

Acoustic Feature Analysis for Hypernasality Detection in Children

Genaro Daza

Universidad Nacional de Colombia Sede Manizales, Colombia

Luis Gonzalo Sánchez

Universidad Nacional de Colombia Sede Manizales, Colombia

Franklin A. Sepúlveda

Universidad Nacional de Colombia Sede Manizales, Colombia

Germán Castellanos D.

Universidad Nacional de Colombia Sede Manizales, Colombia

INTRODUCTION

In the treatment of children with fixed Cleft Lip and Palate (CLP), problems such as hipponasality and hypernasality, which are related to vocal emission and resonance, might appear. Nevertheless, according to the report presented in Castellanos et al. (2006), hypernasality is more frequently found than hipponasality (90% vs. 10%). The interest shown in hypernasality detection is that its occurrence points out problems of anatomical and neurological sort that are also related to the peripheral nervous system (Cairns, Hansen, & Riski, 1996). The presence of hypernasality, understood as the leak of nasal air and compensatory articulations, leads to low intelligibility of speech. This declining of the subject's communication capabilities may end up in behavioral and social changes.

In velopharyngeal learning, the distortion of the acoustic production leads up to nasalized voice. Moreover, since air loss or nasal leak is massive, articulator mechanisms are compromised. The patient can not speak clearly and intelligibly, and thence they replace their velum palatine sphincter by glottal articulation that allows for clearer speech: /p/, /t/, /k/, /b/, /d/, /g/ come from glottal stops, while sounds like /ch/, /s/, /t/, /j/ are accompanied by hoarseness (Habbaby, 2002). Although hard palate has been repaired surgically, it might not provide velopharyngeal competence necessary for a normal speech production. Even if the palate is potentially capable after surgery, previous speech habits might have developed compensatory articulations or physiologic compensations that aimed to approximate to intelligibility that enhance the number of patho-

logical patterns in speech. As a result, compensatory articulations persist generally, even after undergoing post technical or post surgical manipulation that had forecasted a plenty shutting. Thus, they have to be fixed before increasing the performance of the velopharyngeal sphincter throughout language therapy.

In the last years, there has been a growing interest for acoustic speech analysis (ASA) as an alternative method for diagnosis and treatment in identifying functional disorders in children's voices (González, Cervera, & Miralles, 2002; Niedzielska, 2000; Niedzielska, Glijer, & Niedzielski, 2001). This type of analysis exposes great advantages over traditional methods due to its noninvasive nature and potential to obtain a quantitative measure on the clinic state of the larynx and vocal tract.

Acoustic features or objective parameters are frequently used to represent the pathologic voice on held vowels (Hadjitodorov & P.Mitev, 2002; Kent, Vorperian, & Kent, 200; Yu, Ouaknine, Revis, & Giovanni, 2001). However, such vectors are limited in their robustness because of their estimation complexity in real conditions with perturbations of nonstationary structure (Gupta & Gilbert, 2001). Although several analysis of effectiveness have been made around the different kinds of proposed features for objective evaluation of speech disorders (Frohlich & Michaelis, 2000; Yu et al., 2001), they can not be taken as a standard set of parameters for hypernasality identification because each disorder affects differently diverse aspects of speech emission.

The present work analyzes the statistical effectiveness of different acoustic features in the automatic

identification of hypernasality. Acoustic features reflect part of information contained in perceptual analysis; in part, due to their estimation is derived directly or indirectly from the vocal cords behavior. Consequently, it is convenient to apply multivariate analysis techniques in determining the effectiveness of voice features. The effectiveness is studied by using multivariate analysis techniques that are meant for feature extraction and feature selection, as well (latent variable models, heuristic search algorithms).

BACKGROUND

Nasalization and Nasal Emission

Nasalization is defined as the link between nasal cavity and the rest of the vocal tract; while nasal emission refers to abnormal air loss through nasal route. This abnormal leakage reduces intra-oral pressure causing distortion in consonants. When air loss turns into an audible reblowing, the nasal emission is more obstructive and speech is seriously affected. Nasality commonly named hypernasality refers to low speech quality, which results from inappropriate adding of the resonance system to vocal tract. Conversely to nasal emission, nasality does not involve large flows of nasal air, so that there is no significant change in intra-oral air pressure. For this pathology, identification studies based on signal modeling (specialized diagnosis) can be related to acoustic features by using pattern recognition techniques.

Nasalization Model

Considering that normal voice is made of resonances at different frequency formants F_k , the following acoustic model (Cairns et al., 1996) is proposed:

$$S_n(\omega) = \sum_{k=1}^K F_k(\omega)$$

In contrast to normal voice, nasalized voice is the appearance of the antiformants \hat{F}' and nasal formants \hat{F} :

$$S_h(\omega) = \sum_{k=1}^K F_k(\omega) - \sum_{l=1}^L \hat{F}'_l(\omega) + \sum_{m=1}^M \hat{F}_m(\omega)$$

It has been suggested that the intensity in the reduction of the first formant is a primary indicator of nasality. In Cairns et al. (1996), the superior formants are filtered in such a way that filtered normal voices will have one component, while nasalized voices will correspond to a signal with several components, which can be estimated using Teager's instant power operator.

In a general way, patients with CLP manifest nasality with a deficiency of the velopharyngeal port, coupling their nasal cavity to vocalic sounds that gives place to appearance of an additional resonance in the amplitude-frequency feature of the vocal tract, and so a noticeable decrease in the formants F1 and F2 (Baken, 1996) as shown in sections (a) and (b) of Figure 1.

ACOUSTIC FEATURES AND MULTIVARIATE ANALYSIS

Acoustic features can be split into two categories according to the acoustic properties to be measured. Based in additive noise, among them: Harmonic to Noise Ratio (HNR), Normalized Noise Energy (NNE), Glottal Noise Excitation (GNE), defined as the noise estimation and it is based in the assumption that resulting glottal pulses from collisions of vocal folds head to a synchronous excitation of the different band frequencies, Normalized Error prediction (NEP) that can be expressed as the relationship between geometric and arithmetic means of spectral model, and Turbulence Noise Index (TNI).

Other acoustic features are associated to frequency modulation noise, among them pitch or fundamental period of the signal and jitter, which is defined as the average variation percentage between two consecutive values of pitch. In addition, there are considered features associated to parametric models of speech generation. Among them: cepstral coefficients derived from linear prediction analysis (LPC Linear Prediction Coefficients), cepstral coefficients over pounded frequencies scale (MFCC Mel-Frequency Cepstrum Coefficients), and RASTA coefficients (Relative Spectral Transform) (Castellanos, Castrillón, & Guijarro, 2004).

Multivariate Analysis

Latent Variable Models. The set X of measured variables can be seen as the linear combination of generative factors Z and perturbations ϵ over the measurement process.

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/acoustic-feature-analysis-hypernasality-detection/12917

Related Content

Making Integrated eCare a Reality in the UK: Past Failures, Current Successes, and Future Challenges

Mark Gretton (2014). *Achieving Effective Integrated E-Care Beyond the Silos* (pp. 227-239).

www.irma-international.org/chapter/making-integrated-ecare-a-reality-in-the-uk/111385

Adoption of Mobile Devices in the Australian Healthcare: A Conceptual Framework Approach

Vasundhara Rani Sood, Raj Gururajan, Abdul Hafeez-Baigand Nilmini Wickramasinghe (2017). *Handbook of Research on Healthcare Administration and Management* (pp. 662-685).

www.irma-international.org/chapter/adoption-of-mobile-devices-in-the-australian-healthcare/163861

An Intelligent Wearable Platform for Real Time Pilot's Health Telemonitoring

Christos Papadelis, Chrysoula Kourtidou-Papadeli, Fotini Lazaridouand Eleni Perantoni (2008).

Encyclopedia of Healthcare Information Systems (pp. 777-784).

www.irma-international.org/chapter/intelligent-wearable-platform-real-time/13012

Rescheduling Dental Care with GSM-Based Text Messages

Reima Suomianand Ari Serkkola (2007). *Web Mobile-Based Applications for Healthcare Management* (pp. 335-353).

www.irma-international.org/chapter/rescheduling-dental-care-gsm-based/31164

COVID-19 Outbreak Prediction and Analysis of E-Healthcare Data Using Random Forest Algorithms

Debabrata Dansana, Raghvendra Kumar, Aishik Bhattacharjeeand Chandrakanta Mahanty (2022).

International Journal of Reliable and Quality E-Healthcare (pp. 1-13).

www.irma-international.org/article/covid-outbreak-prediction-analysis-healthcare/297075