Cepstral Analysis of Voice in Children With Velopharyngeal Insufficiency After Cleft Palate Surgery

Zengjie Yang, Jincai Fan, Jia Tian, Liqiang Liu, Cheng Gan, Wenlin Chen, and Zhuming Yin, *Beijing, China*

Summary: Objectives. The purpose of this study is to survey the cepstral peak prominence (CPP) of vowel sounds and to compare objective data obtained from patients with velopharyngeal insufficiency after cleft palate surgery with objective data of those with normal healthy controls using acoustic analyzer.

Methods. Participants were divided into a clinical group and a control group. Every participant was recorded phonating the sustained vowel /a/. Each participant in the clinical group was recorded before surgery, before and after speech therapy. All samples were subjected to acoustic analysis using *Praat* software. The vowels were analyzed acoustically by the measurement of smoothed cepstral peak prominence (CPPs).

Results. The results reveal lower values of CPPs in speakers with velopharyngeal insufficiency before and after the operation. And, the results also reveal that there is no significant difference across the control and the clinical groups after speech therapy.

Conclusions. The results reveal lower values of CPPs in the clinical group before surgery and before speech therapy in comparison with the control group, which could be explained because of the body's compensation for the lack of normal intraoral pressure and habit of articulation.

Key Words: Cepstral analysis–Velopharyngeal insufficiency (VPI)–Cleft palate.

INTRODUCTION

A cleft palate is one of the most common congenital malformations in the maxillofacial region. It seriously affects oral and maxillofacial morphology and function, particularly speech. Velopharyngeal insufficiency (VPI) is a common problem resulting from cleft palate repair. Whenever there are structural anomalies in the cavities of the vocal tract, there is a risk for distortion in speech sound, production, and resonance. Therefore, a secondary surgery may be needed to treat VPI, especially because speech therapy is much more effective if it is begun after the normalization of the structure.[1,2]

Auditory-perceptual evaluation has been commonly used to assess the quality of voice and the severity of voice and speech disorders. This method is convenient in terms of expense, time, and complexity of the procedure; however, it entails a limitation of interobserver variability and relies on subjective judgment.

Using recent developments in digital technology and methods for processing speech, it is possible to make effective and objective acoustic diagnostics for the assessment and management of voice disorders. Voice characteristics can now be objectively analyzed with numeral values using an acoustic analyzer. Acoustic analysis programs are able to generate data related to speech disorders of the vocal folds easily and noninvasively in a relatively short time and in a sensitive, objective, and quantitative manner.[3]

Patients with cleft palate often exhibit nasality, which is a distinctive feature and an important target in speech therapy. The voice and speech of patients with cleft palate have been studied using many techniques including spectral, perturbation, and formant analyses. There are several spectral characteristics associated with the hypernasality of patients with a cleft palate, such as a reduction of the first formant (F1), the presence of extra resonance, and a decrease of second formant (F2).[4] Several indexes based on voice spectral analysis have been proposed to evaluate hypernasality, such as the difference between the amplitude of F1 and the amplitude of the extra peak,[5] the difference between the amplitude of F1 and the amplitude of the first nasal peak,[6] and one-third octave spectral analysis.[4] However, there are some limitations of these indexes. These include difficulty detecting the nasal peak when it appears in the vicinity of a formant, variability of the nasal peak when the fundamental frequency is high, and the influence of vocal breathiness.[4] These indexes are now not widely used clinically. Zajac and Linville[7] reported that the cleft palate speakers have larger frequency perturbations (jitter) than normal controls. However, the methods used to calculate perturbations, jitter, and shimmer are only reliable for nearly periodic voice signals and cannot reliably analyze strongly aperiodic signals. And, these measures rely on the ability to accurately identify and track changes in fundamental frequency.

Can we find an ideal acoustic measure to quantify the voice signal independently, without relying on frequency tracking or other variables that may influence the accuracy of the measure? Cepstral peak prominence (CPP) should be one such measure.

Cepstrum is a Fourier transformation of a spectrum. A normal voice with well-defined harmonic structure has a strong cepstral peak compared with the breathy and hoarse voice of a poorly defined harmonic structure.[8] CPP and smoothed CPP

Accepted for publication March 20, 2014.
Disclosure: The authors do not have a financial interest in any products, devices, or drugs mentioned in this article.
From the Ninth Department of Plastic Surgery, Plastic Surgery Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.
Address correspondence and reprint requests to Jincai Fan, Ninth Department of Plastic Surgery, Plastic Surgery Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China. E-mail: free063@yahoo.com
Journal of Voice, Vol. 28, No. 6, pp. 789-792
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.03.015

have been proven to be the best predictors of many vocal diseases and conditions,[9,10] including vocal nodules,[11] adductor vocal fold palsy,[12] and so on. Few articles, however, have been found to report the use of CPP and smoothed CPP in the patients with cleft palate.

The purpose of this study was to survey the smoothed CPP of vowel sounds and compare, using an acoustic analyzer, objective data obtained from patients with VPI after surgery to fix a cleft palate with the objective data of those patients with normal voice control. The assessment of pathological speech is an approach that interests both physicians and speech pathologists.

MATERIALS AND METHODS

The study followed a prospective case-control design.

Participants

Participants were divided into a clinical group and a control group. Every participant was recorded phonating the sustained vowel /a/. Each participant in the clinical group was recorded before surgery, before speech therapy (3–4 months after surgery), and after speech therapy immediately. All samples were subjected to acoustic analysis using *Praat* software *version 4.1.2* (developed by Boersma and Weenink of the University of Amsterdam, the Netherlands). The vowels were analyzed acoustically by measuring CPPs.

The clinical group comprised 20 boys (mean age, 7.1 years; standard deviation [SD], 1.1) and 20 girls (mean age, 7.4 years; SD, 1.3), ranging in age from 6 to 9 years, all of whom experienced VPI after cleft palate surgery. All participants in the clinical group were aged younger than 3 years when they underwent the surgery, a two-flap palatoplasty. There is no fistula in all the palates after surgery. VPI was diagnosed by an experienced laryngologist using an endoscopy procedure when the children were in rest position. The medical treatment was not initiated in any of the participants at the time of voice recording. All the patients in the clinical group presented with various voice disorders. No fistula or dehiscence occurred in any of the patients.

All VPI patients in the clinical group then underwent surgery after the data collection and accepted speech therapy 3–4 months after surgery. All the patients in the clinical group completed the phonating record at each of the three data collection points.

The basic steps of speech therapy are as follows. (1) Determine which phonemes to target first based on stimulability testing and sounds that will have the biggest impact on intelligibility. (2) In general, the speech begins with anterior sound because they are most visible. For continuant sounds, start with the voiceless cognate and then voicing. (3) For plosives, the voiced cognate is easier in some cases. Start with the sound in isolation for continuants or in a consonant-vowel syllable for plosives. (4) Establish correct placement first and then correct manner (including voicing) of production. (5) When transitioning from one sound to the next sound, change only one feature at a time. (6) Teach parents how to work on speech at home with several short practice sessions several times a day, for example 10–15 minutes one time. The speech therapy is performed three to four times in 1 week, 30–45 minutes per time. And, all the speech therapy procedures last for about 4–6 weeks.

The control group comprised 20 boys (mean age, 7.4 years; SD, 1.2) and 20 girls (mean age, 6.9 years; SD, 1.3), ranging in age from 6 to 9 years. All participants had normal voices as evaluated perceptually by three trained speech pathologists. None of them had a history of vocal abuse or misuse. Informed consent was obtained from all parents of the participants, and the institutional ethics committee approved the protocol. All participants in the control group reported no history of speech or hearing anomalies.

All the participants could cooperate the study well. And, they all reported Chinese as their primary language.

Experimental procedure

Each participant was instructed to vocalize extended /a/ vowels three times for at least 5 seconds at a comfortable voice level and tone. All vocalizations were recorded with a microphone (YM-63F; YAMAHA Corporation, Shizuoka-ken, Japan) and *Cool Edit Program* (Version 2.1; Adobe Systems Corporation, San Jose, CA) software in a soundproof room, and the distance of the microphone from the mouth was 10 cm.

Analysis

The vowels analysis of recorded aural signals was performed using the relatively stable vocalization period of the middle 2 seconds, excluding the initial and ending periods. Acoustic analysis was performed using the aforementioned *Praat* software. All vowel samples were digitized on a computer at a 22-kHz sampling rate with 16 bits of resolution. The voice samples were subjected to acoustical analysis using *Praat* software, and Fourier transformation of an acoustic signal was first performed to create spectrum. Cepstrum was arrived at by performing a Fourier transformation of the spectrum. Cepstral parameters were extracted for all voice samples considered for the study.

RESULTS

The present study analyzed the CPPs in patients with cleft palate. Descriptive statistics was used to obtain mean and SD, which are listed in Table 1. The analysis of variance was used in the study.

TABLE 1.
Mean and SD of the CPPs in the Clinical (Before Surgery) and the Control Groups

	Boys		Girls	
Groups	Mean	SD	Mean	SD
Clinical	15.27	1.98	16.31	2.01
Control	23.61	2.36	25.01	1.22