Achados espectrais das vogais [a] e [ɐ̃] em diferentes aberturas velofaríngeas******

Spectral findings for vowels [a] and [ɐ̃] at different velopharyngeal openings

Aveliny Mantovan Lima-Gregio*
Jeniffer de Cássia Rillo Dutka-Souza**
Viviane Cristina de Castro Marino***
Maria Inês Pegoraro-Krook****
Plínio Almeida Barbosa*****

*Fonoaudióloga. Doutoranda em Linguística pelo Instituto de Estudos da Linguagem da Universidade Estadual de Campinas. Endereço para correspondência: R. das Opalas, 26 - Americana - SP. CEP 13471-220 (avelinylima@gmail.com).

**Fonoaudióloga. Doutora em Communication Sciences and Disorders pela University of Florida. Fonoaudióloga do Laboratório de Fonética Experimental do Hospital de Reabilitação de Anomalias Craniofaciais (HRAC) da Universidade de São Paulo (USP).

***Fonoaudióloga. Doutora em Communication Sciences and Disorders pela University of Florida. Professora do Departamento de Fonoaudiologia da Universidade Estadual de São Paulo - Marília.

****Fonoaudióloga. Doutora em Distúrbios da Comunicação Humana pela Universidade Federal de São Paulo. Professora do Departamento de Fonoaudiologia da Faculdade de Odontologia de Bauru da USP.

*****Engenheiro Eletrônico. Doutor em Signal-Image-Parole/Option Parole pelo Institut de la Communication Parlée e Institut National Polytechnique de Grenoble. Professor do Departamento de Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas.

******Trabalho Realizado no Laboratório de Fonética Experimental do HRAC-USP com Auxílio Financeiro do CNPq.

Artigo Original de Pesquisa
Artigo Submetido a Avaliação por Pares
Conflito de Interesse: não

Recebido em 1.12.2009.
Revisado em 19.8.2010; 05.11.2010.
Aceito para Publicação em 30.11.2010.

Abstract

Background: the size control of velopharyngeal opening is an important variable for the acoustic profile characterization of hypernasal speech. Aim: to investigate frequency spectral aspects of F1, F2, F3, nasal formant (FN) and anti-formant, in Hertz, for vowels [a] and [ɐ̃] at different velopharyngeal openings produced in the bulb of a palatal prosthesis replica used by a patient with velopharyngeal insufficiency. Method: speech recordings were obtained for four words ("pato/mato" and "panto/manto") produced within a carrier phrase in 5 conditions of velopharyngeal functioning: prosthesis with no openings (control condition: CC); prosthesis with bulb opening of 10mm^2 (experimental condition with 10mm^2 opening: EC10), prosthesis with a 20mm^2 opening (EC20), prosthesis with a 30mm^2 openning (EC30), and without the prosthesis (ECO). Five speech-language pathologists made a live rating of speech nasality during the reading of an oral passage. The recordings were used for spectral analysis. Results: F1 values were significantly higher for [a] when compared to [ɐ̃] in all conditions. F2 values for [a] in EC20 and EC30 were significantly lower than values in the other conditions, being closer to the values presented for [ɐ̃]. F3 values were not significantly different between the testing conditions. There was a relationship between FN and anti-formants, and the auditory perception of nasality for conditions EC10 and EC20. Conclusion: significant changes were observed in the studied spectral values according to changes in the velopharyngeal opening size.

Key Words: Velopharyngeal Insufficiency; Oral Fistula; Palatal Obturators; Spectrum Analysis.

Resumo

Tema: o controle do tamanho da abertura velofaríngea é uma variável importante na caracterização do perfil acústico da fala hipernasal. Objetivo: investigar os aspectos espectrais das frequências de F1, F2, F3, formante nasal(FN) e anti-formante, em Hertz, para as vogais [a] e [ɐ̃] na presença de aberturas feitas no bulbo de réplicas da prótese de palato de uma paciente com insuficiência velofaríngea. Método: gravações de produções de quatro palavras ("pato/mato" e "panto/manto") inseridas em frase veículo foram obtidas em cinco condições de funcionamento velofaríngeo: prótese sem aberturas (condição controle: CC), prótese com abertura de 10mm^2 no bulbo (condição experimental - CE10), com abertura de 20mm^2 (condição experimental - CE20), com abertura de 30mm^2 (condição experimental - CE30), e sem prótese (condição experimental aberta - CEA). Cinco fonoaudiólogos julgaram a nasalidade de fala ao vivo, durante a leitura de um texto oral. As gravações foram usadas para análise espectral. Resultados: valores de F1 foram significativamente mais altos para [a] que para [ɐ̃] em todas as condições. Valores de F2 para [a] em CE20 e CE30 foram significantemente mais baixos que nas outras condições, aproximando-se dos valores para [ɐ̃]. Valores de F3 não foram significativamente diferentes nas diferentes condições. Houve relação entre os achados de FN e anti-formantes e a percepção de nasalidade para as condições CE10 e CE20. Conclusão: foram observadas mudanças significativas nos valores espectrais estudados de acordo com alterações no tamanho da abertura velofaríngea.

Palavras-Chave: Insuficiência Velofaríngea; Fístula Bucal; Prótese Palatina; Análise Espectral.

Referenciar este material como:
Lima-Gregio AM, Dutka-Souza JCR, Marino VCC, Pegoraro-Krook MI, Barbosa PA. Spectral findings for vowels [a] and [ɐ̃] at different velopharyngeal openings (original title: Achados espectrais das vogais [a] e [ɐ̃] em diferentes aberturas velofaríngeas). Pró-Fono Revista de Atualização Científica. 2010 out-dez;22(4):515-20.

Introduction

Acoustic analysis techniques have been developed on the hypothesis that the perceptive judgment of an individual's speech is related to the information in the acoustic signal(1) of the utterance. Different types of acoustic analysis of speech signals have been proposed for studying resonance disturbances, particularly hypernasality(2-4), and also for verifying the possible relationship between perceptive judgment of nasality and its acoustic correlate(5-6).

The use of spectrographic analysis has benefited the study and understanding of hypernasality, particularly regarding the study of the temporal aspects of nasalization(7-8). However, there has been only limited application of the technique to the diagnosis and treatment of hypernasality, limiting the scope of what could otherwise be a very useful tool in clinical practice as it may foster facilitates pre/post-therapy comparisons(9-10).

Ever since the technique was invented in the 1940s, the most common use made of spectrographic analysis has been in the observation of vowel formant frequencies. In Brazil a few authors have conducted researches(11-14) to obtain descriptions of the oral and nasal vowels in speakers with no pathologies. Vieira(15) conducted a notable study involving the comparison of speakers with and without hypernasality but it did not involve performing any control over the size of the velopharyngeal opening during the production of nasalized speech.

The control of the size of velopharyngeal gap is an important variable in the definition of the acoustic profile of hypernasal speech given that hypernasality is the result of alterations of velopharyngeal functioning(1,16). Despite the complex nature of the relationship between the size of the velopharyngeal gap (gap=failure to close) and speech disturbances and the fact that such a relationship has not been fully described in the literature, variations in perception of hypernasality have been linked to variations in the size of the velopharyngeal gap(16).

Up until now very few studies have formally addressed the relationship of the size of the gap and hypernasality, nevertheless, research involving artificial openings has shown that even an experimental opening of 20mm2 can have a considerable negative effect on speech, especially in regard to hypernasality(17-19). Based on the above considerations, this study sets out to investigate the spectral aspects F1, F2, F3, FN and anti-formant, in Hertz, for the vowels [a] and [□̃] at different velopharyngeal openings. The openings have been introduced in the speech bulb of replicas of a palatal prosthesis worn by a patient with velopharyngeal insufficiency. The assumption made in this study is that the introduction of experimental openings in the pharyngeal bulb would result in hypernasal since while wearing the palatal obturator prosthesis, the patients´ resonance was considered normal during perceptual ratings made by speech-language pathologists (SLPs). The hypothesis was that different sizes of openings would give rise to different spectral measurements.

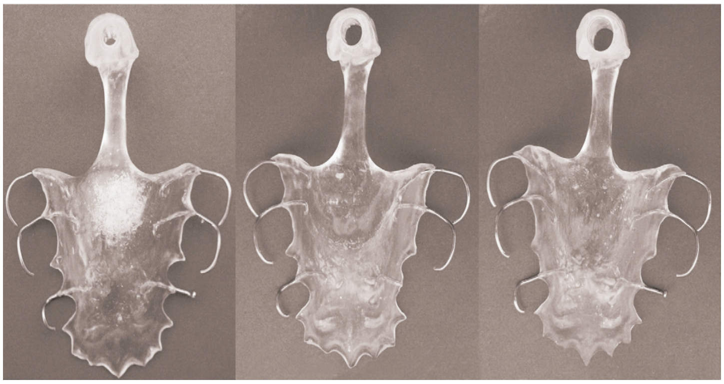
Method

The Ethics Committee of the Hospital de Reabilitação de Anomalias Craniofaciais (HRAC) at the Universidade de São Paulo (USP) approved the study (official document #038/2004 0), and the participant signed a Informed Consent Form (ICF) having been fully informed of the experiment and having received assurances of confidentiality to guard against any form of identification. The participant was also informed that she could withdrawal from the study at any time, all in compliance with the Brazilian National Health Council's Resolution 169 dated October 10, which regulates studies undertaken with human beings.

The speech samples were taken from a 17-year--old female with an operated incomplete palatal cleft. After palatoplasty the patient presented velopharyngeal insufficiency which was successfully corrected with a palatal prosthesis involving a pharyngeal bulb (speech bulb). For the purposes of this experiment it was verified that subject's palatal prosthesis was properly fitted. Prior to the study auditory-perceptual ratings of speech and nasalance scores indicated normal resonance with the prosthesis in place while indicated hypernasal resonance when the prosthesis was removed. The patient had no history of upper respiratory illness, was a non-smoker and was not under the effects of any sort of medication that could affect speech resonance at the dates the study was conducted.

To simulate velopharyngeal openings with different areas, three identical replicas of the patient's prosthesis were prepared but each one with an aperture of a different dimension in the center of the

FIGURE 1. View from above of the experimental acrylic prosthesis with hole of 10mm2, 20mm2 and 30mm2, respectively.



pharyngeal bulb (in the velopharyngeal area). To prepare the apertures, the three experimental bulbs were treated with acrylic and a hole was made in the center of the bulb using drill bits yielding three holes with sizes of approximately 10mm2, 20mm2 and 30mm2 each (figure 1).

The participant was instructed to wear the prosthesis one week prior to the study to adapt to any possible variations between the replicas and her original prosthesis. During that period of adaptation the holes were blocked with wax so that the subject's habitual resonance should not be altered. Nasometric and nasoendoscopic examinations showed that there was no velopharyngeal gap during speech between the pharynx walls and the bulb when the replicas were in place. The three SPLs and the participant could not detect any auditory perceptual differences in speech produced with the three replicas and the original prosthesis.

Control Condition (CC): a statistical comparison between the acoustic measurements extracted from the enunciations produced by the patient using each of the three replicas failed to reveal any significant difference (p=0,51) among them. Subsequently, samples obtained with any of the three replicas with the holes blocked with wax were treated as the control condition for this study, indicative of speech obtained with the original prosthesis.

Experimental Conditions: the four experimental conditions were established as follows: (I) with the subject wearing the first replica with an opening of 10mm2 (EC10); (II) with the subject wearing the second replica with an opening of 20mm2 (EC20); (III) with the subject wearing the third replica with an opening of 30mm2 (EC30); and (IV) with the subject wearing no prosthesis corresponding to velopharyngeal opening of unknown dimensions (ECO).

Two minimum pairs of words with a CVCV structure (consonant-vowel-consonant-vowel) structure were used in which the first consonants of each pair contrasted an oral and a nasal consonant, namely: pato/mato and panto/manto. The phrases were inserted into the carrier phrase "Digo ____ bojudo", chosen on the basis of previous studies for Brazilian Portuguese (BP)(12-14). The participant repeated each phrase five times for each condition, randomly mixing the control and 4 experimental conditions. A total of 100 speech samples (5 repetitions, x 5 conditions, x 4 words) were obtained.

The data was obtained using an Audio Technica AKG C-420 cardioid headset and digitalized by means of a Creative Sound Blaster sound card (Audigy Model) using Sound Forge (version 7) software in mono-channel with a sampling rate of 44100 Hz and 16 bits.

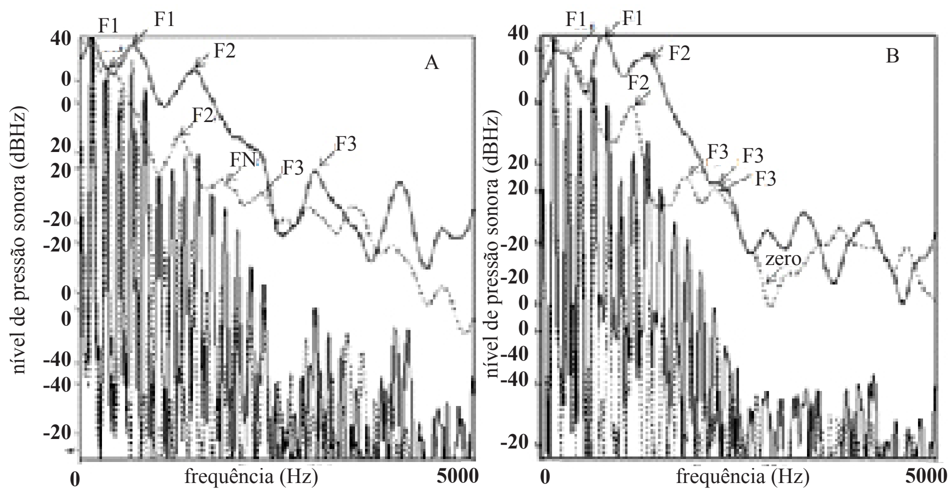
Praat(21) software was used to obtain the measurements of interest for this study. Among the data analyzed were: 1) the frequency in Hertz of the first formant (F1), the second formant (F2), and the third Formant (F3) of the central part of the vowels [a] and [ɐ̃]; 2) the frequency in Hertz of the first nasal formant (FN) in the nasalized portion of [ɐ̃]; 3) and the frequency in Hertz of the first spectral anti-formant (Z) in the nasalized portion of [ɐ̃]. F1N was identified as the closest frequency to F1 of the vowel [ɐ̃] because it has a mandibular configuration close to that of [ɐ̃] (22).

The spectral measurements were obtained using FFT (Fast Fourier Transform) analyses and Cepstral smoothing embedded in the Praat software, and in every case, a comparison was made of the vowel [ɐ̃] with the vowel [a] of the minimum pair, and considering global energy attenuation and the insertion of nasal formants associated to low frequencies as the identifiers of nasality (see figure 2).

During data recording for this study, a live auditory perceptual evaluation was conducted independently by five experienced SLPs. They were instructed to rate only the presence and the absence of hypernasality while the participant was reading out loud an oral passage in Brazilian Portuguese ("A história do urso preto")(23), similar to the "Zoo Passage", involving no nasal sounds. The text was used in order to resembled as far as possible, the situation of spontaneous conversation. The passage was read five times, each under one of the study conditions which were not disclosed to the speech judges (SLPs).

Three factors were the object of Variance Analysis (ANOVA): 1–Conditions: CC, EC10,

FIGURE 2: A) FFT and cepstral analyses of the vowels [a] (unbroken line) and [ɐ̃] (dotted line) of "pato" and "panto". B) FFT and cepstral analyses of the vowels [a] (unbroken line) and [ɐ̃] (dotted line) of "mato" and "manto".



EC20, EC30, ECO; 2–Words: beginning with [p] or [m]; and 3–Vowels: [a] and [ɐ̃]). Formants (F1, F2, F3, in Hertz) were treated as three dependant variable. In the case of the nasal formants and anti--formants, the ANOVA test was applied to two factors only (1–Condition and 2–Word), considering that those parameters were only extracted for the vowel [ɐ̃]. Differences with significance levels of 5% were submitted to the post-hoc Scheffé testing to check the homogeneity of the groups.

Intrajudge reliability during auditory perceptual evaluation of nasality was calculated for the 5 SLPs for each condition described and described in terms of percentages of agreement.

Results

The mean frequencies in Hertz for F1, F2, F3, for the oral vowel [a], in "pato and mato" were in the range of 655 Hz to 764 Hz for F1; 1339 Hz to 1571 Hz for F2; and 2095 Hz to 2627 Hz for F3. The mean frequencies of the formants were lower in the experimental conditions (with the velopharyngeal openings).

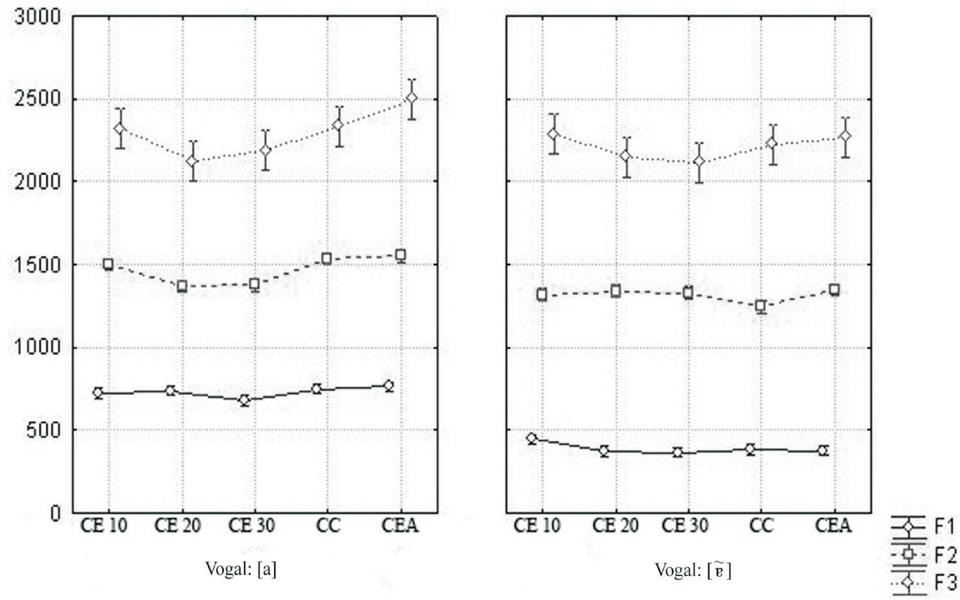
The mean frequencies in Hertz for F1, F2, F3, for the vowel [ɐ̃], in "panto and manto" were in the range of 363 Hz to 487 Hz for F1; 1167 Hz to 1371 Hz for F2; and 2073 Hz to 2296 Hz for F3. The mean frequencies of the oral formants for the nasal vowel [ɐ̃], in "panto and manto" were considerably lower than for the vowel [a], in "pato and mato".

The application of the ANOVA test to the three factors (Conditions, Words, Vowels) and the three dependant variables (F1, F2, F3, in Hertz) showed significant differences associated to the conditions and the vowels [a] and [ɐ̃], for the three formants ($F12,206=7.2$; $p<10$-5), for $\alpha=0.05$. The abscissa in Figure 3 shows the conditions investigated and the ordinate shows the frequency values obtained for the three formants; on the left for the vowel [a] and on the right for the vowel [ɐ̃].

The Scheffé post hoc test for homogeneity showed that the sample for F1 was divided into two groups: higher values for [a] and lower values for [ɐ̃]. In the case of F2, the analysis showed that the sample could also be divided into two groups: (I) the higher values for F2 formed by EC10, ECO and CC of the vowel [a] in "pato" and "mato"; and (II) the lower values for F2 associated to all the control conditions and experimental conditions for the vowel [ɐ̃] in addition to those associated to EC20 and EC30 for the vowel [a]. Thus the EC20 and EC30 for the vowel [a] showed significantly similar behavior to that of all the conditions for the nasal vowel [ɐ̃] in regard to F2 thereby demonstrating an important nasality effect in the final part of the second formant. In regard to the analysis of F3 it was found that the values obtained provided no relevant information on nasality, which was to a certain extent to be expected because the context with the consonant was controlled (and this formant is more strongly affected by the consonantal context, cf.(24)).

The ANOVA test was applied for parameters FN (Hz) and Zero (Hz). In the case of FN, for the interaction Condition versus Word, the test revealed a statistically significant difference ($F4,40=35.7$; $p<10$-5) for $\alpha=0.05$, and the post-hoc Scheffé test indicated that the mean value for ECO was the highest (1804

FIGURE 3. Mean values for F1, F2, F3 (Hz), in conditions EC10, EC20, EC30, CC and ECO.



Hz). For the spectral trough (Zero), the ANOVA statistical test revealed a statistically significant difference for the factors Condition ($F_{4,40}=49.7$; $p<10^{-5}$), Word ($F_{1,18}=14.1$; $p=0.001$) and for the interaction Condition versus Word ($F_{4,40}=8.9$; $p<10^{-5}$), all for $\alpha=0.05$. In this last case the analysis showed that the words with [p] (pato and panto) showed higher values than those with [m] (mato and manto). The post-hoc Scheffé analysis however revealed five different groups; four of them with interactions among one another and another totally distinct group formed of ECO and the words "pato and panto".

The results of the perceptive-auditory evaluation revealed that for conditions CC and EC10 there was 100% agreement regarding absence of hypernasality while for conditions EC30 and ECO there was 100% agreement regarding identification of hypernasality. The only condition that did not result in a unanimous perceptual rating among the SLPs was EC20, which was rated with 60% agreement regarding the presence of hypernasal speech.

Discussion

The higher F1 values obtained for the vowel [a] as compared to the vowel [ɐ̃], in all the conditions were expected, explained by articulatory and nasality aspects that are different between both vowels. Lower F1 values for the nasal vowel [ɐ̃] is characteristic of speech nasality. The main determinant of the F1 values is the position of the mandible, and the height of that articulator – greater in the nasalized vowel(25) – is inversely proportional to the frequency of the formant. Therefore, the lowering of the values obtained for formant F1 can be explained as an effect of the nasalization resulting from conditions of velopharyngeal openings for the vowel [a] has also been registered in the respective literature(26-28).

The F2 values obtained for the vowel [a] in EC20 and EC30, significantly lower than those obtained in CC, EC10 and ECO – even though the last two correspond to conditions of openings –, and coming close to F2 values obtained during the production of the nasal vowel [ɐ̃], are in agreement with what has been registered in the literature(3,24). Considering that it is the position of the tongue along the anterior posterior axis that is the main determinant of the second formant(25), then its posteriorization in conditions of nasality in an attempt to compensate for the loss of air due to the velopharyngeal opening could be the explanation for this finding. In the same way, the fact that F2 values in ECO have been similar to those in CC suggests that the subject may have made other compensations, such as laryngeal and respiratory adjustments (18).

The data analysis did not show that increases in the velopharyngeal opening were accompanied buy a gradual diminishing of the formants (linear relation between the degree of opening and the spectrum findings) as had originally been expected. Coleman Jr. (29) suggested that variability among individuals could be a possible explanation for the non-linearity.

As for the perceptive-auditory evaluation, the presence of anti-resonances proved to be essential for the identification of nasal vowels(30) by the evaluators. The fact that EC10 was evaluated as a normal resonance may be related to the values for FN which were non-existent for [a] and similar to those of CC for [□̃]. Finally, in EC20, a possible explanation for the 60% agreement is that it represents a threshold condition of velopharyngeal opening and also that FN and Zero values were sometimes similar to those of CC and EC10 and sometimes to those of EC30 and ECO.

Conclusion

The hypothesis that different openings would lead to different spectrum measurements was confirmed and although no linear relationship was detected between the experimental conditions and the formant values, significant changes in the spectral values did occur according to the alterations in the size of the velopharyngeal opening.

References

1. Bzoch K. Communicative Disorders: related to cleft lip and palate. Austin: Pro-Ed; 2004.

2. Dalston RM, Warren DW, Dalston ET. Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment. Cleft Palate-Craniofacial Journal. 1991;28(2):185-9.

3. Kent RD, Liss J, Philips BJ. Acoustic analysis of velopharyngeal dysfunction in speech. In: Bzoch K. Communicative Disorders: related to cleft lip and palate. Boston: Little Brown; 1989. p. 258-70.

4. Vogel AP, Ibrahim HM, Reilly S, Kilpatrick N. A comparative study of two acoustic measures of hypernasality. Journal of Speech, Language, and Hearing Research. 2009;52:1640-51.

5. Lee AS-Y, Ciocca V, Whitehill TL. Acoustic correlates of hypernasality. Clinical Linguistics & Phonetics. 2003;17(4-5):259-64.

6. Triona S, Debbie S. Relationship between perceptual ratings of nasality and nasometry in children/adolescents with cleft palate and/or velopharyngeal dysfunction. International Journal of Language & Communication Disorders. 2008;43(3):265-82.

7. Seunghee Ha, Hyunsub Sim, Minje Zhi, David P. Kuehn. An acoustic study of the temporal characteristics of nasalization in children with and without cleft palate. Cleft Palate-Craniofacial Journal. 2004;41(5):535-43.

8. Bae Y, Kuehn DP, Ha S. Validity of the nasometer measuring the temporal characteristics of nasalization. Cleft Palate--Craniofacial Journal. 2007;44(5):506-17.

9. Kent RD. Hearing and believing: some limits to the auditory--perceptual assessment of speech and voice. American Journal of Speech-Language Pathology. 1996;5:7-23.

10.Casal C, Domingues C, Fernández A, Sarget R, Martinez--Celdrán E, Sentís-Vilalta J, Gay-Escoda C. Spectrographic measures of the speech of young children with cleft lip and palate. Folia Phoniatr Logop. 2002;54:247-57.

11. Behlau MS. Uma análise das vogais do Português Brasileiro falado em São Paulo: perceptual, espectrográfica de formantes e computadorizada de freqüência fundamental. [Dissertação]. São Paulo: Escola Paulista de Medicina; 1984.

12. Souza EG. Para caracterização fonético acústica da nasalidade do Português Brasileiro. [Dissertação]. Campinas: Universidade Estadual de Campinas; 1994.

13. Seara IZ. Estudo acústico-perceptual da nasalidade das vogais do Português Brasileiro. [Tese]. Florianópolis: Universidade Federal de Santa Catarina; 2000.

14. Jesus MSV. Estudo fonético da nasalidade vocálica. In: Reis C. Estudos em fonética e fonologia do Português. Belo Horizonte: Fale-UFMG; 2002. p. 205-22.

15. Vieira JM. Perfil espectrográfico da hipernasalidade de fala de mulheres portadoras de fissura palatina. [Dissertação]. São Carlos: Universidade de São Paulo; 2003.

16. Warren DW, Dalston RM, Mayo R. Hypernasality and velopharyngeal impairment. Cleft Palate Craniofacial Journal. 1994;31:257-62.

17. Henningsson G, Isberg A. Influence of palatal fistulae on speech and resonance. Folia Phoniatrica. 1987;39:183-91.

18. Sapienza CM, Brown WS, Williams WN, Wharton PW, Turner GE. Experimental coupling of oral and nasal cavities. Cleft Palate-Craniofacial Journal. 1996;33(2):118-26.

19. Marino VC, Williams WN, Wharton PW, Paul MF, Dutka--Souza JCR, Schulz GM. Immediate and sustained changes in lingual movements after an experimental after an experimental palatal fistula: a case study. Cleft Palate-Craniofacial Journal. 2005;42(3):286-96.

20. Warren DW. Aerodynamic assessments and procedures to determine extent of velopharyngeal inadequacy. In: Bzoch KR. Communicative Disorders: related to cleft lip and palate. Austin: Pro-Ed; 1997. p. 411-37.

21. Boorsman P, Weenink D. Praat: doing phonetics by computer. Download da versão 4.3.37, em 2006, do site: http://www.praat.org.

22. Stevens KN. Acoustics phonetics. Massachusetts: MIT Press; 2000. p. 315-6.

23. Suguimoto MLCP, Pegoraro-Krook, MI. Avaliação nasométrica em adultos normais falantes do português brasileiro. Pró-Fono Revista de Atualização Científica. 1995;7(2):3-9.

24. Fant G. Acoustic theory of speech production. The Hague: Mouton; 1970.

25. Kent RD. Acoustic Phonetics. In: Kent RD. The Speech Sciences. Singular: San Diego; 1997. p. 329-70.

26. Delattre P. Les attributs acoutiques de la nasalité vocalique et consonantique. Studia Linguistica. 1954;8:103-9.