

Avaliação da nasalidade de fala na fissura labiopalatina

Assessment of speech nasality in individuals with cleft palate

Edna Zakrzewski Padilha¹, Jeniffer de Cássia Rillo Dutka², Viviane Cristina de Castro Marino³, José Roberto Pereira Lauris⁴, Mariana Jales Felix da Silva⁵, Maria Inês Pegoraro-Krook²

RESUMO

Objetivo: Descrever os resultados da nasalidade de fala de indivíduos com fissura labiopalatina e comparar os achados de nasalidade estabelecidos por meio do julgamento perceptivo-auditivo realizado ao vivo com os achados estabelecidos por análise de gravações por juízes, em dois tipos de amostras de fala. **Métodos:** O estudo envolveu a análise retrospectiva dos resultados de avaliações perceptivo-auditivas da nasalidade de fala realizadas ao vivo por uma fonoaudióloga e o julgamento prospectivo, por consenso de juízas de 100 gravações de amostras de fala, obtidas durante a produção de dois conjuntos de estímulos de fala: um com consoantes de alta pressão (CAP, n=100) e outro com consoantes de baixa pressão (CBP, n=100). Os dados pertenciam a pacientes de ambos os gêneros, com idades entre 5 e 12 anos, que tiveram a fissura labiopalatina operada por um mesmo cirurgião. **Resultados:** A ausência de hipernasalidade foi constatada em 69% dos julgamentos ao vivo. Quando presente, a hipernasalidade leve foi constatada em 23% dos casos, enquanto a hipernasalidade moderada em 8%. Para os julgamentos das amostras gravadas, 50% foram identificadas com hipernasalidade durante a produção das amostras CAP e 62% durante a das amostras CBP. Diferença significativa foi encontrada entre o resultado do julgamento ao vivo e o julgamento pelas juízas nas amostras CAP. A concordância entre as modalidades de avaliação variou de 79% para as amostras CAP e 80% para as amostras CBP, sendo considerada moderada. **Conclusão:** O julgamento perceptivo ao vivo da nasalidade de fala pode detectar melhor a ausência de hipernasalidade, seguida pela hipernasalidade de grau leve, em comparação com o julgamento realizado por juízes múltiplos, a partir de amostras gravadas. Contudo, tem a desvantagem de os dados não poderem ser reproduzidos, nem quantificados, nem compartilhados por outros membros da equipe.

Descritores: Fissura palatina; Insuficiência velofaríngea; Diagnóstico; Fala; Distúrbios da fala

ABSTRACT

Purpose: To describe the results of speech nasality of individuals with cleft lip and palate, and to compare auditory-perceptual judgments of nasality between live ratings and multiple judges ratings of recorded speech, for two sets of speech stimuli. **Methods:** The study involved the retrospective analysis of the results of auditory-perceptual assessments of speech nasality performed live by a single speech-language pathologist and a prospective judgment of 100 recordings of speech samples obtained during production of two sets of speech stimuli: one with high pressure consonants (HPC, n=100) and another with low pressure consonants (LPC, n=100). The data belonged to patients, of both genders, with ages between 5 and 12 years, with cleft lip and palate operated by the same surgeon. **Results:** The absence of hypernasality was found for 69% of the patients during live assessment. When present, mild hypernasality was found for 23% and moderate for 8% of the patients. For judge ratings of recorded samples, 50% was identified as hypernasal during production of samples with high pressure consonants, and 62% for the samples with low pressure consonants. A statistically significant difference was found between the live perceptual judgments and judges' ratings of the recorded samples only for the stimuli with high pressure consonants. The agreement between the methods of assessment was 79% for HPC samples and 80% for LPC samples, within the moderate range. **Conclusion:** Live perceptual judgment of speech nasality revealed higher occurrence of absence of hypernasality followed by presence of mild hypernasality, when compared to multiple judges of recorded samples. The live clinical assessment of speech, however, has the disadvantage that the data may not be reproduced, quantified or shared by other team members.

Keywords: Cleft palate; Velopharyngeal insufficiency; Diagnosis; Speech; Speech disorders

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

(1) Programa de Pós-graduação (Mestrado) em Ciências, Departamento de Fonoaudiologia, Faculdade de Odontologia de Bauru, Universidade de São Paulo – USP – Bauru (SP), Brasil.

(2) Departamento de Fonoaudiologia, Faculdade de Odontologia de Bauru, Universidade de São Paulo – USP – Bauru (SP), Brasil; Programa de Pós-Graduação em Ciências da Reabilitação, Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo – USP – Bauru (SP), Brasil.

(3) Departamento de Fonoaudiologia, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP – Marília (SP), Brasil.

(4) Departamento de Ortodontia, Odontopediatria e Saúde Coletiva, Faculdade de Odontologia de Bauru, Universidade de São Paulo – USP – São Paulo (SP), Brasil.

(5) Programa de Pós-Graduação (Mestrado) em Ciências da Reabilitação, Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo – USP – Bauru (SP), Brasil.

Financiamento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), como bolsa de mestrado para a primeira autora.

Conflito de interesses: Não

Contribuição dos autores: EZP pesquisador principal, elaboração da pesquisa, elaboração do cronograma, levantamento da literatura, coleta e análise dos dados, redação do artigo, submissão e trâmites do artigo; JCRD elaboração da pesquisa, análise dos dados, correção da redação do artigo; VCC coleta e análise dos dados; JRPL análise dos dados; MJFS coleta e análise dos dados; MIPK orientadora, elaboração da pesquisa, elaboração do cronograma, análise dos dados, correção da redação do artigo, aprovação da versão final.

Endereço para correspondência: Maria Inês Pegoraro-Krook, Departamento de Fonoaudiologia, Faculdade de Odontologia de Bauru, Universidade de São Paulo. Al. Octávio Pinheiro Brisolla, 9/75, Bauru (SP), Brasil, CEP: 17012-901. E-mail: mikrook@usp.br

Recebido em: 8/7/2014; **Aceito em:** 11/2/2015

INTRODUÇÃO

A disfunção velofaríngea é considerada um dos principais fatores etiológicos que influenciam as habilidades de produção de fala de indivíduos com fissura de palato⁽¹⁾. Os resultados de fala após a palatoplastia primária são usados como indicadores do resultado da cirurgia e podem refletir na eficácia dos protocolos de tratamento usados pelos serviços que gerenciam essa anomalia⁽²⁻⁵⁾.

O instrumento de rotina usado para a avaliação dos resultados da correção cirúrgica da fissura labiopalatina é o julgamento perceptivo-auditivo da fala, realizado ao vivo por um fonoaudiólogo⁽⁶⁾. Um estudo envolvendo fonoaudiólogos de equipes craniofaciais norte-americanas constatou que 99% das equipes utilizavam a avaliação perceptivo-auditiva como padrão ouro para estabelecer os resultados de fala durante avaliação da função velofaríngea⁽⁵⁾. Considerando-se que os sintomas da disfunção velofaríngea são perceptíveis, é possível entender porque o julgamento perceptivo-auditivo é considerado o padrão ouro para a avaliação da fala^(6,7) e porque este julgamento deve ser realizado por ouvintes treinados^(8,9).

A identificação da ocorrência da hipernasalidade por equipes craniofaciais é realizada, geralmente, por meio de julgamentos perceptivo-auditivos, envolvendo o uso de escalas binárias (normal *versus* anormal), ou usando-se categorias descritivas envolvendo escalas com intervalos iguais, como, por exemplo, uma escala de 4 pontos, onde 1=normal, 2=leve, 3=moderada e 4=grave. Observa-se, que a estimativa da magnitude direta e as comparações pareadas (com ou sem amostras de referência) também têm sido usadas para identificar a hipernasalidade⁽⁶⁾. A literatura reporta, com maior frequência, o uso de categorias descritivas e a escala de intervalos iguais, assumindo-se, nesta última, que a distância entre duas posições na escala é a mesma, apesar de ser possível a variação do grau da nasalidade com escalas com intervalos de 3, 4, 5 ou mais graus^(6,10).

Além das escalas com diferentes intervalos de pontuação, existem outras diferenças metodológicas encontradas na avaliação da fala de pacientes com fissura labiopalatina, em equipes de diferentes partes do mundo. Na classificação da hipernasalidade, por exemplo, são utilizados diferentes estímulos para obtenção das amostras de fala, variando entre palavras isoladas, sentenças ou fala espontânea, amostras balanceadas de acordo com o tipo de vogal (vogais altas *versus* demais vogais) e, ainda, amostras balanceadas de acordo com o tipo de consoante, incluindo consoantes orais de alta pressão (como plosivas e fricativas), consoantes orais de baixa pressão (como líquidas), ou, ainda, consoantes orais combinadas a consoantes nasais. Podem ser encontradas, também, diferenças no método para elicitar a amostra de fala (como leitura, nomeação ou repetição) e, ainda, diferenças para documentação dos resultados de fala (como a avaliação ao vivo e o uso de gravações de áudio ou vídeo, por exemplo)^(10,11).

A gravação do sinal de áudio é o sistema de documentação de fala mais comum entre equipes craniofaciais, particularmente devido à facilidade de captura, arquivo, edição e análise das gravações por juízes múltiplos. Além disso, o uso de gravações de amostras de fala é essencial para a obtenção de medidas de confiabilidade intra e interjuízes, além de permitir a comprovação dos achados clínicos obtidos durante avaliação ao vivo^(10,12,13). Pesquisadores da Escandinávia e do Reino Unido conduziram um estudo multicêntrico entre equipes craniofaciais (Projeto *Scandleft*, 1997), tendo como um dos objetivos padronizar a captura e a análise de amostras de fala obtidas, para documentar os resultados da correção cirúrgica da fissura labiopalatina. Os pesquisadores ressaltaram que as gravações de áudio e/ou de vídeo deveriam ser parte do protocolo de documentação de todos os pacientes e que o equipamento de gravação deve ser de boa qualidade para permitir a avaliação de diferentes variáveis da qualidade de fala⁽¹³⁾.

Vários autores têm apontado a importância de se estabelecer um protocolo clínico criterioso para a avaliação da nasalidade de fala, sugerindo a padronização dos julgamentos^(6,10,14). Cuidados com a metodologia são essenciais para possibilitar comparações de resultados de fala e assegurar que não haja perda ou comprometimento dos dados decorrentes da qualidade das amostras de fala gravadas⁽¹⁴⁾.

Para o gerenciamento da fissura labiopalatina, de uma forma geral, os procedimentos para o julgamento da nasalidade de fala precisam focar, primeiramente, a identificação da ausência ou presença de hipernasalidade, visto que esta é considerada um dos principais indicadores do resultado cirúrgico e é, também, um dos sintomas primários da disfunção velofaríngea. Para avaliar adequadamente a nasalidade de fala, portanto, é necessário fazer uso de procedimentos que sejam eficientes, precisos e confiáveis e que também possam colaborar para a consistência dos achados.

O presente estudo teve como objetivos descrever e comparar os resultados da nasalidade de fala de indivíduos com fissura labiopalatina, obtidos por meio do julgamento perceptivo-auditivo realizado ao vivo e por meio de gravações de amostras de fala.

MÉTODOS

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa em Seres Humanos do Hospital de Reabilitação de Anomalias Craniofaciais da Universidade São Paulo (HRAC-USP) (346/2012).

Amostras de fala

Foram identificados, inicialmente, 100 pacientes com fissura labiopalatina unilateral, sem outras anomalias associadas, operados, consecutivamente, por um mesmo cirurgião, durante um período de cinco anos. As palatoplastias primárias do grupo estudado foram conduzidas entre 9 e 18 meses de idade (média

= 12 meses), na mesma instituição. Foram, então, selecionadas 100 gravações e os respectivos registros de avaliações perceptivo-auditivas realizadas numa mesma visita do paciente à instituição e obtidas antes de outros procedimentos cirúrgicos, sendo 41 gravações de pacientes do gênero feminino e 59 do masculino, idades entre 5 e 12 anos, todas obtidas no período entre agosto de 2006 e maio de 2010.

Avaliação perceptivo-auditiva da nasalidade ao vivo

Neste estudo, os dados da avaliação perceptivo-auditiva ao vivo foram obtidos por meio de análise retrospectiva das informações registradas nos prontuários dos pacientes. Três fonoaudiólogas, com mais de cinco anos de experiência diária na avaliação e no tratamento de pacientes com fissura labiopalatina, realizaram as avaliações ao vivo entre 2006 e 2010 e registraram a ocorrência e o grau de hipernasalidade nos prontuários dos pacientes, como parte do protocolo de avaliação pós palatoplastia primária conduzido na instituição. Utilizaram, para este fim, uma escala de 4 pontos, na qual o zero sugere ausência de hipernasalidade e as pontuações 1, 2 e 3 sugerem a presença de hipernasalidade (leve, moderada e grave, respectivamente). A avaliação ao vivo foi estabelecida a partir de uma amostra de fala não padronizada, envolvendo uma breve conversa e a repetição de palavras e/ou frases variadas (que poderiam ou não apresentar recorrência de sons).

Avaliação perceptivo-auditiva da nasalidade por juízas

Uma análise prospectiva da nasalidade de fala foi realizada por outras três fonoaudiólogas, em consenso, a seguir denominadas juízas, experientes na avaliação de fala de pacientes com fissura labiopalatina. Foram utilizadas gravações de fala dos 100 pacientes estudados, existentes em um banco de dados da instituição onde foi conduzido o estudo. Do banco de dados foram selecionadas as gravações das amostras de fala de dois conjuntos de frases com sons orais: um com frases contendo consoantes de alta pressão, (amostras CAP) e outro contendo frases somente com consoantes de baixa pressão (amostras CBP). As amostras CAP incluíram as seguintes frases: *Papai caiu da escada, Fábio pegou o gelo, O palhaço chutou a bola, Teresa fez pastel e A árvore dá frutos e flores*. As amostras CBP incluíram as seguintes frases: *O louro ia olhar a lua, Laura lia ao luar, A leoa é leal, Lili era loira e Lulu olha a arara*. As amostras foram capturadas por um microfone de cabeça, condensado/unidirecional, modelo AKG C420®, posicionado a, aproximadamente, 5 cm da lateral da comissura labial. Foram gravadas diretamente no computador IBM-PC Intel Pentium® 4 2.8 GHz, equipado com placa de som *Sound Blaster Audigy 2*, programa Sony® Sound Forge, versão 7.0 (2003), com taxa de amostragem de 44100 Hz, em monocal, 16 Bits, em arquivos com extensão do tipo *wave*. Todas as gravações foram realizadas em ambiente silencioso e acusticamente tratado, conforme rotina de documentação de resultados de fala da instituição.

As gravações foram editadas utilizando-se o programa Sony® Sound Forge Pro-10 (Sony® Media Software, 2009). Dois arquivos de amostras de áudio, formato *wave*, foram preparados: um para amostras CAP e outro para amostras CBP, criando-se a pasta denominada “amostras para julgamento”. Para cada amostra, foi atribuído um número, além de uma identificação, indicando se pertencia a pacientes do gênero feminino ou masculino. Informações sobre o gênero dos participantes foram fornecidas para possibilitar o acesso às amostras de referência, que foram apresentadas de acordo com o gênero do falante.

As amostras de referência foram estabelecidas para calibrar as juízas na identificação dos quatro intervalos da escala que foi usada neste estudo. Estas referências foram estabelecidas após terem sido julgadas como representativas de cada grau da escala com 100% de concordância entre as juízas.

Para que as juízas tivessem acesso às amostras de referência, durante as análises das gravações, foram criados quatro arquivos representativos de cada um dos quatro intervalos da escala (ausência de hipernasalidade, hipernasalidade leve, moderada e grave). Em cada um, foram inseridas amostras representativas dos dois gêneros (feminino e masculino) e dos dois tipos de amostras (CAP e CBP). Ou seja, no arquivo com amostras representativas da ausência de hipernasalidade, foram criadas quatro pastas: uma contendo amostras de referência com as frases com consoantes de alta pressão para a voz feminina (CAP-feminino); uma contendo amostras de referência com as frases com consoantes de alta pressão para a voz masculina (CAP-masculino); uma contendo amostras de referência com as frases com consoantes de baixa pressão para a voz feminina (CBP-feminino) e outra contendo amostras de referência com as frases com consoantes de baixa pressão para a voz masculina (CBP-masculino).

Como o mesmo procedimento foi realizado para as amostras representativas dos outros três intervalos da escala (hipernasalidade leve, moderado e grave), 16 pastas com amostras de referência foram criadas, todas usando amostras que não pertenciam aos pacientes incluídos no estudo (dois gêneros + dois tipos de amostras + quatro intervalos da escala = 16). Estas amostras foram usadas para calibrar as juízas quanto ao uso da escala de 4 pontos e, também, para que ficassem acessíveis para consultas durante os julgamentos das gravações de interesse.

As amostras para o julgamento e as amostras de referência foram salvas em um *compact disc* (CD). Antes do início do julgamento, as juízas passaram por breve treinamento, durante o qual tiveram oportunidade de ouvir as amostras de referência, e usar a escala de 4 pontos para comparar os diferentes intervalos da escala, nas vozes masculina e feminina. O julgamento perceptivo-auditivo foi realizado em uma sala silenciosa, onde as três juízas ouviram as amostras para julgamento, usando o programa Windows Media Player (Microsoft Windows®), cada uma utilizando um fone de ouvido AKG® K414P, todas conectadas ao computador. Deste modo, as três juízas ouviram

as amostras ao mesmo tempo e indicaram o resultado do julgamento em duas fichas de anotações, uma para as amostras CAP e outra para as amostras CBP. Era permitido ouvir cada gravação quantas vezes fossem necessárias para realizar a classificação. As juízas foram instruídas a julgar apenas o aspecto hipernasalidade em todas as amostras, tendo quatro opções de escolha: ausência de hipernasalidade, hipernasalidade leve, hipernasalidade moderada e hipernasalidade grave. Em caso de dúvida no julgamento, ou de desacordo, as amostras de referência eram ouvidas novamente, para ajudar na diferenciação entre um e outro grau da escala. Era permitido que as juízas discutissem o resultado, até que fosse estabelecido um julgamento único (100% de consenso), para cada amostra CAP e CBP. A classificação da nasalidade foi realizada em um único dia, com a realização de uma sessão no período da manhã, para as amostras CAP, e no período da tarde, para as amostras CBP. Cada sessão de classificação durou em torno de três horas, com um intervalo de dez minutos.

A análise descritiva dos dados envolveu medidas de porcentagem de ocorrência da hipernasalidade de fala e a porcentagem para cada grau da escala de 4 pontos. Para a análise inferencial dos achados, as medidas obtidas usando a escala de 4 pontos foram transformadas em escala binária, sendo indicadas apenas a ausência ou a presença da hipernasalidade. A comparação para verificar a significância das diferenças dos resultados obtidos nas diferentes modalidades de julgamento, no que se refere à proporção de dados com ausência e presença da hipernasalidade, foi realizada por meio do teste de McNemar, com nível de significância de $p < 0,05$. A concordância das modalidades de avaliação foi expressa com a porcentagem de concordância e também com o coeficiente Kappa⁽¹⁵⁾.

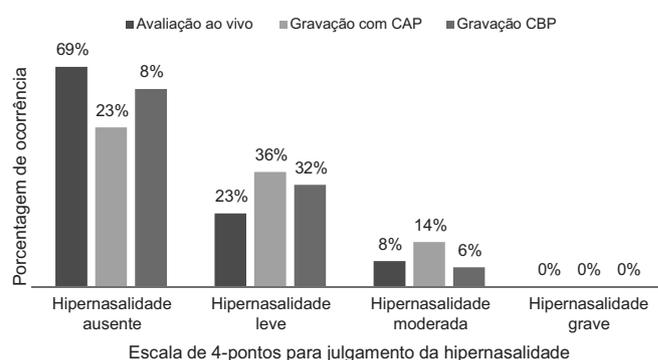
RESULTADOS

No levantamento retrospectivo dos dados do prontuário, foram encontrados 99 registros de julgamento perceptivo-auditivo ao vivo ($n=99$). Constatou-se que 68 (69%) indicavam ausência de hipernasalidade e 31 (31%) apresentavam hipernasalidade, sendo que destes 31 com hipernasalidade, 23 (23%) indicavam grau leve e 8 (8%), grau moderado. O grau de hipernasalidade grave não foi observado para esse grupo de pacientes, ou seja, se o resultado da palatoplastia primária na instituição onde o estudo foi conduzido for baseado apenas na identificação da presença ou ausência de hipernasalidade estabelecida por um único fonoaudiólogo, durante julgamento perceptivo-auditivo ao vivo, 31% dos pacientes deste estudo apresentaram um sintoma indicativo de disfunção velofaríngea.

A análise perceptivo-auditiva das gravações pelas juízas foi estabelecida inicialmente com 100% de consenso para 80% das amostras julgadas. Isto é, para 20% das amostras, as três juízas não concordaram 100% quanto à ausência ou presença de hipernasalidade (e o grau) e tiveram que rever os julgamentos depois de acessar as amostras de referência para recalibrar os

intervalos da escala. Após discussão e revisão das amostras, foi estabelecido consenso para 100% das gravações entre as juízas.

Ao considerarmos os 200 julgamentos das amostras gravadas pelas juízas (100 julgamentos para as amostras CAP e 100 julgamentos para as amostras CBP), os resultados revelaram que 112 (56%) delas foram julgadas com ausência de hipernasalidade, enquanto 88 (44%) foram julgadas com presença de hipernasalidade. Especificamente com relação às amostras CAP, foram encontradas 50 (50%) com ausência de hipernasalidade, 36 (36%) com hipernasalidade leve, 14 (14%) com hipernasalidade moderada e nenhuma (0%) com hipernasalidade grave. Nas amostras CBP foram encontradas 62 (62%) com ausência de hipernasalidade, 32 (32%) com hipernasalidade leve, 6 (6%) com hipernasalidade moderada e nenhuma (0%) com hipernasalidade grave (Figura 1).



Legenda: CAP = consoantes de alta pressão; CBP = consoantes de baixa pressão

Figura 1. Ocorrência de hipernasalidade em relação aos julgamentos perceptivo-auditivos da nasalidade de fala, nas amostras de alta e de baixa pressão

Ao comparar os resultados do julgamento perceptivo ao vivo obtido dos prontuários dos pacientes com o julgamento perceptivo das gravações por juízas, constatou-se que houve diferença entre o resultado do julgamento perceptivo ao vivo (1%) e o julgamento perceptivo por juízas nas amostras CAP (50%). As juízas identificaram hipernasalidade na produção das amostras CAP de 19 indivíduos que foram julgados sem hipernasalidade de fala no julgamento ao vivo ($p < 0,001$, teste Mc Nemar).

A porcentagem de concordância entre o julgamento perceptivo ao vivo e o julgamento perceptivo por juízas para as amostras foi de 79% para as amostras CAP e de 80% para as amostras CBP, com coeficiente Kappa (K) indicando concordância moderada ($K=0,57$ e $0,55$).

DISCUSSÃO

No ano de 2002, a Organização Mundial de Saúde⁽¹⁶⁾ lançou um relatório sobre pesquisa em anomalias craniofaciais, enfatizando a necessidade de se estabelecer métodos e determinar medidas de resultados válidas. Recomendou que a variável nasalidade fosse avaliada quanto a sua ocorrência

(ausência/presença), assim como fosse estabelecido o grau de hipernasalidade.

Realizar uma avaliação de nasalidade, especialmente da hipernasalidade, não é uma tarefa fácil, haja vista que inúmeras variáveis podem influenciar os julgamentos. Dentre elas, destacam-se as características individuais da casuística envolvida, o estímulo de fala usado na captura das amostras a serem julgadas, os procedimentos para a avaliação perceptiva, a escolha dos métodos estatísticos utilizados para comparação, dentre outras^(5,11,12,17,18). No presente estudo, tomou-se o cuidado de selecionar as amostras de fala de indivíduos com o mesmo tipo de fissura e operados pelo mesmo cirurgião, minimizando-se as variáveis externas que pudessem ter impacto nos resultados de nasalidade após a palatoplastia primária para correção da fissura labiopalatina.

A dificuldade em se obter reprodução nos julgamentos de fala tem sido evidenciada na literatura, sendo que a avaliação perceptivo-auditiva da nasalidade de fala, particularmente, tem se mostrado uma tarefa desafiadora^(6,10,17,19,20). Neste estudo, houve diferença entre a avaliação perceptivo-auditiva ao vivo (resgatada dos prontuários dos pacientes) e a avaliação de gravações de amostras por juízes, especificamente para as amostras CAP. Ao analisar as amostras CAP e CBP, observou-se que o conjunto de amostras CBP possuía maior ocorrência de vogais altas. Pesquisadores relataram que diferentes vogais produzem padrões de produção significativamente diferentes na nasalidade, com a altura da língua influenciando as características acústicas de cada amostra e, portanto, nos graus de percepção da nasalidade⁽²¹⁾. Particularmente para indivíduos com hipernasalidade e fissura de palato operada, é enfatizado por alguns autores^(21,22) que os ouvintes percebem as vogais altas como mais nasalizadas do que vogais baixas tanto na forma isolada como em sentenças. Uma tendência para os falantes deste estudo perceberem o conjunto de amostras CAP (com maior porcentagem de vogais altas) como mais nasais do que o conjunto de amostras CBP (com menor porcentagem de vogais altas), entretanto era esperada. A literatura^(21,22) também tem enfatizado que os ouvintes percebem vogais baixas como mais nasais do que vogais altas em falantes normais. Um estudo encontrou que vogais baixas são produzidas por falantes normais com menor força no fechamento velofaríngeo do que em vogais altas⁽²³⁾. Pode-se inferir que vogais altas, tanto quanto consoantes de alta pressão, requerem um fechamento velofaríngeo mais preciso do que vogais baixas e consoantes de baixa pressão, e o achado de maior ocorrência de ausência de hipernasalidade para a amostra CBP pode ser decorrente da diferença, em conteúdo fonético, entre estímulos CAP e CBP. Futuros estudos com maior controle do contexto-fonético podem contribuir para melhor esclarecer estes aspectos.

Outro fator que pode ter interferido no julgamento perceptivo das gravações por juízas do presente estudo foi o comprimento e o contexto fonético do estímulo de fala julgado, uma vez que cada conjunto a ser analisado era constituído por cinco

frases, além da variedade do contexto fonético das amostras CAP. Alguns autores⁽²⁴⁾ sumarizaram os achados da literatura sobre o efeito do comprimento do estímulo na confiabilidade do ouvinte. Os dados obtidos neste estudo mostraram que a confiabilidade do ouvinte em julgar nasalidade foi maior para sentenças do que para palavras isoladas, e maior para palavras isoladas do que para vogais isoladas. Entretanto, os resultados não esclarecem se estímulos mais curtos ou estímulos com menor variedade de sons podem favorecer os julgamentos perceptivo-auditivos de amostras gravadas. No caso do contexto fonético muito variado, em particular, existe o risco dos ouvintes prestarem atenção a partes isoladas do estímulo de fala, ou ainda, a aspectos não relacionados à nasalidade de fala, como distorção provocada pelo escape de ar nasal audível, ronco nasal, uso de articulação compensatória ou mesmo disфония, levando a maior variação nas medidas de confiabilidade dos juízes.

Há indicação na literatura que, tanto para calibrar os juízes, quanto para melhorar a reprodução dos julgamentos, estudos envolvendo análise perceptiva têm utilizado estímulos “âncora”⁽²⁵⁾. No presente estudo, estes estímulos foram chamados de amostras de referência e foram usados tanto no treinamento quanto na calibração das juízas e durante os julgamentos, quando houve falta de consenso. As amostras de referência foram usadas durante o julgamento perceptivo, especialmente quando as amostras apresentavam hipernasalidade de grau leve. O limiar entre a normalidade e uma alteração de fala é de grande importância clínica, pois após identificar a hipernasalidade, particularmente, o clínico geralmente encaminha o paciente para outras avaliações e/ou monitoramentos e/ou sugestões de intervenção (cirurgia secundária, por exemplo). Usualmente, avaliações instrumentais são obtidas com exames mais invasivos, que permitem a visualização das estruturas do mecanismo velofaríngeo, como a nasoendoscopia e a videofluoroscopia. Instrumentos que oferecem possibilidade de análise do sinal acústico (como nasometria) ou aerodinâmico da fala (como técnica de fluxo-pressão) também podem ser usados, uma vez que permitem ao clínico inferir sobre a adequação ou inadequação do funcionamento do mecanismo velofaríngeo, confirmando os achados perceptivos^(5,6,8,26). Os dados deste estudo mostram a importância de combinar o uso de diferentes modalidades de avaliação da fala e da função velofaríngea na prática clínica, tomando-se o cuidado para a definição da conduta mais adequada para cada caso, especialmente quando as alterações detectadas sugerem uma alteração leve de fala e disfunção velofaríngea marginal. Os achados deste estudo, portanto, concordam com trabalhos na literatura⁽¹⁴⁾ que apontam para a dificuldade do ouvido humano em trabalhar com alterações muito próximas do limiar de normalidade.

Nos resultados do presente estudo, mais da metade da amostra foi julgada com ausência de hipernasalidade (69% pelo perceptivo ao vivo, 50% pelo perceptivo por juízas das amostras CAP e 62% pelo perceptivo por juízas das amostras

CBP), seguida por uma representativa quantidade de amostras com hipernasalidade leve (23% pelo perceptivo ao vivo, 36% pelo perceptivo por juízas das amostras CAP e 32% por juízas das amostras CBP). Estes dados revelam que quem fez o julgamento teria que optar exatamente entre a ausência ou a presença de hipernasalidade e como o julgamento ao vivo não foi realizado simultaneamente à realização da gravação, os pacientes com nasalidade limítrofe estavam passíveis de apresentar flutuações na nasalidade de fala. O nível de correlação encontrado, neste estudo, entre as duas modalidades de avaliação perceptivo-auditiva (ao vivo *versus* gravação) pode ter sido influenciado pela dificuldade do ouvido humano em usar escalas com intervalos, especialmente com as amostras no limiar entre normalidade e alteração.

O resultado da análise Kappa, em especial, é sempre afetado quando não existe uma amostra distribuída igualmente em todos os graus da escala, o que foi verdade neste estudo e é a realidade nesta prática clínica. A maioria das gravações foi representativa de fala sem hipernasalidade, seguidas do grupo com hipernasalidade leve, com apenas alguns casos nas categorias moderada e grave. Assim sendo, mesmo tendo uma alta porcentagem de concordância entre a avaliação ao vivo e a análise das gravações por juízas, em ambos os estímulos de fala (CAP=79% e CBP=80%), o teste Kappa revelou concordância de 0,57 e 0,55, respectivamente. Interessante observar que ao comparar o julgamento ao vivo com o julgamento por juízas na amostra CAP, houve diferença significativa na ocorrência de hipernasalidade. A diferença entre os julgamentos ao vivo e os julgamentos das gravações com amostras CBP, entretanto, não foi significativa mesmo que o teste Kappa tenha sido muito similar para ambas as amostras. Este achado também foi reportado em um estudo na literatura⁽²⁷⁾.

Uma variável importante, presente entre as modalidades de avaliação perceptivo-auditiva, são os vieses daqueles que realizam o julgamento. No presente estudo, no julgamento perceptivo-auditivo ao vivo, a avaliação clínica foi realizada por uma dentre três fonoaudiólogas pertencentes à equipe craniofacial e, embora experientes na avaliação de fala de pacientes com fissura labiopalatina, poderiam apresentar padrões de julgamentos diferentes. Quando ouvintes classificam características da fala ou voz com algum critério de qualidade, comparam o estímulo apresentado a um padrão interno (uma “escala pessoal”). Estes padrões internos desenvolvem-se, são mantidos na memória e acabam por ser diferentes de ouvinte para ouvinte. Além disso, são inerentemente instáveis e podem ser influenciados por fatores internos, como lapsos de memória e atenção e por variáveis externas, tais como contexto acústico e tarefas de escuta^(25,28). Alguns estudos recomendam que a análise de dados de fala em indivíduos com fissura labiopalatina seja feita por fonoaudiólogos independentes⁽¹²⁾ e outros recomendam que o julgamento perceptivo da hipernasalidade tem mais credibilidade e confiabilidade quando é feito por fonoaudiólogos em consenso, utilizando amostras de fala gravadas em áudio e/ou vídeo^(11,13).

A medida perceptiva tem a vantagem de ser o padrão ouro, por não requerer instrumentos especiais e por não ser cara, porém, como desvantagem, não avalia diretamente a função e as estruturas do mecanismo velofaríngeo, é subjetiva e suscetível a pobre reprodutibilidade entre os examinadores⁽²⁶⁾. A gravação em áudio de amostras de fala é o sistema de documentação da fala mais usado, pois tem a vantagem de ser facilmente recuperada, editada e apresentada para julgamento perceptivo-auditivo por juízes, o que permite a obtenção de medidas de confiabilidade intra e interavaliadores, favorecendo, ainda, a confirmação dos achados clínicos obtidos durante avaliação ao vivo. Porém, há a necessidade de protocolos padronizados para coleta das amostras de fala e equipamentos de gravação de boa qualidade^(10,12,13). Ao se utilizar gravações, ao invés de julgamentos ao vivo, as informações da fala são obtidas sem a influência das expectativas do clínico, das informações visuais, ou de outras informações individuais do paciente. O julgamento da nasalidade na situação ao vivo considera todos os dados necessários para que o clínico faça um julgamento pleno. Entretanto, como não permite uma avaliação cega da fala, é mais sujeito aos vieses do avaliador e não oferece a possibilidade de medidas de confiabilidade^(12,14).

A concordância moderada entre as duas modalidades de avaliação talvez tenha sido também resultado do fato de as amostras gravadas serem mais limitadas quanto ao conjunto de informações disponíveis para o ouvinte, no momento de uma avaliação ao vivo, ou seja, embora as gravações pudessem ser ouvidas quantas vezes fossem necessárias para o julgamento, não havia a possibilidade de se obter informações visuais ou individuais do paciente.

O fonoaudiólogo que avaliou o paciente ao vivo teve acesso às informações do prontuário, às avaliações preexistentes e a outras informações, como expressão facial e corporal. Estudos que ofereçam maior controle da modalidade de avaliação ao vivo, ou o mesmo acesso ao prontuário e filmagem para análise das gravações de áudio, são necessários para melhor esclarecer esses aspectos.

Outra importante limitação a ser considerada em futuros estudos, é o momento em que foram conduzidas as avaliações da nasalidade, sendo que, para a análise pelas juízas, as três fonoaudiólogas avaliaram todas as gravações em um único dia, enquanto o julgamento ao vivo foi conduzido durante a prática clínica e ao longo de cinco anos (2006-2010).

Ao considerarmos os resultados de nasalidade obtidos nas duas modalidades de avaliação, de uma forma geral, observa-se grande similaridade entre os achados, mesmo quando levamos em conta as diferentes variáveis discutidas acima. Neste estudo, as juízas puderam comparar a amostra a ser julgada com as amostras de referência, para a classificação da hipernasalidade durante a análise das gravações de áudio, e, portanto, tiveram suas respostas “ancoradas”, confirmando achados de outros autores⁽²⁵⁾. Gerrat et al.⁽²⁵⁾, ao compararem o julgamento perceptivo de vozes normais e levemente ásperas

com uma escala de intervalos iguais de 5 pontos e outra, com estímulos de referência “âncora” para cada ponto da escala, relataram que as avaliações que utilizaram a escala “ancorada” foram mais confiáveis do que as “não ancoradas”.

Finalmente, para aumentar o entendimento sobre hipernasalidade e sua avaliação, diferentes aspectos devem ser ainda mais investigados, como o efeito do treinamento auditivo, o uso de amostras de referência como âncoras, o tipo de escala usada e, ainda, a diferenciação entre os vários intervalos da escala. São necessários, também, estudos sobre a percepção da nasalidade de fala na presença de alterações associadas, como uso de articulação compensatória, escape de ar nasal audível com e sem turbulência (ronco nasal), fraca pressão intraoral e distorções decorrentes de alterações dento-oclusais.

CONCLUSÃO

O julgamento perceptivo realizado ao vivo pode detectar melhor a ausência de hipernasalidade de fala, assim como a hipernasalidade de grau leve, em comparação ao julgamento realizado por juízes múltiplos, a partir de amostras gravadas. Contudo, tem a desvantagem de os dados não poderem ser reproduzidos, nem quantificados, nem compartilhados por outros membros da equipe e nem comparados com dados de outros centros. Futuros estudos com maior controle do contexto fonético das amostras de fala devem ser realizados para melhorar a confiabilidade do julgamento perceptivo-auditivo da nasalidade de fala.

REFERÊNCIAS

1. D'Antonio L, Scherer N. Communication disorders associated with cleft palate. In: Losee JE, Kirschner RE, organizers. *Comprehensive cleft care*. New York: Mc Graw Hill Professional; 2008. p.114-35.
2. Bae YC, Kim JH, Lee J, Hwang SM, Kim SS. Comparative study of the extent of palatal lengthening by different methods. *Ann Plast Surg*. 2002;48(4):359-62. <http://dx.doi.org/10.1097/00000637-200204000-00004>
3. Farzaneh F, Becker M, Peterson AM, Svensson H. Speech results in adult Swedish patients born with unilateral complete cleft lip and palate. *Scand J Plast Reconstr Surg Hand Surg*. 2008;42(1):7-13. <http://dx.doi.org/10.1080/02844310701694522>
4. Koh KS, Kang BS, Seo DW. Speech evaluation after repair of unilateral complete cleft palate using modified 2-flap palatoplasty. *J Craniofac Surg*. 2009;20(1):111-4. <http://dx.doi.org/10.1097/SCS.0b013e318195ab0a>
5. Kummer AW, Clark SL, Redle EE, Thomsen LL, Billmire DA. Current practice in assessing and reporting speech outcomes of cleft palate and velopharyngeal surgery: a survey of cleft palate/craniofacial professionals. *Cleft Palate Craniofac J*. 2012;49(2):146-52. <http://dx.doi.org/10.1597/10-285>
6. Kuehn D, Moller K. Speech and language issues in the cleft palate population: the state of the art. *Cleft Palate Craniofac J*. 2000;37(4):348-83. [http://dx.doi.org/10.1597/1545-1569\(2000\)037<0348:SALIIT>2.3.CO;2](http://dx.doi.org/10.1597/1545-1569(2000)037<0348:SALIIT>2.3.CO;2)
7. Brunnegård K, Lohmander A. A cross-sectional study of speech in 10-year-old children with cleft palate: results and issues of rater reliability. *Cleft Palate Craniofac J*. 2007;44(1):33-44. <http://dx.doi.org/10.1597/05-164>
8. Smith B, Guyette TW. Evaluation of cleft palate speech. *Clin Plast Surg*. 2004;31(2):251-60. [http://dx.doi.org/10.1016/S0094-1298\(03\)00123-8](http://dx.doi.org/10.1016/S0094-1298(03)00123-8)
9. Kummer AW. Perceptual assessment of resonance and velopharyngeal function. *Semin Speech Lang*. 2011;32(2):159-67. <http://dx.doi.org/10.1055/s-0031-1277718>
10. 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. <http://dx.doi.org/10.1597/02-136>
11. Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Universal parameters for reporting speech outcomes in individuals with cleft palate. *Cleft Palate Craniofac J*. 2008;45(1):1-17. <http://dx.doi.org/10.1597/06-086.1>
12. Sell D. Issues in perceptual speech analysis in cleft palate and related disorders: a review. *Int J Lang Commun Disord*. 2005;40(2):103-21. <http://dx.doi.org/10.1080/13682820400016522>
13. Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutter B. Methodology for speech assessment in the Scandcleft project-an international randomized clinical trial on palatal surgery: experiences from a pilot study. *Cleft Palate Craniofac J*. 2009;46(4):347-62. <http://dx.doi.org/10.1597/08-039.1>
14. Sweeney T, Sell D. Relationship between perceptual ratings of nasality and nasometry in children/adolescents with cleft palate and/or velopharyngeal dysfunction. *Int J Lang Commun Disord*. 2008;43(3):265-82. <http://dx.doi.org/10.1080/13682820701438177>
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74. <http://dx.doi.org/10.2307/2529310>
16. World Health Organization. *Global strategies to reduce the health-care burden of craniofacial anomalies: report of WHO meetings on International Collaborative Research on Craniofacial Anomalies; Geneva; 2000 May 24-26*. Geneva: World Health Organization; 2002.
17. Wyatt R, Sell D, Russell J, Harding A, Harland K, Albery E. Cleft palate speech dissected: a review of current knowledge and analysis. *Br J Plast Surg*. 1996; 49(3):143-9. [http://dx.doi.org/10.1016/S0007-1226\(96\)90216-7](http://dx.doi.org/10.1016/S0007-1226(96)90216-7)
18. Kent R. Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *Am J Speech Lang Pathol*. 1996;5(3):7-23. <http://dx.doi.org/10.1044/1058-0360.0503.07>
19. Keuning KH, Wieneke GH, Dejonckere PH. The intrajudge reliability of the perceptual rating of cleft palate speech before and after pharyngeal flap surgery: the effect of judges and speech samples. *Cleft Palate Craniofac J*. 1999;36(4):328-33. [http://dx.doi.org/10.1597/1545-1569\(1999\)036<0328:TIOTP>2.3.CO;2](http://dx.doi.org/10.1597/1545-1569(1999)036<0328:TIOTP>2.3.CO;2)

20. Persson C, Lohmander A, Elander A. Speech in children with an isolated cleft palate: A longitudinal perspective. *Cleft Palate Craniofac J.* 2006;43(3):295-309. <http://dx.doi.org/10.1597/04-071.1>
21. Bae Y, Kuehn DP, Ha S. Validity of the nasometer measuring the temporal characteristics of nasalization. *Cleft Palate Craniofac J.* 2007;44(5):506-17. <http://dx.doi.org/10.1597/06-128.1>
22. Lewis KE, Watterson T, Quint T. The effect of vowels on nasalance scores. *Cleft Palate Craniofac J.* 2000;37(6):584-9. [http://dx.doi.org/10.1597/1545-1569\(2000\)037<0584:TEOVON>2.0.CO;2](http://dx.doi.org/10.1597/1545-1569(2000)037<0584:TEOVON>2.0.CO;2)
23. Kuehn DP, Moon JB. Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts. *J Speech Lang Hear Res.* 1998;41(1):51-62. <http://dx.doi.org/10.1044/jslhr.4101.51>
24. Watterson T, Lewis K, Allord M, Sulprizio S, O'Neill P. Effect of vowel type on reliability of nasality ratings. *J Commun Disord.* 2007;40(6):503-12. <http://dx.doi.org/10.1016/j.jcomdis.2007.02.002>
25. Gerratt BR, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing internal and external standards in voice quality judgments. *J Speech Hear Res.* 1993;36(1):14-20. <http://dx.doi.org/10.1044/jshr.3601.14>
26. Smith BE, Kuehn DP. Speech evaluation of velopharyngeal dysfunction. *J Craniofac Surg.* 2007;18(2):251-61. <http://dx.doi.org/10.1097/SCS.0b013e31803ecf3b>
27. Watterson T, Lewis KE, Deutsch C. Nasalance and nasality in low pressure and high pressure speech. *Cleft Palate Craniofac J.* 1998;35(4):293-8. [http://dx.doi.org/10.1597/1545-1569\(1998\)035<0293:NANILP>2.3.CO;2](http://dx.doi.org/10.1597/1545-1569(1998)035<0293:NANILP>2.3.CO;2)
28. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J Speech Hear Res.* 1993;36(1):21-40. <http://dx.doi.org/10.1044/jshr.3601.21>