz deve escolher entre "whisper/breathy", "N" (ajuste neutro) ou "creaky/harsh". Por exemplo, para descrever uma voz que no VPAS original seria considerada áspera no grau 2, o juiz a marcaria como "creaky/harsh", sem a opção de especificar o grau. Isso impede de saber apenas pelo roteiro se a voz é áspera ou crepitante (ou se possui ambas as características), e em qual grau. Sobre o VPAS simplificado, San Segundo e Mompean (2017) defendem que é um protocolo confiável para a caracterização perceptiva de qualidade de voz, e que pode ser uma ferramenta útil para a avaliação da similaridade do falante. Nessa versão, os autores reduzem o número de configurações do roteiro original e retiram as graduações, e acreditam que o roteiro em versão simplificada apresenta potencial em campos como fonética forense, nas áreas relacionadas de pesquisa de voz e prática profissional.

Método

Todos os participantes deste estudo consentiram voluntariamente, assinando o Termo de Consentimento Livre e Esclarecido (TCLE). O projeto de pesquisa recebeu aprovação do Comitê de Ética e Pesquisa da Pontifícia Universidade Católica de São Paulo - PUC-SP, sob o registro número 2.153.565.

Na fase inicial, amostras de fala semi-espontânea foram gravadas por um grupo de indivíduos denominados 'falantes audiogravados'. Este grupo consistia em 10 homens, com idade variando entre 34 e 49 anos, todos com ensino médio como nível mínimo de escolaridade. Esses falantes audiogravados residem no estado do Rio de Janeiro e são militares do Corpo de Bombeiros Militar do Estado do Rio de Janeiro (CBMERJ). Critérios de exclusão foram

aplicados: presença de problemas vocais ou auditivos, e o uso da voz para fins profissionais, como locução, radiodifusão, atuação, advocacia, Fonoaudiologia, docência, telemarketing, atividades eclesiásticas.

As informações sociolinguísticas dos participantes foram obtidas por meio de um questionário específico, denominado questionário sociolinguístico. Este questionário abordou tópicos como nome, idade, escolaridade, cidade de nascimento, locais de residência, origens dos pais, ocupação profissional e a presença de qualquer problema vocal ou auditivo. Caso positivo, os participantes foram solicitados a descrever suas dificuldades, e se bilíngues, a identificar suas línguas de proficiência.

O quadro 1 a seguir, apresenta os perfis sociolinguísticos dos falantes audiogravados.

Sujeito	Idade	Escolaridade	Cidade onde nasceu	Cidades onde morou	Cidade atual	Onde os pais nasceram e viveram	Profissão	É bilíngue
1	41	Ensino médio	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	Maranhão Minas Gerais Rio de Janeiro	Bombeiro militar	Não
2	34	Superior completo	Brasília	Brasília Rio de Janeiro	Rio de Janeiro	Brasília Goiás Rio de Janeiro	Bombeiro militar	Não
3	48	Superior completo	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	Recife Rio de Janeiro	Bombeiro militar e professor de educação física (não atua)	Não
4	40	Ensino médio	Rio de Janeiro	Rio de Janeiro Niterói Valença São Gonçalo	Rio de Janeiro	Espírito Santo Rio de Janeiro	Bombeiro militar	Não
5	40	Ensino médio	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	Rio Grande do Norte Brasília Rio de Janeiro	Bombeiro militar	Não
6	43	Superior incompleto	Rio de Janeiro	Rio de Janeiro Duque de Caxias	Rio de Janeiro	Rio de Janeiro	Bombeiro militar	Não
7	41	Pós-graduado	Rio de Janeiro	Niterói Rio de Janeiro São Gonçalo	Rio de Janeiro	Rio de Janeiro	Bombeiro militar	Não
8	45	Superior completo	Duque de Caxias/RJ	Duque de Caxias Mari/PB Rio de Janeiro	Duque de Caxias	Paraíba Rio de Janeiro	Bombeiro militar	Não
9	49	Superior completo	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	São Paulo Rio de Janeiro	Bombeiro militar	Não
10	38	Superior completo	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro	Bombeiro militar	Não

Quadro 1 – Perfil sociolinguístico dos falantes audiogravados.

Os equipamentos utilizados para as gravações foram: gravador portátil Zoom H5, celular Samsung Galaxy A5 com sistema Android e processador Quad Core de 1,2 GHz, e o aplicativo de gravação telefônica AutomaticCall Recorder. As gravações seguiram a frequência de amostragem de 44.100 Hz, 16 bits e foram salvas em extensão .wav, alinhando-se com os padrões adotados no Laboratório Integrado de Análise Acústica e Cognição (LIAAC) da PUC-SP.

Foram coletadas amostras de fala semi-espontânea do grupo de falantes audiogravados. Cada falante leu a frase 'o nosso verdadeiro objeto de estudo é essa poderosa e complexa face sonora da linguagem: a fala', além de fornecer duas narrativas sobre experiências pessoais significativas. Para manter a uniformidade na coleta das amostras, cada narrativa pessoal teve uma duração média de um minuto nas duas gravações.

Na segunda fase do experimento, Juízes A, experientes na área, realizaram a tarefa de avaliação perceptiva. Eles descreveram a qualidade vocal dos falantes audiogravados utilizando o protocolo VPAS. Essas descrições foram utilizadas para elaborar os perfis de qualidade vocal das amostras de fala coletadas na fase 1. As descrições desses perfis foram estabelecidas seguindo critérios específicos:

i) O ajuste foi considerado na descrição quando duas das três juízas identificaram o mesmo ajuste e graduação.

ii) Se apenas uma juíza identificou o ajuste, não foi considerado na descrição.

iii) Em casos em que duas juízas identificaram o mesmo ajuste, mas com graduações diferentes, foi calculada a média da gradação.

Na fase 3, os autores converteram a descrição da qualidade vocal dos aspectos fonatórios realizadas pelas juízas com base no VPAS para descrições com base no VPAS simplificado, com o objetivo de realizar a comparação.

No escopo metodológico deste estudo, a delimitação para análise exclusiva do aspecto fonatório foi estabelecida para garantir uma comparação direta e específica entre os protocolos perceptivos selecionados. Essa abordagem focada permite uma investigação pormenorizada dos elementos fonatórios, eliminando interferências de outras variáveis e oferecendo maior compreensão das diferentes abordagens adotadas por esses protocolos na avaliação fonatória.

Resultados

Os perfis de qualidade vocal, construídos a partir das minuciosas descrições realizadas pelas juízas por meio do Protocolo de Avaliação Perceptiva da Voz (VPAS), estão detalhadamente apresentados no quadro a seguir.

Perfil de qualidade vocal do falante audiogravados	Aspectos do trato vocal	Aspectos fonatórios	Aspectos de dinâmica vocal	Aspectos prosódicos
1	lábios com extensão diminuída grau 2, mandíbula aberta e com extensão diminuída em grau 1, corpo de língua abaixado em grau 1, altura de laringe abaixada em grau 1.	voz crepitante em grau 1.	pitch habitual abaixado em grau 2, pitch com extensão diminuída em grau 2.	----
2	lábios com extensão diminuída grau 1, mandíbula aberta em grau 1, ponta de língua avançada em grau 2, corpo de língua elevado em grau 1, constrição faríngea em grau 1, hiperfunção de trato vocal em grau 1.	----	na história 2, adotou ajustes diferentes com tremor e irregularidade laríngea e continuidade interrompida (quando foi utilizada a história 2 na tarefa perceptiva, os ajustes prosódicos foram incluídos na descrição).	----
3	mandíbula aberta em grau 1, ponta de língua avançada em grau 1, corpo de língua abaixado em grau 1, altura de laringe abaixada em grau 1.	----	voz crepitante em grau 1 e áspera em grau 2.	taxa de elocução rápida em grau 2 (e loudness habitual aumentado em grau 1 na leitura de frase. Quando a frase foi utilizada na tarefa perceptiva, o ajuste loudness habitual aumentado foi incluído na descrição).
4	língua recuada em grau 1, corpo de língua recuado em grau 2 e elevado em grau 1, mandíbula com extensão aumentada em grau 1, constrição faríngea em grau 2, hiperfunção de trato vocal em grau 2, hiperfunção laríngea em grau 2.	voz crepitante em grau 2 e áspera em grau 2.	taxa de elocução lenta em grau 1 e pitch habitual abaixado em grau 1.	----
5	lábios com extensão diminuída em grau 1, mandíbula aberta em grau 1, ponta de língua avançada em grau 1, corpo de língua com extensão diminuída em grau 1, constrição faríngea em grau 1, hiperfunção de trato vocal em grau 1.	voz áspera em grau 1.	taxa de elocução lenta em grau 1, pitch com variabilidade diminuída em grau 1 e pitch habitual elevado em grau 1. Loudness habitual abaixado em grau 1 na leitura de frase (que foi incluído na descrição quando a frase foi utilizada na tarefa).	----

6	lábios com extensão diminuída em grau 1 e protraídos em grau 2, mandíbula com extensão aumentada em grau 1, ponta de língua recuada em grau 1, corpo de língua abaixado em grau 1 e recuado em grau 1, altura de laringe abaixada em grau 1, hiperfunção de trato vocal em grau 1, hiperfunção laríngea em grau 1.	voz crepitante em grau 1 e áspera em grau 1.	pitch com extensão diminuída em grau 1 e com variabilidade diminuída em grau 1.	----
7	lábios estirados em grau 1, mandíbula com extensão aumentada em grau 1, ponta de língua avançada em grau 2, corpo de língua avançado em grau 1 e abaixado em grau 1, expansão faríngea em grau 2, laringe abaixada em grau 1.	----	pitch com variabilidade aumentada em grau 1, Lourdes habitual diminuído em grau 2.	----
8	lábios com extensão aumentada em grau 1, mandíbula com extensão aumentada em grau 1, ponta de língua avançada em grau 2, constrição faríngea em grau 1, altura de laringe elevada em grau 2, hiperfunção laríngea em grau 1.	voz soprosa em grau 2 e áspera em grau 1.	taxa de elocução rápida em grau 1 e pitch habitual elevado em grau 2.	----
9	ponta de língua avançada em grau 1, corpo de língua elevado em grau 1 e com extensão diminuída em grau 2, mandíbula aberta em grau 1, hiperfunção de trato vocal em grau 1, laringe alta em grau 2.	voz crepitante em grau 1	pitch habitual elevado em grau 2 e loudness habitual elevado em grau 2.	----
10	lábios com extensão diminuída em grau 1, mandíbula com extensão diminuída em grau 1, ponta de língua avançada em grau 2.	----	loudness habitual diminuído em grau 1.	----

Quadro 2 - Descrições de qualidade vocal dos participantes, utilizando o VPAS.

O Quadro 3 apresenta os resultados da análise dos aspectos fonatórios de 10 perfis de qualidade vocal, utilizando os protocolos VPAS e VPAS simplificado. Cada perfil é identificado por um número e, para cada um, são listados os ajustes fonatórios encontrados,

acompanhados das descrições correspondentes nos dois protocolos. Essa comparação direta entre as avaliações fonatórias realizadas pelos dois métodos permite uma análise detalhada das diferenças e semelhanças nas classificações dos ajustes vocais feitas pelos juízes para cada perfil.

Perfis	VPAS	VPAS simplificado
Perfil 1	Voz crepitante em grau 1	*Creaky*/*harsh*
Perfil 2	Voz modal	N
Perfil 3	Voz crepitante em grau 1 e áspera em grau 2	*Creaky*/*harsh*
Perfil 4	Voz crepitante em grau 2 e áspera em grau 2	*Creaky*/*harsh*
Perfil 5	Voz áspera em grau 1	*Creaky*/*harsh*
Perfil 6	Voz crepitante em grau 1 e áspera em grau 1	*Creaky*/*harsh*
Perfil 7	Voz modal	N
Perfil 8	Voz soprosa em grau 2 e áspera em grau 1	*Whisper*/*breathy*
Perfil 9	Voz crepitante em grau 1	*Creaky*/*harsh*
Perfil 10	Voz modal	N

Quadro 3 – Ajustes fonatórios dos 10 perfis de qualidade vocal dos falantes audiogravados nesta pesquisa, utilizando o VPAS e o VPAS simplificado.

Discussão

A concentração específica no aspecto fonatório, conforme investigado neste estudo, foi essencial para oferecer uma análise mais precisa e clara das diferenças e semelhanças entre os protocolos perceptivos examinados em relação aos aspectos fonatórios. Esta abordagem permitiu uma interpretação mais direta dos resultados fonatórios, destacando de forma nítida o impacto das distintas abordagens adotadas por cada protocolo na avaliação desses elementos específicos.

Ademais, vale ressaltar que, além da análise focada no aspecto fonatório, este estudo também contemplou análises dos demais tipos de ajustes presentes nos protocolos perceptivos selecionados. A avaliação minuciosa de outros ajustes, como os supraglóticos e de tensão muscular, proporcionou uma visão abrangente das abordagens adotadas pelos protocolos na análise da qualidade vocal. Essa análise multifacetada revelou nuances importantes sobre a eficácia e a sensibilidade dos protocolos na identificação e avaliação de diferentes aspectos da qualidade vocal, contribuindo para uma compreensão mais holística e detalhada das técnicas de análise vocal empregadas. A possibilidade de graduação e de detalhamento da combinação de ajustes fonatórios torna fácil a identificação das diferenças existentes entre as amostras quando utilizamos o VPAS. Observando o Quadro 2 é fácil distinguir os perfis e perceber que em relação aos ajustes fonatórios, os perfis que apresentam exatamente as mesmas características

fonatórias são: perfis 1 e 9 (voz crepitante em grau 1) e perfis 2 , 7 e 10 (voz modal). Os perfis 3, 4, 5, 6, e 8 apresentam ajustes fonatórios diferentes entre os demais.

Utilizando o VPAS simplificado (Quadro 3), temos dois grupos de perfis com características em comum: o grupo N (perfis 2, 7 e 10), o grupo "*Creaky*/*harsh*" (perfis 1, 3, 4, 5, 6 e 9), e o grupo "*Whisper*/*breathy*" (perfil 8).

O grupo de perfis classificado pelo VPAS simplificado como "*creaky*/*harsh*" apresenta ajustes isolados, combinados, e com graduações distintas. O perfil 8, apesar de apresentar voz áspera em grau 1, é classificada como "*whisper*/*breathy*" porque apresenta voz soprosa em grau 2.

Observando apenas os ajustes fonatórios, foi possível perceber que o objetivo de "reduzir problemas típicos associados à multidimensionalidade da qualidade da voz e permitir uma quantificação fácil da similaridade dos falantes" (SAN SEGUNDO, MOMPEAN, 2017), dificulta o trabalho de avaliação perceptiva para a tarefa de comparação de locutores, uma vez que o roteiro restringe as possibilidades de classificação dos ajustes.

Encontramos no estudo de San Segundo e Mompean (2017), equívocos quanto a questões fisiológicas da fonação. De acordo com os autores,

> Voice type (ie, phonation features) is probably the setting for which SVPA is less suitable, or at least that for which more training is required to improve agreement. Combined phonatory qualities are frequent. Laver mentions some of them: "harsh whispery voice" or "harsh creaky voice", for instance. The latter does not cause any problem in our SVPA, as both harsh and creaky belong to the tense larynx typology (SAN SEGUNDO, MOMPEAN, 2017).

Na discussão da pesquisa, os autores confirmam que a descrição de ajustes fonatórios do VPAS simplificado não é adequada, mas incorrem em um equívoco que demonstra falta de compreensão da produção dos tipos de vozes.

A voz crepitante não, necessariamente, é produzida com tensão, tanto que esse tipo de voz foi introduzido com finalidade terapêutica por Boone e McFarlane (1988), com o nome de som basal. O exercício consiste na emissão relaxada da vogal "a". O fluxo aéreo mínimo propicia a produção do som por longo tempo. O exercício som basal comprovadamente beneficia a melhor coaptação glótica, e pode ser utilizado com objetivo de redução ou fechamento de fendas glóticas. Além disso, reduz a tensão de laringe e aumenta a amplitude de vibração da mucosa (CARRARA, 1991).

Analisando os ajustes de trato vocal, é possível identificar outros problemas no que tange a fundamentação para fusão de ajustes. De acordo com os autores, a fusão dos ajustes de corpo

de língua foi realizada por existir uma tendência do ajuste de corpo de língua avançado vir acompanhado do ajuste de corpo de língua elevado e, ao contrário, do ajuste de corpo de língua recuado vir acompanhado do ajuste de corpo de língua abaixado: "*if he is non-neutral as concerns the lingual body, he will either tend to present a fronted and raised tongue body or a backed and lowered tongue body* " (SAN SEGUNDO, MOMPEAN, 2017).

Dos perfis de qualidade vocal dos 10 falantes audiogravados desta pesquisa, 8 apresentam ajustes não neutros de corpo de língua, e, desses 8 perfis, encontramos dois com ajustes de corpo de língua que fogem da tendência apontada pelos autores. Como foi possível observar no Quadro 2, o falante 4 apresenta corpo de língua recuado em grau 2 e elevado em grau 1, enquanto o falante 7 apresenta corpo de língua avançado em grau 1 e abaixado em grau 1.

Outra redução que consideramos equivocada é a dos ajustes de extensão diminuída e aumentada de lábios, mandíbula e corpo de língua. Dizem os autores:

> Finally, all the extensive and minimized range variants in Mackenzie Beck (ie, extensive and minimized mandibular, labial, or lingual setting) were discarded, as they were deemed to be covered by other settings: "open jaw" can be used to describe all extensive configurations and "close jaw" the minimized configurations (SAN SEGUNDO, MOMPEAN, 2017).

Assim, de acordo com os autores, os ajustes de extensão diminuída são, no VPAS simplificado, marcados como ajustes de mandíbula fechada e, os ajustes de extensão aumentada, marcados como ajustes de mandíbula aberta. Se as juízas tivessem utilizado o VPAS simplificado para traçar os perfis de qualidade vocal deste estudo, teriam dificuldade na descrição do perfil vocal do sujeito 5, que apresenta, nos aspectos de trato vocal, entre outros ajustes, lábios com extensão diminuída em grau 1, mandíbula aberta em grau 1 e corpo de língua com extensão diminuída em grau 1.

> Com a justificativa de que os ajustes de trato vocal, mandíbula protraída e labiodentalização são raros, os autores retiraram estas características do roteiro In fact, in San Segundo et al none of these two settings were found in a normophonic population of 100 male speakers of Standard Southern British English, aged 18–25 (DyViS corpus). Because of its low incidence also in Spanish, those non-neutral configurations were discarded from the mandibular and labial settings in the SVPA protocol (SAN SEGUNDO, MOMPEAN, 2017).

Entendemos que a retirada de ajustes como, por exemplo, os de mandíbula protraída e labiodentalização, com a justificativa de serem "raros", são um equívoco, pois a experiência clínica em Fonoaudiologia mostra que tais ajustes ocorrem com frequência (ANDRADE F.V.

et al., 2005; MACIEL C.T.V. et al., 2006; DAENECKE S. et al., 2006; TAUCCI R.A. e BIANCHINI E.M.G, 2007; ALÉSSIO C.V. et al., 2007; PEREIRA A.C. et al., 2005; GENARO K.F. et al., 2009; METZGER A.L.T. et al., 2009; SÍGOLO S. et al., 2009; MEZZOMO C.L. et al., 2011; MARTINELLI R.L.C., 2011; PEREIRA J.B.A. e BIANCHINI E.M.G., 2011).

Outra questão importante a ser analisada, no que se refere ao modelo reduzido do VPAS, é a falta de graduação dos ajustes. Para exemplificarmos o equívoco que esta falta provoca, podemos imaginar um exame de identificação de falantes, em que duas vozes são descritas como do tipo soproso, mas uma apresenta soprosidade em grau leve, e a outra está próxima da afonia. O uso do roteiro VPAS em versão reduzida, por não ter a devida graduação, pode levar a equivalência entre vozes com características distintas.

Merece atenção a correlação estabelecida entre os autores entre trato vocal relaxado e taxa de elocução lenta, e trato vocal tenso e taxa de elocução rápida, que consideramos, no mínimo, equivocada. De acordo com San Segundo e Mompean (2017):

> Indeed, the number of possible articulatory settings that would be associated with either lax or tense vocal tract is quite large (eg, different degrees of nasality and pharyngeal constriction). Furthermore, prosodic aspects seem to be associated with vocal tract tension, with faster tempo characterizing a high tense vocal tract and slower tempo a lax vocal tract. The number of acoustic correlates, although not all of them empirically tested yet, makes this a perfect candidate setting to increase agreement in future auditory evaluations, provided that perceptual assessment is aided by acoustic analysis.

Os perfis de qualidade vocal dos falantes audiogravados desta pesquisa demonstram que a taxa de elocução pode não estar relacionada com a tensão ou relaxamento do trato vocal. O falante 3 apresenta taxa de elocução rápida em grau 2 e não apresenta tensão de trato vocal. O falante 4 apresenta taxa de elocução lenta em grau 1 e hiperfunção de trato vocal em grau 2. O falante 5 apresenta taxa de elocução lenta em grau 1 e hiperfunção de trato vocal em grau 1.

Além disso, existem patologias que apresentam como característica tônus muscular excessivo, com ritmo regular e velocidade lentificada, como na disartria espástica e na hipocinética (BEHLAU ET AL, 2001). As questões de dinâmica vocal envolvem, entre outros aspectos, o contexto situacional, a personalidade e intenção do falante, e não estão, necessariamente, relacionados com característica de relaxamento ou tensão de trato vocal.

Conclusão

A versão completa do protocolo VPAS se destaca por sua capacidade de descrever detalhadamente os ajustes de trato vocal, fonatórios e de dinâmica vocal, permitindo uma caracterização abrangente do perfil de qualidade vocal do falante. Reduzir um instrumento tão minucioso como o VPAS pode, porém, restringir suas possibilidades de aplicação, especialmente no contexto forense, onde sua precisão é crucial, como na realização de laudos de comparação de locutores por peritos em fonética forense.

Ao comparar os dois protocolos, o quadro evidencia possíveis desafios enfrentados por peritos ao utilizar o VPAS simplificado. A simplificação do instrumento pode apresentar limitações na obtenção de resultados precisos, um fator preocupante em contextos forenses nos quais a exatidão e a confiabilidade dos dados são essenciais para as conclusões periciais.

Dessa forma, embora a versão simplificada do VPAS possa ter seu mérito em determinados contextos, é essencial ponderar suas limitações e considerar cuidadosamente a adequação do instrumento para diferentes propósitos, especialmente em situações que demandam análises detalhadas e rigorosas, como na fonética forense, onde a integridade dos resultados é fundamental para o desfecho de investigações.

Essas considerações destacam a importância da precisão do instrumento no contexto forense e a necessidade de cautela ao considerar versões simplificadas, ressaltando a relevância de manter a qualidade e a confiabilidade das análises em investigações judiciais.

Referências

ALÉSSIO, C.V.; MEZZOMO, C.L.; KÖRBES, D. Intervenção fonoaudiológica nos casos de pacientes classe III com indicação à cirurgia ortognática. **Arquivos em Odontologia**, v. 43, n. 3, jul/set. 2007.

ANDRADE, F.V., ANDRADE, D.V., ARAUJO, A.S., RIBEIRO, A.C.C., DECCAX, L.D.G., NEMR, K. Alterações estruturais de órgãos fonoarticulatórios e más oclusões dentárias em respiradores orais de 6 a 10 anos. **Rev. CEFAC**, São Paulo, v.7, n.3, p. 318-25, jul-set. 2005.

BEHLAU M, AZEVEDO R, PONTES P. Conceito de voz normal e classificação das disfonias. In: BEHLAU M. **Voz: o livro do especialista.** Rio de Janeiro: Revinter, 2001.

BOONE, D., MACFARLANE, S. **The voice and voice therapy**. Englewood Cliffs: Prentice Hall, 1988.

CAMARGO, Z.A; MADUREIRA, S. Avaliação vocal sob a perspectiva da fonética: investigação preliminar. **Rev. Distúrb. Comun**, São Paulo, v.20, n. 1, p.77-96, 2008.

CARRARA, E. **Análise comparativa da configuração laríngea, perceptual auditiva e espectrográfica acústica da qualidade vocal pré e pós emissão vocal em registro basal.** Monografia (Especialização) – Escola Paulista de Medicina. São Paulo, 1991.

DAENECKE, S., BIANCHINI, E.M.G., SILVA, A.P.B.V. Medidas antropométricas de comprimento de lábio superior e filtro. **Pró-Fono Revista de Atualização Científica**, v. 18, n. 3, set/dez. 2006.

DEJONCKERE PH, BRADLEY P, CLEMENTE P, CORNUT G, CREVIER-BUCHMAN L , FRIEDRICH G, VAN DE HEYNING P, REMACLE M, WOISARD V. 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**, (258), p. 77-82, 2001

FRENCH, P.; FOULKES, P.; HARRISON, P.; HUGHES, V.; SAN SEGUNDO E STEVENS, L. The vocal tract as a biometric: output measures, interrelationships, and efficacy. **Phonetic Sciences**, Glasgow, 2015.

GENARO, K.F., BERRETIN-FELIX, G., REHDER, M.I.B.C, MARQUESAN, I.Q. Avaliação miofuncional orofacial – Protocolo MGBR. **Rev. CEFAC** v.11, n. 2, p. 237-255, abr-jun. 2009.

GOLD, E.; FRENCH, P. International practices in forensic speaker comparison. **International Journal of Speech, Language and the Law**, v.18, p. 293-307, 2011.

HIRANO, M. **Clinical examination of voice**. New York: Springer-Verlag; 1981.

ISSHIKI, N., OKAMURA, M., TANABE, M., MORIMOTO, M. Aproach to the objective diagnosis of hoarseness. **Folia Phoniatric**, v. 18, p. 393-400, 1966.

MACIEL, C.T.V., BARBOSA, M.H., TOLDO, C.A., FAZA, F.C.B, CHIAPPETTA, A.L.M.L. Disfunções orofaciais nos pacientes em tratamento ortodôntico. **Rev. CEFAC**, São Paulo, v. 8, n. 4, p. 456-66, out-dez. 2006.

MACKENZIE-BECK, J. Perceptual analysis of voice quality: the place of vocal profile analysis. In: Hardcastle, W.J.; Mackenzie-Beck, J. **A figure of speech: a festschrift for John Laver.** Hillsdale: Lawrence Erlbrum Ass; 2005.

MARTINELLI R.L.C; FORNARO, E.F.; OLIVEIRA, C.J.M.; FERREIRA, L.M.D.B.; REHDER, M.I.B.C. Correlações entre alterações de fala, respiração oral, dentição e oclusão. **Rev. CEFAC**, v. 13, n.1, p. 17-26, jan-fev. 2011.

MEZZOMO, C.L.; MACHADO, P.G.; PACHECO, A.B.; GONÇALVES, B.F.T.; HOFFMANN, C.F. As implicações da classe II de Angle e da desproporção esquelética do tipo classe II no aspecto miofuncional. **Rev. CEFAC**, v. 13, n.4, p. 728-734, jul-ago. 2011.

METZGER A.L.T.; CAMPIOTTO, A.R., MUZY, P.C. Interferência do tipo de má oclusão nas medidas dos movimentos mandibulares: um estudo realizado com o apoio do exército brasileiro. **Rev. CEFAC**, São Paulo, v. 11, n. 1, jan./mar. 2009.

PEREIRA, A.C.; JORGE, T.M.; RIBEIRO JUNIOR, P.M.; BERRETIN-FELIX, G. Características das funções orais de indivíduos com má oclusão classe III e diferentes tipos faciais. **R Dental Press e Ortodontia e Ortopedia Facial**, Maringá, v. 10, n. 6, nov/dez., 2005.

PEREIRA, J.B.A.; BIANCHINI, E.M.G. Caracterização das funções estomatognáticas e disfunções temporomandibulares pré e pós cirurgia ortognática e reabilitação fonoaudiológica da deformidade dentofacial classe II esquelética. **Rev. CEFAC**, v. 13, n. 6, p.1086-1094, nov-dez. 2011.

SAN SEGUNDO, E; MOMPEAN, J.A. A simplified vocal profile analysis protocol for the assessment of voice quality and speaker similarity. **Journal of Voice**, 2017. Disponível em: http://dx.doi.org/10.1016/j.voice.2017.01.005

SÍGOLO, S.; CAMPIOTTO, A.R.; SOTELO, M.B. Posição habitual de língua e padrão de deglutição em indivíduos com oclusão classe III, pré e pós cirurgia ortognática. **Rev. CEFAC**, v. 11, n. 2, p. 256-260, abr-jun. 2009.

TAUCCI, R.A.; BIANCHINI, E.M.G. Verificação da interferência das funções temporomandibulares na articulação da fala: queixas e caracterização dos movimentos mandibulares. **Soc. Bras. Fonoaudiol**, v. 12, n. 4, p. 274-80, 2007.

VIEIRA, RC. **Identificação de falante: um estudo perceptivo de qualidade de voz.** Tese (Doutoramento) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2018.

Anexo 1 – VPAS (Adaptado para o Português Brasileiro por Camargo e Madureira, 2008)

QUALIDADE VOCAL	PRIMEIRA PASSADA		SEGUNDA PASSADA							
	Neutro	Não neutro	AJUSTE	Moderado			Extremo			
				1	2	3	4	5	6	
A. ELEMENTOS DO TRATO VOCAL										
1.Lábios			Arredondados/protraídos							
			Estirados							
			Labiodentalização							
			Extensão diminuída							
			Extensão aumentada							
2. Mandíbula			Fechada							
			Aberta							
			Protraída							
			Extensão diminuída							
			Extensão aumentada							
3.Língua ponta/lâmina			Avançada							
			Recuada							
4. Corpo de língua			Avançado							
			Recuado							
			Elevado							
			Abaixado							
			Extensão diminuída							
			Extensão aumentada							
5.Faringe			Constrição							
			Expansão							
6.Velofaringe			Escape nasal audível							
			Nasal							
			Denasal							
7. Altura de laringe			Elevada							
			Abaixada							
B. TENSÃO MUSCULAR GERAL										
8. Tensão do trato vocal			Hiperfunção							
			Hipofunção							
9. Tensão laríngea			Hiperfunção							
			Hipofunção							
C. ELEMENTOS FONATÓRIOS										
	AJUSTE		Presente		Graus de escala					
			Neutro	Não Neutro	Moderado			Extremo		
					1	2	3	4	5	6
10. Modo de fonação	Modal									
	Falsete									
	Crepitância/ *vocal fry*									
	Voz crepitante									
11. Fricção laríngea	Escape de ar									
	Voz soprosa									
12.Irregularidade laríngea	Voz áspera									

DINÂMICA VOCAL		Neutro	AJUSTE	Moderado			Extremo		
				1	2	3	4	5	6
D. ELEMENTOS PROSÓDICOS									
13.*Pitch (f0)*	Habitual		Elevado						
			Abaixado						
	Extensão		Diminuída						
			Aumentada						
	Variabilidade		Diminuída						
			Aumentada						
14.*Loudness (intensidade)*	Habitual		Aumentado						
			Diminuído						
	Extensão		Diminuída						
			Aumentada						
	Variabilidade		Diminuída						
			Aumentada						
15. Tempo									
Continuidade			Interrompida						
Taxa de elocução			Rápida						
			Lenta						
16.OUTROS ELEMENTOS									
Suporte respiratório			Adequado						
			Inadequado						

Anexo 2 – VPAS simplificado (San Segundo, Mompean, 2017)

Simplified Vocal Profile Analysis (SVPA)

A. Featural ***(tick the appropriate box)***

Major Setting Groups	Settings	Numerical Labels for One Neutral (N) and Two Non-Neutral Configurations		
		−1	0	+1
Vocal tract settings	Labial	Spreading	N	Rounding
	Mandibular	Close	N	Open
	Apical	Retracted	N	Advanced
	Dorsal	Backed and lowered	N	Fronted and raised
	Velopharyngeal	Denasal	N	Nasal
	Pharyngeal	Constricted	N	Expanded
	Laryngeal height	Lowered	N	Raised
Overall muscular tension	Vocal tract tension	Lax	N	Tense
	Laryngeal tension	Lax	N	Tense
Phonation	Voice type	Whisper/Breathy	N	Creaky/Harsh

B. Holistic

(fill with qualitative input; comments, etc)