



# Articulation Analysis in the Speech of Children with Cleft Lip and Palate

H. A. Carvajal-Castaño<sup>1,2</sup> and Juan Rafael Orozco-Arroyave<sup>1,3</sup>(✉)

<sup>1</sup> Research Group on Applied Telecommunications - GITA, Electronic Engineering and Telecommunications Department, Faculty of Engineering, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia

[rafael.orozco@udea.edu.co](mailto:rafael.orozco@udea.edu.co)

<sup>2</sup> Bioinstrumentation and Clinical Engineering Research Group - GIBIC, Bioengineering Department, Faculty of Engineering, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia

<sup>3</sup> Pattern Recognition Lab, University of Erlangen-Nürnberg, Erlangen, Germany  
<http://www.udea.edu.co>, <http://www.fau.eu>

**Abstract.** Hypernasality is a speech deficit that affects children with cleft lip and palate (CLP). It is characterized by the lack of control of the velum, which causes problems when controlling the amount of air passing from the oral to the nasal cavity while speaking. The automatic evaluation of hypernasality could help in the monitoring of speech-language therapies and in the design of better oriented exercises. Several articulation features have been used for the automatic detection of hypernasal speech. This paper evaluates the suitability of classical articulation features for the automatic classification of hypernasal and healthy speech recordings. Two different databases are considered with recordings collected under different acoustic conditions and with different audio settings. Besides the evaluation of the proposed approach upon each database separately, non-parametric statistical tests are performed to evaluate the possibility of merging features from the two databases with the aim of finding more robust systems that could be used in different acoustic conditions. The results indicate that the proposed approach has a high sensitivity, which indicates that it is suitable to detect hypernasal speech samples. We believe that promising results could be obtained with this approach in future experiments where the degree of hypernasality is evaluated.

**Keywords:** Cleft lip and palate · Hypernasality · Articulation measures · Classification

## 1 Introduction

Cleft Lip and Palate (CLP) is a craniofacial malformation that occurs in about one in every 700 live births [1]. This malformation is characterized by the incomplete formation of tissues that separate oral and nasal cavities, which generates

© Springer Nature Switzerland AG 2019

I. Nyström et al. (Eds.): CIARP 2019, LNCS 11896, pp. 575–585, 2019.

[https://doi.org/10.1007/978-3-030-33904-3\\_54](https://doi.org/10.1007/978-3-030-33904-3_54)

several speech disorders such as hypernasality, hyponasality, glottal stops, and others. During speech production it is necessary to control the amount of air that comes out through the nasal cavity. The amount of air is controlled by the velum, which opens and closes the connection between the oral and nasal cavity depending on which sound is intended to be produced by the speaker (e.g., nasal or no-nasal). When there is excess of air coming out through the nasal cavity the speech is perceived as hypernasal, which is the speech pathology suffered by the majority of CLP children [2, 3].

The evaluation of CLP patients is subjective and time-consuming because highly depends on the speech and language therapist's expertise. The research community has been interested since several decades in the development of systems that allow the objective evaluation of speech in children with CLP. One of the most suitable approach has been the use of features that reflect articulation deficits. The most common ones include Mel-frequency cepstral coefficients (MFCCs), the first two vocal formants, vowel space area (VSA), and non-linear measures like the Teager Energy Operator (TEO). For instance, in [4] speech recordings of CLP patients are studied with the aim of discriminating the severity of hypernasality. The extracted features were based on the fundamental frequency of voice, energy content, and MFCCs. The authors reported accuracies of up to 80.4% when discriminating four different grades of hypernasality. In [5] the five Spanish vowels are considered to evaluate hypernasality in children with CLP considering acoustic and noise features. The authors reported accuracies of up to 89%. In [6] the authors introduced a method for the classification of hypernasal patients and healthy control subjects using the VSA and 13 MFCCs with their corresponding first and second derivatives. According to the authors, the proposed approach has an accuracy of 86.89%. In [7] features based on acoustic, noise, cepstral analysis, nonlinear dynamics [8] and entropy measures are used for hypernasality detection. Accuracies of around 92% are reported when the five Spanish vowels and the words *coco* and *gato* are considered.

Although articulation-related features have been extensively used in the literature to evaluate hypernasality in the speech of CLP patients, all of the works report results on single databases. Thus, there is no evidence about the robustness of those features when are used in different databases recorded in different acoustic conditions and with different recording settings. This study evaluates the suitability of classical articulation-related features to discriminate between hypernasal and healthy speech signals collected from children with CLP from two different clinics for children in Colombia, using different recording settings and under different acoustic conditions. The features set extracted for study includes 12 MFCCs and their first and second derivatives, the first two vocal formants ( $F_1$  and  $F_2$ ) and their first and second derivatives, TEO formant centralization ratio (FCR), and VSA. Four statistical functionals are calculated upon each feature vector: mean, standard deviation, kurtosis and skewness, and two different classifiers are evaluated: a support vector machine (SVM) and a Random Forest (RF).

## 2 Participants

### 2.1 CLP Manizales

This database was provided by Grupo de Control y Procesamiento Digital de Señales - (GCyPDS) at Universidad Nacional de Colombia, Manizales. This database contains recordings of the five Spanish vowels pronounced by children between 5 and 15 years old. A total of 140 audio registers were collected, 84 labeled as hypernasal by a phoniatriy expert and the remaining 56 were labeled as healthy. The signals were recorded in a quiet room but under non-controlled acoustic conditions using a non-professional audio setting with a sampling frequency of 44100 Hz and 16 bit-resolution.

### 2.2 CLP Clínica Noel

This data set was recorded in the Clínica Noel from Medellín and contains recordings of the five Spanish vowels pronounced by children between 5 and 15 years old. The data contain 95 audio recordings of 53 children with CLP and 42 healthy controls (HC). All of the 53 children included in the CLP group were labeled as hypernasal by a phoniatriy expert. The recordings were collected in a quiet room using a professional audio setting with a sampling frequency of 44100 Hz and 16 bit-resolution. Further details of this corpus can be found in [9].

## 3 Methods

### 3.1 Feature Extraction

**Mel-Frequency Cepstral Coefficients (MFCCs).** Since several decades, these features have been widely used in applications for automatic speech recognition and for speaker identification. However, since about a decade the MFCCs started to be used in applications of pathological speech analysis like laryngeal pathologies [10] and dysarthria in Parkinson's disease [11]. The coefficients are based on the Mel scale which considers the perceived frequency of a tone and the actual measured frequency. The scale emulates the frequency-response of the human hearing system.

To estimate MFCCs considers that  $s_i(k)$  is the  $i^{th}$  frame of the speech signal and  $S_i(k)$  is the Discrete Fourier Transform (DFT), the mel-spectrum is [12]:

$$\text{MF}_i[r] = \frac{1}{A_r} \sum_{k=L_r}^{U_r} |V_r[k]S_i[k]|^2, \quad r = 1, 2, \dots, R \quad (1)$$

where  $R$  is the number of filters,  $V_r[k]$  is a weigh function for each mel filter,  $L_r$  and  $U_r$  are the lower and upper range of the filter, and  $A_r$  is a normalization factor for the  $r^{th}$  filter. The MFCC are finally found as:

$$\text{MFCC}_i[n] = \frac{1}{R} \sum_{r=1}^R \log(\text{MF}_i[r]) \cos \left[ \frac{2\pi}{R} \left( r + \frac{1}{2} \right) n \right] \quad (2)$$

MFCC<sub>*i*</sub>[*n*] is evaluated for  $n = 1, 2, \dots, N_{\text{MFCC}}$  where  $N_{\text{MFCC}}$  is the number of desired coefficients ( $N_{\text{MFCC}} < R$ ). Besides, the first and second derivatives of MFCCs is compute. First and second derivatives of the MFCCs are also extracted before computing the four statistical functionals, forming a 144-dimensional feature vector.

**First and Second Vocal Formants ( $F_1$  and  $F_2$ ).** These two frequencies correspond to the first two peaks that appear, when the envelope of a voice spectrum is calculated (typically by a linear predictive filter).  $F_1$  and  $F_2$  provide information about resonances that occur in the vocal tract during a phonation. Thus they are related to the shape of organs and tissues in the vocal cavity, e.g., the tongue. Typically, the first two formants are used to evaluate the capability of a speaker to keep the tongue in a certain position during the phonation of a given vowel [6]. In this study, we calculate  $F_1$  and  $F_2$  considering the spectral envelope found with a Linear Predictive Coding (LPC) filter. This method assumes that each frame,  $s_i(k)$ , of the speech signal, can be approximated as a linear combination of the past  $p$  samples:

$$\hat{s}_i = \sum_{k=1}^p a_k s_{i-k} \quad (3)$$

where  $\mathbf{a} = a_1, \dots, a_p$  is a vector with  $p$  coefficients. The main aim is to minimize the mean-square-error such that:

$$\mathbf{a} = \arg \min_{\mathbf{a}} \frac{1}{N} \sum_{n=1}^N (\hat{s}_i - s_i)^2 \quad (4)$$

The optimal vector  $\mathbf{a}$  that minimizes the MSE comprises the envelope of the speech spectrum and the value of  $p$  determines the shape of such an envelope. As in the case of the MFCC, the first and second derivatives of  $F_1$  and  $F_2$  are also calculated before computing the four statistical functional, forming a 24-dimensional feature vector.

**Teager Energy Operator (TEO).** Consider the signal  $x(n)$ . Its associated TEO was defined in [13] as:

$$\Psi[x(n)] = x^2(n) - x(n+1)x(n-1) \quad (5)$$

One of the most important characteristics of TEO is the sensibility to composed signals. If we consider a composed signal as  $x(n) = s(n) + g(n)$ , the TEO is computed as:

$$\Psi[x(n)] = [s(n) + g(n)]^2 - [s(n+1) + g(n+1)][s(n-1) + g(n-1)] \quad (6)$$

if a cross-correlation term ( $\Psi_{\text{cross}}[s(n), g(n)]$ ) between the two signals is considered as  $\Psi_{\text{cross}}[s(n), g(n)] = s(n)g(n) - g(n+1)s(n-1)$ , we obtain:

$$\Psi[x(n)] = \Psi[s(n)] + \Psi[g(n)] + \Psi_{\text{cross}}[s(n), g(n)] + \Psi_{\text{cross}}[g(n), s(n)] \quad (7)$$

Equation 7 shows that the superposition theorem does not apply to TEO. This property is useful in cases where the original signal is composed by several components. A regular speech spectrum has a typical profile where its peaks are characterized by the vocal formants. In the case of a hypernasal speech spectrum, additional peaks (additional formants) appear and also additional anti-formants (additional valleys in the speech spectrum) appear. TEO seems to be a good strategy to model such additional components that result in the spectrum of hypernasal signals [14]. The TEO is extracted from windows of 40 ms-length, then the four statistical functionals are computed to create a 4-dimensional feature vector per speaker.

**Formant Centralization Ratio (FCR)** is an alternative measure to represent articulatory problems in speakers. It offers the advantage of maximizing the sensitivity to vowel centralization and minimizing the sensitivity to inter-speaker variability. Thus, provides more robust and stable information of the vowel space produced by a speaker. The FCR was proposed in [15] as:

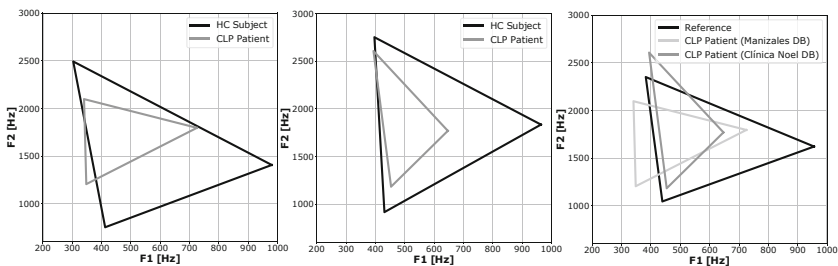
$$FCR = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}} \tag{8}$$

Where  $F_{1a}$ ,  $F_{2a}$ ,  $F_{1i}$ ,  $F_{2i}$ ,  $F_{1u}$  and  $F_{2u}$  are the first and second format of the corner vowels /a/, /i/ and /u/ respectively.

**Vowel Space Area (VSA)** is the most common way of measuring vowel centralization using  $F_1$  and  $F_2$  of corner vowels (/a/, /i/, /u/). It is given by the area of the triangle formed by the vertexes ( $F_1$ ,  $F_2$ ) in the vowel space created for the three corner vowels. VSA is computed as [15]:

$$VSA = \left| \frac{F_{1i}(F_{2a} - F_{2u}) + F_{1a}(F_{2u} - F_{2i}) + F_{1u}(F_{2i} - F_{2a})}{2} \right| \tag{9}$$

Figure 1 shows the resulting vocal triangle for two different children from the two databases. Note that in both cases the triangle of the patients is compressed which is a typical indicator of reduced articulation capability.



**Fig. 1.** Vocal triangle computed for two different children from the two databases.

At the end of the feature extraction procedure, each vowel pronounced by each speaker is modeled by a 172-dimensional feature vector. Additionally, FCR and VSA are calculated per speaker. Since the five Spanish vowels are considered together, each speaker is finally represented by a 862-dimensional feature vector.

### 3.2 Classification

With the aim of comparing the robustness of two different classification approaches, two classifiers are considered: a Support Vector Machine (SVM) with a Gaussian kernel, and a Random Forest (RF). The parameters of the classifiers were optimized following a 5-fold cross-validation strategy, where 4 folds were used for training and the remaining one for test. Within the 4 folds used for training, another 5-fold cross validation was performed. The optimization criterion is based on the accuracy in training. The parameters were optimized using a grid-search over the training folds. For SVM,  $\mathbf{C}$  and  $\gamma \in \{10^{-6}, 10^{-5}, \dots, 10^4\}$  and for RF, the number of trees  $\mathbf{N} \in \{5, 10, 20, \dots, 100\}$  and the depth of the decision trees  $\mathbf{D} \in \{2, 5, 10, 20, \dots, 100\}$ . Optimal parameters are selected according to the mode across the 5-fold cross-validation procedure.

### 3.3 Statistical Analysis and Merging of Features

Apart from calculating the features and performing the classification between the two classes, we wanted to evaluate the possibility of merging those features extracted from two different datasets that were collected under different acoustic conditions. The analysis to decide which features were able to be merged between datasets was performed according to Kruskal-Wallis statistical tests. The null hypothesis was that the given feature has the same distribution in both databases. Thus, if the p-value  $< 0.05$  the null hypothesis is rejected. The test was applied over all features of both databases and only the features that successfully passed the test (p-value  $\geq 0.05$ ) were included in the merging process. At the end of the procedure a total of 508 features passed the test.

## 4 Experiments and Results

In this study three experiments are performed: (1) classification of CLP vs. HC with the Manizales database, (2) classification of CLP vs. HC with the Clínica Noel database and (3) classification of CLP vs. HC with the fusion of both databases only considering those features that successfully passed the statistical test.

### 4.1 Experiment 1 - Manizales DB

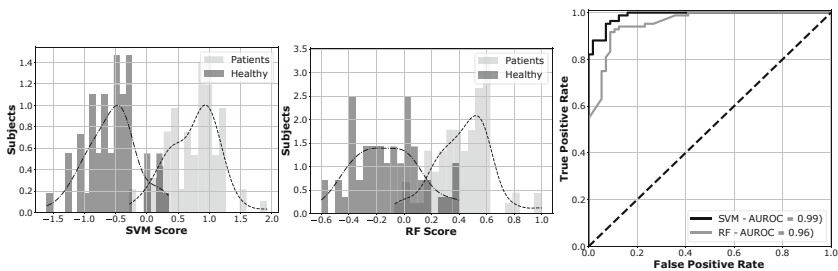
Results for the classification of CLP patients vs HC subjects are presented in Table 1. Note that the best results are always obtained using the SVM.

Figure 2 shows histograms of the scores obtained during the classification process. Additionally, Receiver Operating Characteristic (ROC) curves obtained with the two classifiers are also included. Note that the scores in an SVM are the distance of each sample to the separating hyperplane. For RF the scores are the probability of a sample to belong to the selected class. From the histograms it is possible to observe that most of the CLP patients are correctly classified, which confirms the high sensitivity obtained in the experiments (96.4%). For the HC subjects the result is not as high but still competitive with a specificity of 89.6%. The values of AUROC allow to perform a more compact analysis of the system’s performance considering the classification accuracy of the two classes at the same time (sensitivity and specificity).

**Table 1.** CLP vs. HC using the Manizales DB

Classifier	ACC [%]	Sen [%]	Spe [%]	AUROC	Parameters
SVM	93.6 ± 6	96.4 ± 5	89.6 ± 9	0.99	C = 10, $\gamma = 10^{-4}$
RF	86.4 ± 5	95.2 ± 6	73.0 ± 16	0.96	D = 2, N = 10

ACC: accuracy in the test set, AUROC: Area under ROC curve. Sen: Sensitivity, Spe: Specificity.  
 C and  $\gamma$ : complexity parameter and bandwidth of the kernel in the SVM classifier.  
 N and D: Number of trees and depth of the decision trees in the RF classifier



**Fig. 2.** Histograms of the scores and ROC curves (Manizales DB)

### 4.2 Experiment 2 - Clínica Noel DB

The results obtained with the two classifiers with data of the Clínica Noel are indicated in Table 2. Similar to Experiment 1, Fig. 3 the histograms of the scores and the ROC curves are presented. Note that as in the experiments with the other dataset, the best results are obtained with the SVM. This result confirms the robustness of these classifiers, which have been extensively used in the literature in problems of pathological speech processing. When this results are compared to those presented in Fig. 2, the histograms have a larger overlapping which reduces the performance of the classifiers.

**Table 2.** CLP vs. HC using the Clínica Noel DB

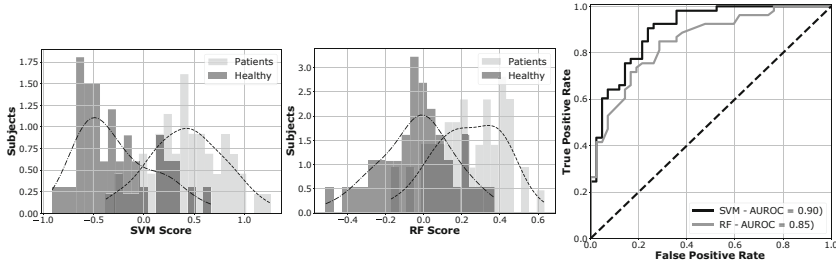
Classifier	ACC [%]	Sen [%]	Spe [%]	AUROC	Parameters
SVM	84.3 ± 7	90.8 ± 6	76.6 ± 11	0.90	C = 10, $\gamma = 10^{-4}$
RF	76.9 ± 7	88.4 ± 16	62.6 ± 19	0.85	D = 10, N = 70

ACC: accuracy in the test set, AUROC: Area under ROC curve. Sen: Sensitivity,

Spe: Specificity.

C and  $\gamma$ : complexity parameter and bandwidth of the kernel in the SVM classifier.

N and D: Number of trees and depth of the decision trees in the RF classifier

**Fig. 3.** Histograms of the scores and ROC curves (Clínica Noel DB)

### 4.3 Experiment 3 - Fusion of both DB

Only those features that successfully passed the Kruskal-Wallis test explained in Sect. 3.3 are considered in this experiment. The results obtained with the resulting sub-set of features in the Clínica Noel and Manizales databases are included in Tables 3 and 4. The results after merging both datasets are presented in Table 5. Note that the results after selecting those features that passed the statistical tests are lower than those obtained when considering the complete set of features. This result can be likely explained because only features robust against different acoustic conditions were included after the statistical test, hence those features are not necessary the most suitable to model articulation deficits. On the other hand, Table 5 indicates that when the two databases are merged the results improve. It seems like the difference in the acoustic conditions of the two databases allow the selected features to complement among them and the result is the improvement in the classification accuracies compared to those obtained when the selected features are used upon each database separately. Figure 4 shows the histograms of the scores obtained in the classification process and the ROC curves. Similarly to the previous experiments, there is a high sensitivity obtained with both classifiers. This result indicates that the proposed approach is suitable to detect CLP patients then it seems to be sensitive to articulation deficits exhibited by children with CLP. We think that it could be used in future studies to evaluate the degree of nasalization. We are currently collecting more data with those labels with the aim of performing these kinds of experiments.



**Table 3.** CLP vs. HC using the Manizales DB with selected features

Classifier	ACC [%]	Sen [%]	Spe [%]	AUROC	Parameters
SVM	79.2 ± 4	90.4 ± 8	62.8 ± 6	0.86	$C = 10, \gamma = 10^{-4}$
RF	70.1 ± 7	92.8 ± 2	36.0 ± 16	0.80	D = 90, N = 30

ACC: accuracy in the test set, AUROC: Area under ROC curve. Sen: Sensitivity.

Spe: Specificity.

C and  $\gamma$ : complexity parameter and bandwidth of the kernel in the SVM classifier.

N and D: Number of trees and depth of the decision trees in the RF classifier

**Table 4.** CLP vs. HC using the Clinica Noel DB with selected features

Classifier	ACC [%]	Sen [%]	Spe [%]	AUROC	Parameters
SVM	73.5 ± 5	84.8 ± 8	59.2 ± 11	0.78	$C = 10^2, \gamma = 10^{-5}$
RF	68.2 ± 12	79.0 ± 10	54.8 ± 19	0.79	D = 60, N = 60

ACC: accuracy in the test set, AUROC: Area under ROC curve. Sen: Sensitivity,

Spe: Specificity.

C and  $\gamma$ : complexity parameter and bandwidth of the kernel in the SVM classifier.

N and D: Number of trees and depth of the decision trees in the RF classifier

**Table 5.** CLP vs. HC using the merged DB with selected features

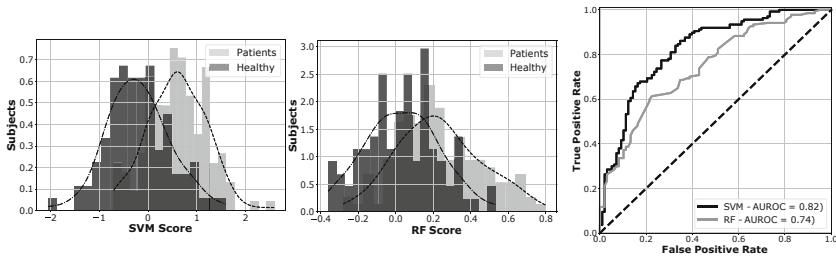
Classifier	ACC [%]	Sen [%]	Spe [%]	AUROC	Parameters
SVM	83.4 ± 2	91.2 ± 5	72.4 ± 9	0.90	$C = 10, \gamma = 10^{-3}$
RF	80.0 ± 4	92.8 ± 2	62.2 ± 9	0.89	D = 2, N = 70

ACC: accuracy in the test set, AUROC: Area under ROC curve. Sen: Sensitivity,

Spe: Specificity.

C and  $\gamma$ : complexity parameter and bandwidth of the kernel in the SVM classifier.

N and D: Number of trees and depth of the decision trees in the RF classifier



**Fig. 4.** Histograms of the scores and ROC curves (merged DB)

## 5 Conclusions

The proposed approach, based on articulation measures is effective for the classification of hypernasality in children. High accuracies were obtained with the SVM classifier (above 90%) in the two databases. When Kruskal-Wallis tests are applied as the selection criterion to include features before merging the two databases, accuracies of around 80.0% are obtained. The scores obtained with the Clínica Noel DB show less separability between CLP patients and HC patients. Conversely, the results obtained with the Manizales DB are higher and the associated histograms show less overlapping between the two classes. This can be likely explained due to the difference in the acoustic conditions of both databases. The sensitivity of the model was consistently high along the three experiments presented in this paper. This result may indicate that the proposed approach is suitable to evaluate the degree of hypernasality. Further research with more data and additional labels are required to confirm this hypothesis. Our team is currently working on the collection of more speech samples such that allow the evaluation of different degrees of hypernasality considering sustained vowel phonations and continuous speech signals.

**Acknowledgement.** This work was partially funded by CODI at UdeA grant # PRG2018-23541 and SOS18-2-01\_ES84180137.

## References

1. World Health Organization: Global registry and database on craniofacial anomalies. Report of a WHO registry meeting on craniofacial anomalies (2001)
2. Golabbakhsh, M., et al.: Automatic identification of hypernasality in normal and cleft lip and palate patients with acoustic analysis of speech. *J. Acoust. Soc. Am.* **141**(2), 929–935 (2017)
3. Kummer, A.W.: *Cleft Lip, Palate and Craniofacial Anomalies*. CENGAGE Learning, Boston (2014)
4. He, L., Zhang, J., Liu, Q., Yin, H., Lech, M., Huang, Y.: Automatic evaluation of hypernasality based on a cleft palate speech database. *J. Med. Syst.* **39**, 61 (2015)
5. Rendón, S.M., Orozco Arroyave, J.R., Vargas Bonilla, J.F., Arias Londoño, J.D., Castellanos Domínguez, C.G.: Automatic detection of hypernasality in children. In: Ferrández, J.M., Álvarez Sánchez, J.R., de la Paz, F., Toledo, F.J. (eds.) *IWINAC 2011*. LNCS, vol. 6687, pp. 167–174. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-21326-7\\_19](https://doi.org/10.1007/978-3-642-21326-7_19)
6. Dubey, A.K., Tripathi, A., Prasanna, S.R.M., Dandapat, S.: Detection of hypernasality based on vowel space area. *J. Acoust. Soc. Am.* **143**(5), 412–417 (2018)
7. Orozco Arroyave, J.R., Arias Londoño, J.D., Vargas Bonilla, J.F., Nöth, E.: Automatic detection of hypernasal speech signals using nonlinear and entropy measurements. In: *Proceedings of INTERSPEECH*, pp. 2027–2030 (2012)
8. Orozco Arroyave, J.R., Vargas Bonilla, J.F., Arias Londoño, J.D., Murillo Rendón, S., Castellanos Domínguez, C.G., Garcés, J.F.: Nonlinear dynamics for hypernasality detection in Spanish vowels and words. *Cogn. Comput.* **5**(4), 448–457 (2012)

9. Carvajal Castaño, H.A.: Metodología para la reducción de ruido aditivo de fondo en sistemas basados en procesamiento de voz. Master's thesis, Universidad de Antioquia (2013)
10. Godino Llorente, J.I., Gómez Vilda, P.: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *IEEE Trans. Biomed. Eng.* **51**(2), 380–384 (2004)
11. Orozco Arroyave, J.R., Hönl, F., Arias Londoño, J.D., Vargas Bonilla, J.F., Nöth, E.: Spectral and cepstral analyses for Parkinson's disease detection in Spanish vowels and words. *Expert Syst.* **32**(6), 688–697 (2015)
12. Rabiner, L., Schafer, R.W.: *Theory and Applications of Digital Speech Processing*. Prentice Hall, Upper Saddle River (2011)
13. Kaiser, J.F.: On a simple algorithm to calculate the 'energy' of a signal. In: *Proceedings of ICASSP*, pp. 381–384 (1990)
14. Cairns, D., Hansen, J., Riski, J.: A noninvasive technique for detecting hypernasal speech using a nonlinear operator. *IEEE Trans. Biomed. Eng.* **43**(1), 35–45 (1996)
15. Sapir, S., Ramig, L.O., Spielman, J.L., Fox, C.: Formant centralization ratio (FCR): a proposal for a new acoustic measure of dysarthric speech. *J. Speech Lang. Hear. Res.* **53**(1), 114–125 (2010)