

Comunicação Breve

Brief Communication

Rafaeli Higa Scarmagnani¹
 Adriana Cristina de Almeida Santos
 Furlan de Oliveira¹
 Ana Paula Fukushiro^{1,2}
 Manoel Henrique Salgado³
 Inge Elly Kiemle Trindade^{1,2}
 Renata Paciello Yamashita¹

Descritores

Fala
 Fissura Palatina
 Insuficiência Velofaríngea
 Reprodutibilidade dos Testes
 Rinomanometria

Keywords

Speech
 Cleft Palate
 Velopharyngeal Insufficiency
 Reproducibility of Results
 Rhinomanometry

Endereço para correspondência:

Rafaeli Higa Scarmagnani
 Rua Silvio Marchione, 3-20, Bauru (SP),
 Brasil, CEP: 17012-900.
 E-mail: rafaelihiga@usp.br

Recebido em: 28/04/2014

Aceito em: 28/07/2014

O impacto da concordância entre avaliadores no julgamento perceptivo da nasalidade da fala

Impact of inter-judge agreement on perceptual judgment of nasality

RESUMO

Objetivo: Investigar o efeito da concordância entre diferentes avaliadores no julgamento perceptivo da hipernasalidade na previsão do fechamento velofaríngeo (FVF). **Métodos:** Dois modelos de regressão logística foram desenvolvidos com o objetivo de se prever o FVF utilizando as seguintes características: FVF (adequado, marginal e inadequado), determinado pela técnica fluxo-pressão, grau da hipernasalidade (ausente, leve, moderado, grave) e a presença/ausência de emissão de ar nasal e ronco nasal, determinados perceptivamente por três fonoaudiólogos experientes. O primeiro modelo foi composto de 100 amostras de fala com índice de concordância moderado para a hipernasalidade (coeficiente kappa: 0,41) e o segundo, de 43 amostras com concordância total entre os avaliadores. O teste do χ^2 foi utilizado para comparar os modelos ($p \leq 0,05$). **Resultados:** No primeiro modelo, 65 das 100 amostras foram classificadas na categoria correta de FVF, sendo 42 adequado e 23 inadequado. O FVF marginal não foi previsto. O segundo modelo classificou 31 das 43 amostras na categoria correta, sendo 21 FVF adequado, 5 marginal e 5 inadequado. Não houve diferença ($p=0,526$) entre os dois modelos. No entanto, o segundo mostrou proporção de acerto mais alta (7%) do que o primeiro modelo, além de ter conseguido prever o FVF marginal. **Conclusão:** Esses resultados mostram a importância da concordância total entre diferentes avaliadores quando se utilizam parâmetros subjetivos de avaliação da fala, especialmente quando comparados com a avaliação instrumental, e sugerem a necessidade de estratégias para o treinamento e calibração de avaliadores no julgamento perceptivo a fim de aumentar a confiabilidade da avaliação perceptivo-auditiva da fala.

ABSTRACT

Purpose: To investigate the effect of perceptual inter-judge agreement of hypernasality on velopharyngeal (VP) closure prediction. **Methods:** Two logistic regression models were developed aiming to verify the possibility of predicting the VP closure using the following characteristics: rating of VP closure (adequate, borderline, inadequate), determined by the pressure-flow technique, degree of hypernasality (absent, mild, moderate, severe), and the presence/absence of nasal air emission and nasal rustle determined perceptually by three experienced speech language pathologists. In the first model, 100 speech samples with a moderate agreement rate of hypernasality (kappa coefficient: 0.41) were used. In the second model, 43 speech samples with a perfect agreement among judges were included. The χ^2 -test was used to compare the models ($p \leq 0.05$). **Results:** In the first model, 65 of the 100 samples were rated in the correct VP closure category, with 42 adequate and 23 inadequate. The borderline VP closure was not predicted. The second model rated 31 of the 43 samples in the correct category, with 21 adequate VP closure, 5 in the borderline VP closure, and 5 inadequate. There was no difference ($p=0.526$) between the two models. However, the second model showed a higher proportion of accuracy (7%) than the first one, and it has also predicted the borderline VP closure. **Conclusion:** These results showed the importance of high index of inter-judge agreement when using subjective parameters of speech evaluation, especially when compared to an instrumental evaluation. This suggests the need for strategies for training and calibration of judges in the perceptual judgment to improve the reliability of auditory-perceptual assessment.

Trabalho realizado no Laboratório de Fisiologia, Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo – USP – Bauru (SP), Brasil.

(1) Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo – USP – Bauru (SP), Brasil.

(2) Faculdade de Odontologia de Bauru, Universidade de São Paulo – USP – Bauru (SP) Brasil.

(3) Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP – Bauru, (SP), Brasil.

Fonte de financiamento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

A avaliação perceptivo-auditiva da fala é um importante instrumento para prever a função velofaríngea, pois as características de fala identificadas nesse tipo de avaliação fornecem pistas sobre a dimensão da falha velofaríngea⁽¹⁻³⁾. Entre os sintomas de fala relacionados diretamente à disfunção velofaríngea (DVF), está a hipernasalidade, definida como um fenômeno perceptivo e, portanto, subjetivo, fato este que interfere na confiabilidade da avaliação perceptivo-auditiva da fala e tem sido alvo de pesquisas há muitos anos, conforme revisão realizada recentemente⁽⁴⁾.

A fim de se garantir maior confiabilidade dos resultados de pesquisas envolvendo a fala de indivíduos com DVF, a literatura recomenda o uso de diferentes avaliadores para o julgamento perceptivo dos sintomas de fala^(5,6). No entanto, um alto índice de concordância pode ser difícil de ser obtido, em função de variáveis diversas.

Considerando a possibilidade de que o índice de concordância entre diferentes avaliadores no julgamento dos sintomas de fala pode influenciar a previsão do fechamento velofaríngeo, definiu-se a proposta do presente estudo.

Este trabalho buscou investigar a influência do índice de concordância entre diferentes avaliadores no julgamento perceptivo da hipernasalidade para a previsão do fechamento velofaríngeo.

MÉTODOS

Este estudo teve aprovação do Comitê de Ética em Pesquisa do Hospital de Reabilitação de Anomalias Craniofaciais da Universidade de São Paulo (ofícios n° 360/2010 e n° 254/2012 SVAPEPE-CEP) e todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

A concordância interavaliadores quanto ao julgamento da hipernasalidade foi obtida por meio do coeficiente Kappa.

Dois modelos de regressão logística foram desenvolvidos com o objetivo de verificar a possibilidade de se prever o fechamento velofaríngeo utilizando as seguintes características de fala: fechamento velofaríngeo (adequado, marginal e inadequado) determinado pela técnica fluxo-pressão, grau da hipernasalidade (ausente, leve, moderado, grave) e a presença ou ausência de emissão de ar nasal e ronco nasal aferidos, perceptivamente, por três fonoaudiólogas com experiência de 12 anos, em média, na avaliação perceptivo-auditiva da fala em pacientes com fissura labiopalatina. No primeiro modelo, foram utilizadas 100 amostras de fala com índice de concordância moderado para a hipernasalidade (coeficiente kappa: 0,41). No segundo modelo, foram utilizadas 43 das 100 amostras de fala que obtiveram concordância total entre as avaliadoras. Posteriormente, as proporções de acertos referentes a cada modelo foram comparadas por meio do teste do χ^2 a um nível de significância de 5%.

RESULTADOS

A Tabela 1 mostra o número de amostras, previsto de acordo com o modelo logístico em função da real classificação do fechamento velofaríngeo (FVF) determinado pela técnica fluxo-pressão.

Tabela 1. Distribuição das 100 amostras de fala prevendo o fechamento velofaríngeo, a partir do grau de hipernasalidade e da presença ou ausência de emissão de ar nasal e ronco nasal

FVF real	Previsão do FVF segundo o modelo		Total
	FVF	FVF	
	adequado	inadequado	
FVF adequado	42*	3	45
FVF marginal	7	13	20
FVF inadequado	12	23*	35
Total	61	39	100

*Acertos do modelo

Legenda: FVF = fechamento velofaríngeo

De acordo com esse modelo, elaborado a partir de 100 amostras que obtiveram índice de concordância moderado entre as avaliadoras no julgamento perceptivo do grau da hipernasalidade, 65 foram previstos na categoria correta. No entanto, o FVF marginal não foi previsto.

A Tabela 2 mostra o número de previsões de acordo com o segundo modelo logístico em função da real classificação do FVF, utilizando 43 amostras que obtiveram índice de concordância perfeito entre as avaliadoras no julgamento perceptivo do grau da hipernasalidade.

Tabela 2. Distribuição das 43 amostras de fala prevendo o fechamento velofaríngeo em função da hipernasalidade, emissão de ar nasal e ronco nasal

FVF real	Previsão do FVF segundo o modelo			Total
	FVF	FVF	FVF	
	adequado	marginal	inadequado	
FVF adequado	21*	1	0	22
FVF marginal	4	5*	2	11
FVF inadequado	2	3	5*	10
Total	27	9	7	43

*Acertos do modelo

Legenda: FVF = fechamento velofaríngeo

De acordo com esse modelo, 31 das 43 amostras foram classificadas corretamente. Esse modelo previu o FVF marginal.

Embora a comparação entre os dois modelos não tenha mostrado diferença ($p=0,526$), verificou-se que o segundo modelo composto pela concordância total entre as avaliadoras quanto ao grau de hipernasalidade mostrou proporção de acerto 7% maior do que o primeiro modelo, além de ter conseguido prever o FVF marginal.

DISCUSSÃO

É consenso na literatura que a avaliação perceptivo-auditiva da fala é o principal meio pelo qual o fonoaudiólogo pode identificar as alterações de fala, classificar a sua gravidade e, assim, definir a conduta e avaliar a efetividade dos tratamentos realizados⁽⁷⁻¹¹⁾. Contudo, trata-se de um método subjetivo e, portanto, sujeito a erros e influências de vários fatores. A principal influência

está relacionada ao padrão interno de cada ouvinte, ou seja, as referências próprias de cada avaliador que diferem de um para o outro. Isso levou clínicos e pesquisadores a buscar estratégias para aperfeiçoar a avaliação perceptiva, a fim de melhorar a sua confiabilidade. A principal mudança foi a utilização de gravações de amostras de fala, o que possibilitou o julgamento da fala por mais de um ouvinte. A partir daí, vários estudos mostraram a importância da presença da análise dos sintomas de fala por diferentes avaliadores e a obtenção da concordância entre eles.

Especificamente no que se refere à hipernasalidade, um alto índice de concordância entre diferentes avaliadores é difícil de ser obtido. Isso porque se trata de um fenômeno perceptivo que é influenciado por vários fatores, como, por exemplo, a experiência do ouvinte que determina o padrão interno de cada ouvinte⁽¹²⁾. Ainda que os avaliadores consultados apresentassem experiência de 12 anos, em média, na avaliação da fala na presença da fissura palatina e, ainda, formação acadêmica semelhante no que se refere ao processo de avaliação da fala, tendo recebido treinamento no mesmo centro de referência, diferenças relativas aos padrões internos de cada avaliador podem ser esperadas. Especula-se que essa tenha sido a razão para a obtenção do índice moderado de concordância quanto ao grau de hipernasalidade verificado no primeiro modelo, resultado também obtido em outros estudos⁽¹³⁻¹⁵⁾. Isso, provavelmente, foi o que levou à reduzida porcentagem de acertos mostrados nesse modelo, tornando questionável a confiabilidade da previsão do FVF. Assim, elaborou-se um segundo modelo logístico com base nas 43 amostras cujo índice de concordância entre os avaliadores no julgamento da hipernasalidade foi perfeito. Nesse caso, houve um aumento da porcentagem de previsões corretas do FVF baseado nas características perceptivas da fala. Além disso, nesse segundo modelo, o FVF marginal foi previsto, o que não havia ocorrido no modelo anterior. Esse resultado foi decorrente da concordância total dos avaliadores quanto ao grau de hipernasalidade, mostrando, portanto, a influência do índice de concordância entre diferentes avaliadores quando parâmetros subjetivos de fala estão sendo julgados, especialmente aqueles que envolvem a percepção do ouvinte.

CONCLUSÃO

O alto índice de concordância entre diferentes avaliadores quanto ao grau de hipernasalidade influenciou positivamente a previsão do FVF. Isso significa que, além do julgamento perceptivo feito por mais de um ouvinte, na avaliação perceptivo-auditiva das características de fala, é essencial o uso de estratégias que garantam altos índices de concordância entre eles, a fim de melhorar a confiabilidade dos seus resultados.

**RHS foi responsável pela ideia original do estudo, coleta de dados, análise dos dados e redação do artigo; ACASFO colaborou na coleta e tabulação dos dados; APF colaborou na coleta, análise dos dados e redação do artigo; MHS participou da análise estatística dos dados e redação do artigo; IEKT participou da redação do artigo; RPY foi responsável pelo projeto e delineamento do estudo e orientação geral das etapas de execução e elaboração do manuscrito.*

REFERÊNCIAS

1. Warren DW, Dalston RM, Mayo R. Hypernasality and velopharyngeal impairment. *Cleft Palate Craniofac J.* 1994;31(4):257-62.
2. Kummer AW, Briggs M, Lee L. The relationship between the characteristics of speech and velopharyngeal gap size. *Cleft Palate Craniofac J.* 2003;40(6):590-6.
3. Ma L, Shi B, Li Y, Zheng Q. Velopharyngeal function assessment in patients with cleft palate: perceptual speech assessment versus nasopharyngoscopy. *J Craniofac Surg.* 2013;24(4):1229-31.
4. Bressmann T, Sell D. Plus ça change: selected papers on speech research from the 1964 issue of the cleft palate journal. *Cleft Palate Craniofac J.* 2014;51(2):124-8.
5. Lohmander A, Olsson M. Methodology for perceptual assessment of speech in patients with cleft palate: a critical review of the literature. *Cleft Palate Craniofac J.* 2004;41(1):64-70.
6. Henningson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Speech parameters group. Universal parameters for reporting speech outcomes in individuals with cleft palate. *Cleft Palate Craniofac J.* 2008;45(1):2-17.
7. Dotevall H, Lohmander-Agerskov A, Ejnell H, Bake B. Perceptual evaluation of speech and velopharyngeal function in children with and without cleft palate and the relationship to nasal airflow patterns. *Cleft Palate Craniofac J.* 2002;39(4):409-24.
8. Whitehill TL, Lee ASY, Chun JC. Direct magnitude estimation and Interval scaling of hypernasality. *J Speech Lang Hear Res.* 2002;45(1):80-8.
9. Smith BE, Kuehn DP. Speech evaluation of velopharyngeal dysfunction. *J Craniofac Surg.* 2007;18(2):251-61.
10. Kummer AW. Resonance disorders and velopharyngeal dysfunction (VPD). In: Kummer AW. *Cleft palate and craniofacial anomalies: the effects on speech and resonance.* San Diego: Singular; 2008. p. 176-213.
11. Baylis AL, Munson B, Moller KT. Perceptions of audible nasal emission in speakers with cleft palate: a comparative study of listener judgments. *Cleft Palate Craniofac J.* 2011;48(4):399-411.
12. Lee A, Whitehill TL, Ciocca V. Effect of listener training on perceptual judgement of hypernasality. *Clin Linguist Phon.* 2009;23(5):319-34.
13. John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech-augmented: a validated and reliable measure for auditing cleft speech. *Cleft Palate Craniofac J.* 2006;43(3):272-88.
14. Scarmagnani RH. Correlação entre as dimensões do orifício velofaríngeo, hipernasalidade, emissão de ar nasal audível e ronco nasal em indivíduos com fissura de palato reparada [dissertação]. Bauru: Universidade de São Paulo; 2013.
15. Barbosa DA, Scarmagnani RH, Fukushima AP, Trindade IE, Yamashita RP. Surgical outcome of pharyngeal flap surgery and intravelar veloplasty on the velopharyngeal function. *CoDAS.* 2013;25(5):451-5.