Voice Perturbations of Children With Perceived Nasality and Hoarseness

David J. Zajac, M.A.
Raymond N. Linville, Ph.D.

Cycle-to-cycle variations in voice fundamental frequency (jitter) and amplitude (shimmer) were derived by electroglottography for 10 children with velopharyngeal insufficiency (VPI). Jitter was found to be positively correlated with ratings of perceived nasality, whereas shimmer was found to be positively correlated with ratings of perceived hoarseness. Theoretic implications for a regulatory model of speech aerodynamics are discussed. Additionally, clinical applications, in terms of using electroglottography as a supplemental assessment procedure, are suggested.

KEY WORDS: *voice perturbations, electroglottography, speech pathology*

The presence of voice disorders in individuals with cleft palate has been previously reported (McDonald and Koepp-Baker, 1951; Westlake, 1953; Hess, 1959; Brooks and Shelton, 1963; Bzoch, 1964; McWilliams et al, 1969; Marks et al, 1971; D'Antonio et al, 1988). Vocal problems such as hoarseness (both with and without vocal cord pathology), breathiness, reduced loudness, deviant pitch, restricted pitch range, and tense-strained vocal quality have been observed.

D'Antonio et al (1988) reported that 41 percent of a group of individuals with velopharyngeal insufficiency exhibited "laryngeal abnormalities" and/or abnormal vocal quality. Their sample (n = 85) consisted of 42 individuals with cleft palate (49 percent), 10 with submucous cleft palate (12 percent), and 33 without clefts (39 percent). The prevalence of voice disorders was not significantly different between individuals with clefts and those without clefts. D'Antonio et al (1988) did note a significant relationship between the presence of vocal findings and increased subglottal pressure, although a cause-and-effect relationship was not established.

McWilliams et al (1969) reported that bilateral vocal cord nodules occurred in 72 percent of a group of 32 children with both cleft palate and chronic hoarseness. The majority of these children were found to have "borderline" velopharyngeal competence. McWilliams et al (1969) suggested that the children may have had laryngeal compensations other than glottal stops that contributed to the development of their vocal pathology. McWilliams et al (1969) concluded that a "logical" (although unspecified) relationship existed between VPI and vocal fold nodules.

Curtis (1968) suggested that individuals with cleft palate may need to exert greater respiratory effort to achieve normal intensity level because of acoustic damping in the nasal tract. Analog model studies (House and Stevens, 1956) and in vivo studies (Bernthal and Beukelman, 1977) have demonstrated that the overall sound energy of vowels is reduced as a consequence of oronasal coupling. Warren et al (1969) reported that speakers with VPI expend twice the normal volume of air during speech production. Warren et al (1988) have further suggested that individuals with VPI may increase respiratory effort as a way to develop adequate intraoral air pressures. However, the increased respiratory effort may contribute to vocal abuse.

Hamlet (1973) reported that the vibratory characteristics of the vocal cords were altered when nasality was present. Specifically, Hamlet (1973) found that the opening phase of the glottal cycle was reduced for nasalized vowels compared with nonnasalized vowels, even when the vowels were matched for intensity. Hamlet interpreted these results to indicate that nasalization increased the "force" of vocal cord adduction independent of the level of vocal effort. Leder and Lerman (1985) provide additional evidence for the effects of VPI on laryngeal function. They reported that adults with clefts and "clinically significant hypernasality" demonstrated inappropriate vocal cord adduction and voicing during the production of voiceless stop plosives. Leder and Lerman (1985) speculated that phonation was facilitated by transglottal pressure changes that resulted from inadequate velopharyngeal function. They further suggested that inappropriate voicing may serve to reduce nasal air emission.

The purpose of the present study was to investigate voice perturbations of children with perceived nasality and hoarseness. Vocal perturbations are usually defined as the cycle-to-cycle variation in fundamental frequency (jitter) and amplitude (shimmer). These measures have been shown to reflect the regularity of vocal cord oscillations. Specifically, the primary purpose of the study was to determine the

David Zajac is Graduate Research Assistant, Cleft Palate–Craniofacial Center, University of Pittsburgh, Pittsburgh, PA. Dr. Linville is Assistant Professor, Department of Speech and Hearing Sciences, Texas Tech University, Lubbock, TX.

Presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA, 1987.

This study was supported in part by NIH research grant DE-01696 and by the Pittsburgh Eye and Ear Foundation.

Reprint requests: David Zajac, M.A., Graduate Research Assistant, Cleft Palate–Craniofacial Center, University of Pittsburgh, 313 Salk Hall, Pittsburgh, PA 15261.

relationship between vocal perturbations and perceived nasality in children diagnosed with VPI, regardless of etiology. If a definite relationship was found, this would be further evidence for an interaction between laryngeal and velopharyngeal events. A secondary purpose of the study was to determine the relationship between vocal perturbations and perceived hoarseness exhibited by the children. Previous research has indicated that perturbation measures are correlated positively with perceived vocal dysfunction in adult speakers.

METHOD

Subjects

The subjects for this study were 10 children with VPI and five normal children. The children with VPI were patients at the University of Pittsburgh Cleft Palate–Craniofacial Center (Table 1). They had been diagnosed as having at least borderline velopharyngeal incompetence within 1 year prior to the initiation of the study. They ranged in age from 6 years, 8 months to 11 years, 10 months. The mean age was 9 years, 0 months. Additional selection criteria for subjects with VPI included (1) that they be free of vocal pathology, upper respiratory infections, and middle ear infections at the time of the study and (2) that they be free of any additional craniofacial syndromes. All subjects with VPI had clefts of the palate or of the lip and palate, except for one subject (FJ), who developed VPI subsequent to adenotonsillectomy at age 4 years. Eight of the 10 subjects had surgical repair of the palate; two subjects also had pharyngeal flap procedures. All subjects with VPI were free of oronasal fistulas and exhibited normal articulation. Finally, audiograms of all subjects with VPI revealed hearing to be within normal limits.

The five normal children served as controls. They consisted of three boys and two girls who ranged in age from 7 years, 7 months to 12 years, 2 months of age, with a mean age of 9 years, 2 months. They were free of any known craniofacial abnormality and were judged to exhibit normal phonation, articulation, and resonance.

Speech Sample

Subjects produced the steady state vowels /i/, /a/, and /u/ for approximately 3 seconds at a comfortable loudness level. The sentences "Mama made lemon jam," "Sissy sees the sky," "Put the baby in the buggy," "Go get a cookie for Kate," and counting from 60 to 65 were also produced.

TABLE 1 Gender, Age, and Palatal Condition of Subjects With Velopharyngeal Insufficiency

Subject	*Sex*	*Age (Years/Months)*	*Palatal Condition*
KD	M	6–8	Unilateral cleft
MP*	M	7–0	Cleft of soft palate
GN	F	7–1	Cleft of soft palate
PD	M	7–5	Bilateral cleft
JK	M	7–8	Unilateral cleft
MD	M	9–8	Cleft of soft palate
PA†	M	10–0	Submucous cleft
SA	F	10–10	Bilateral cleft
FJ*	M	11–8	No cleft (acquired VPI)
ML	M	11–10	Unilateral cleft

* Subject had additional pharyngeal flap surgery.
† Palate not repaired because of cardiovascular disorder.

Instrumentation

A Kay Elemetrics Laryngograph monitored vocal cord contact. The laryngograph or electroglottograph (EGG) is a noninvasive device that represents vocal cord contact by monitoring transglottal impedance variations detected by electrodes placed on the surface of the neck. An Omnidyne microphone positioned 7 inches from the subject's mouth transduced the audio signals. Both the audio and EGG signals were displayed on a Tektronix dual-channel storage oscilloscope (model T912) and were recorded on a Teac four-channel taperecorder (model A-3440), at 15 inches per second using Scotch 226 audio recording tape. For data analysis, the EGG signals were low pass filtered at a cutoff frequency of 5 kHz using a Krohn-Hite filter (model 3200), which was digitized at a rate of 20 kHz with 12-bit resolution using a Data Translation A/D converter (model DT-2821) and were then stored in computer memory (Fig. 1). Perturbation analysis was performed using an IBM-AT computer and Interactive Laboratory Systems (ILS) software, version 6.0 (Signal Technology, Inc.).

Procedures

Subjects were seated in a sound attenuated booth, and the EGG electrodes were positioned superficially on each side of the thyroid laminae. A velcro fastener secured the electrodes in position. Subjects were instructed to count to 10 using a comfortable loudness level while the EGG signal was monitored on the oscilloscope.

The speech samples were presented to the subjects both visually and verbally. The subjects first produced the vowels followed by the sentences. The order of each utterance was randomized. Two repetitions of each vowel were produced. This sequence was repeated a second time, which

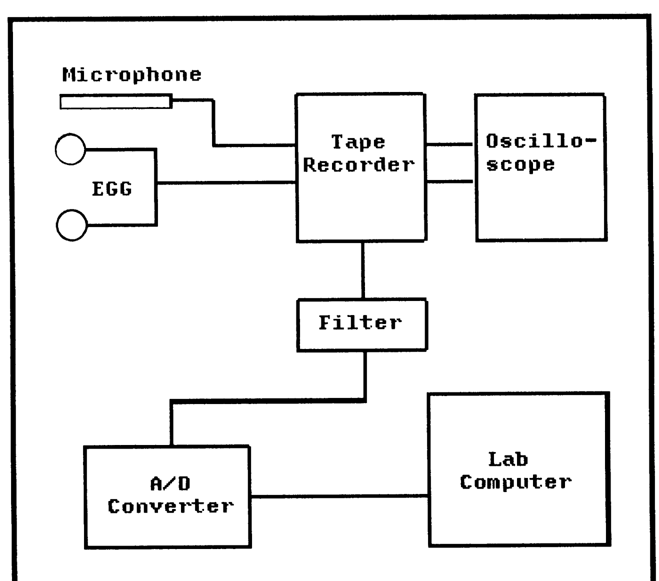


FIGURE 1 Signal processing diagram.

resulted in a total of 12 vowels and 10 sentences produced by each subject.

Data Analysis

A computer program using a zero crossing algorithm, developed at the University of Pittsburgh Speech Science Laboratory, was used to detect peak-to-peak amplitude and period values from the EGG signal. Jitter, shimmer, and fundamental frequency were determined from 50 consecutive glottal cycles taken from the midportion of each vowel. The computational formulas for jitter and shimmer are shown in Figure 2. Jitter and shimmer were calculated as ratios of the sum of the absolute cycle-to-cycle variability to the mean frequency (jitter) and amplitude (shimmer) respectively, multiplied by 100.

To calibrate the analysis system, 100, 200, and 300 Hz sinusoidal waves generated by a B&K Dual Channel Frequency Analyzer (Type 2032) were recorded and digitized. Mean shimmer was 0.37 percent, 0.27 percent, and 0.30 percent, respectively. Mean jitter was 0.17 percent, 0.15 percent, and 1.02 percent, respectively. The jitter values are similar to those reported by Titze et al (1987) for a sampling rate of 20 KHz without interpolation (digitized directly).

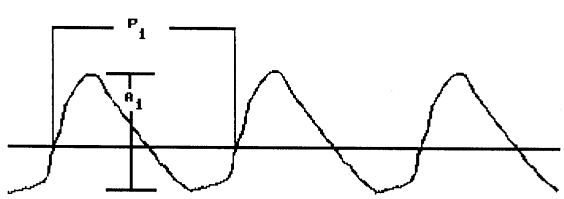
Perceptual Ratings

Three speech samples were selected for each subject. Counting, "Mama made lemon jam," and one additional nonnasal sentence were selected at random. These were recorded onto audiotape in random order. Five of these sentences were selected at random and repeated a second time to check reliability.

Three experienced speech-language pathologists served as judges. All judges passed a pure tone hearing screening at 25 dB HL. They were instructed to rate the sentences for perceived nasality using a seven-point, equal interval scale; lower numbers represented lesser degrees of perceived nasality. Nasality was not defined for the judges. Twelve sample sentences that represented various degrees of nasality from mild to severe were presented first to the judges. This was done to allow each judge to determine his or her own end point of the rating scale. The audiotape containing the randomized speech samples was then presented to the judges for rating. They marked their responses on prepared forms. Approximately 15 seconds separated each speech sample.

Following the nasality rating procedure, similar procedures were employed for rating hoarseness. A five-point, equal interval scale was used for these ratings. Fewer intervals were selected for the hoarseness scale to allow comparisons to previous studies, which employed four- or five-point rating scales (Haji et al, 1986).

Interjudge and intrajudge agreements were determined. For nasality, the three judges averaged 93 percent interjudge agreement within two scale points (100 percent between judges A and B, 93 percent between judges A and C, and 87 percent between judges B and C). All three judges obtained 100 percent intrajudge agreement within two scale points. The three judges averaged 80 percent interjudge agreement within one scale point for hoarseness (100 percent between judges A and B, 73 percent between judges A and C, and 67 percent between judges B and C). Both judges A and B obtained 100 percent intrajudge agreement, while judge C obtained 80 percent intrajudge agreement (within one scale point).



$$\%\ \text{Jitter} = \frac{\Sigma |P_i - P_{i+1}|}{N - 1} \Big/ \frac{\Sigma P_i}{N} \times 100$$

$$\%\ \text{Shimmer} = \frac{\Sigma |A_i - A_{i+1}|}{N - 1} \Big/ \frac{\Sigma A_i}{N} \times 100$$

Where: P_i = period of i^{th} cycle

A_i = amplitude of i^{th} cycle

N = number of cycles

FIGURE 2 Computational formulas for jitter and shimmer. Increasing amplitude represents decreasing glottal impedance, i.e., increasing vocal fold contact. Jitter and shimmer are expressed as ratios of the sum of the absolute cycle-to-cycle differences to the mean frequency and amplitude, respectively, multiplied by 100.

RESULTS

Table 2 shows the means, standard deviations, and ranges of fundamental frequency (f_o), jitter, shimmer, and perceptual ratings for the subjects with VPI. Mean f_o for all subjects was 253 Hz (SD = 25). The mean f_o range was 225 Hz to 297 Hz. Mean jitter for all subjects was 1.61 percent (SD = 0.36). Mean jitter ranged from 1.01 to 2.26 percent. Mean shimmer for all subjects was 2.27 percent (SD = 0.83). Mean shimmer ranged from 1.05 to 3.70 percent.

The mean rating for perceived nasality for subjects with VPI was 2.04 scale points (SD = 1.42). Mean nasality ranged from 1.22 to 6.00 scale points. The mean rating for perceived hoarseness for subjects with VPI was 1.90 scale points (SD = 0.56). Mean hoarseness ranged from 1.11 to 2.89 scale points. The relatively small group nasality rating probably reflects the constraint of long-term treatment of the subjects. The majority of the subjects had "borderline" VPI and exhibited relatively little hypernasality, as was the

TABLE 2 Means, Standard Deviations (SDS), and Ranges of Fundamental Frequency (f_o), Jitter, Shimmer, Nasality, and Hoarseness for Subjects With VPI (n=10)

	Mean	*SD*	*Range*
f_o (Hz)	253.00	25.00	225.00–297.00
Jitter (%)	1.61	0.36	1.01–2.26
Shimmer (%)	2.27	0.83	1.05–3.70
Nasality (scale pts)	2.04	1.42	1.22–6.00
Hoarseness (scale pts)	1.90	0.56	1.11–2.89

case in the McWilliams et al (1969) study. The relatively small group hoarseness rating may reflect the constraint of the study's selection criteria (i.e., subjects free of vocal pathology at the time of the study).

Table 3 shows the means, standard deviations, and ranges of f_o, jitter, and shimmer for the control subjects. Mean f_o for all subjects was 258 Hz (SD = 45), with a mean f_o range of 207 to 328 Hz. Mean jitter for all subjects was 1.20 percent (SD = 0.34), with a mean jitter range of 0.73 to 1.68 percent. Mean shimmer for all subjects was 1.66 percent (SD = 0.62), with a mean shimmer range of 0.85 to 2.58 percent. Mann-Whitney U-Tests revealed that jitter was significantly different between groups (0.05 level), while the difference for shimmer was nonsignificant.

Spearman rank order correlation coefficients were computed between the perturbation measures and the perceptual ratings for subjects with VPI. All coefficients were positive, thereby indicating that both fundamental frequency and amplitude instability tended to increase as the perception of nasality and hoarseness increased (Table 4). The correlation between nasality and jitter was significant ($r_s = 0.63$, $p<0.05$, two-tail test). The correlation between hoarseness and shimmer also was significant ($r_s = 0.74$, $p<0.05$, two-tail test).

Table 4 also includes correlations between jitter, shimmer, nasality, hoarseness, age, and fundamental frequency for the subjects with VPI. The correlation between jitter and shimmer was positive and significant ($r_s = 0.71$, $p<0.05$, two-tail test). Both jitter and shimmer were correlated positively although nonsignificantly with age ($r_s = 0.32$, and $r_s = 0.58$, respectively). Fundamental frequency was correlated positively with jitter ($r_s = 0.20$) and negatively with shimmer ($r_s = -0.09$). Neither correlation was significant.

To determine the possible influence of the other variables on the observed relationships between jitter-nasality and shimmer-hoarseness, Kendall partial correlation coefficients were computed. The partial correlation coefficients between jitter and nasality were essentially the same when age, f_o, shimmer, and hoarseness were held constant. Likewise, the partial correlation coefficients between shimmer and hoarseness were the same when f_o, jitter, and nasality were held constant. The partial correlation with age, however, accounted for approximately 6 percent less of the variance than did the first-order correlation between shimmer and hoarseness. This indicated that age contributed slightly to the observed relationship.

DISCUSSION

The results show that voice perturbations of the children with VPI were correlated moderately with perceived nasality and hoarseness. Additionally, the results suggest that voice perturbations (at least jitter) of these children were significantly greater than those of the children without VPI. The positive relationship between jitter and perceived nasality provides additional evidence for a link between laryngeal and velopharyngeal events. Although causality cannot be determined from this study, several explanations for the relationship between jitter and perceived nasality seem plausible.

TABLE 3 Means, Standard Deviations (SDS), and Ranges of Fundamental Frequency (f_o), Jitter, and Shimmer for Control Subjects (n=5)

	Mean	*SD*	*Range*
f_o (Hz)	258.00	45.00	207.00–328.00
Jitter (%)	1.20	0.34	0.73–1.68
Shimmer (%)	1.66	0.62	0.85–2.58

TABLE 4 Spearman Rank-Order Correlation Coefficients Among All Variables for Subjects With VPI

	f_o	*Jitter*	*Shimmer*	*Nasality*	*Hoarse-ness*	*Age*
f_o	1.00	0.20	−0.09	0.19	−0.11	−0.01
Jitter		1.00	0.71*	0.63*	0.35	0.32
Shimmer			1.00	0.30	0.74*	0.58
Nasality				1.00	0.12	0.08
Hoarseness					1.00	0.44
Age						1.00

* $p<0.05$ (two-tail test)

Aerodynamic Regulation

It is known that some amount of voice perturbation is normal and may reflect random aerodynamic and neuromuscular events (Titze et al, 1987). The increased jitter levels of the children with VPI suggest that laryngeal aerodynamic and/or neuromuscular processes may be altered as a result of oronasal coupling. When the velopharyngeal port is open during vowel production, airflow has two alternative paths to follow (analogous to a parallel electric circuit). In a parallel circuit, the total current flow is greater than that in either branch (i.e., additive). A similar situation in the vocal tract would result in changes in glottal volume velocity (flow rate) and transglottal pressure changes as suggested by Leder and Lerman (1985). However, this situation assumes conditions of constant glottal resistance and respiratory effort without compensatory responses.

Individuals with VPI may attempt to regulate actively vocal tract resistances as a compensation (Warren, 1986). Increased glottal resistance during vowel production, for example, would decrease flow rate and facilitate regulation of subglottal pressure required to sustain phonation. Additionally, compensatory changes in chest wall dynamics may occur either to increase or to decrease subglottal pressure as needed. Therefore, attempts to regulate respiratory and laryngeal aerodynamic and neuromuscular processes when inappropriate oronasal coupling exists may contribute to increased vocal perturbations.

Intensity Levels

Intensity levels among subjects were not controlled. Research has indicated that children may produce greater intensity levels than adults when instructed to talk at a "comfortable loudness" level (Stathopoulos, 1986). Differences in intensity levels among the children may have affected jitter and shimmer. Zajac and Linville (1988), for example, reported increased jitter for adult speakers when phonating at greater than normal loudness levels. However, Glaze et al (1988) reported that acoustically derived voice perturbations of children decreased with increased loudness. These findings may be attributable to differences in laryngeal anatomy between children and adults.

The explanations for the relationship between perceived nasality and jitter require further experimental study. Also, it is possible that other mechanisms may account for the present findings or that the above explanations are not mutually exclusive.

The results of the present study also show a positive correlation between shimmer and perceived hoarseness. This finding has been reported previously for adult speakers with known vocal pathologies (Haji et al, 1986). That this relationship existed for children with VPI but with no known vocal cord pathology further supports this finding. For children with VPI who may be susceptible to the development of vocal pathology, this relationship may have important clinical implications. However, the results of both this study and that of Glaze et al (1988) suggest that vocal shimmer may be quite variable among children.

Clinical Significance

Several potential clinical applications are suggested by the results of this study. First, electroglottography may be a valuable screening procedure for vocal dysfunction in children with VPI. For example, children who exhibit relatively large shimmer values (with or without accompanying perceptual correlates) may be at risk for the development of vocal pathology. Based on the results of this study, shimmer values that exceed 3 percent (approximately one SD above the mean for the children with VPI) may identify such children. Electroglottographic monitoring of these children with possible intervention in terms of a prophylactic vocal reduction program may be indicated. Since electroglottography is a safe and noninvasive procedure, its usefulness as a screening and monitoring technique is enhanced. The establishment of a normative data base for children's voice perturbations is needed to facilitate such diagnostic decisions.

Electroglottography also may serve as a valuable adjunct to direct viewing techniques once vocal pathology is found. Anastaplo and Karnell (1988) have described a procedure that uses EGG simultaneously with videostroboscopic imaging of the vocal cords. This technique provides the potential for gaining greater information pertaining to vocal cord oscillations than from either method alone. Finally, treatment outcomes of either behavioral or physical vocal management techniques may be quantified by EGG and perturbation analysis.

A Clinical Case

The following is a description of the application of EGG as an adjunct assessment procedure to a clinical case. The main purpose is to illustrate how EGG used on a routine basis may facilitate clinical decisions pertaining to vocal function. The case, however, also provides data in support of a theoretic model for airway regulation.

The patient was an 11-year-old female with a repaired cleft of the palate, a pharyngeal flap, and a small anterior oronasal fistula thought to be nonsymptomatic for speech. Her speech, however, was judged to be mildly hypernasal. Vocal quality was judged to be normal. She was referred for aerodynamic evaluation to estimate the magnitude of nasal air emission. Nasal airflow measurements were conducted first with the fistula open and then with the fistula occluded with dental wax. These measurements (obtained during the repetition of the syllable /pi/) revealed that nasal airflow was substantially reduced when the fistula was occluded (69 ml per second compared with 135 ml per second when the fistula was unoccluded). This indicated that the fisula was indeed symptomatic for speech production. The treatment team decided to close the fistula and then reevaluate the patient.

Perturbation data additionally were collected during the evaluation of this patient. Mean jitter and shimmer with the fistula unoccluded were 2.1 and 4.0 percent respectively. When the fistula was occluded, mean jitter and shimmer were 1.7 and 2.1 percent respectively. The substantial reduction in vocal shimmer suggests the involvement of nasal airflow as a precipitating factor. It should be noted that the patient previously had been evaluated with aerodynamic and EGG instrumentation and that those results were consistent with the later findings when the fistula was unoccluded. Upon questioning, the patient revealed that she frequently engaged in vocally abusive behaviors with siblings in the home. Based on the relatively high shimmer value, it was recommended that she receive postsurgical voice monitoring and that a vocal abuse reduction program be considered.

Conclusions

The results of the present investigation indicate positive relationships between vocal perturbations and perceived nasality and hoarseness in children with VPI. The relationship between nasality and jitter suggests possible interactions between laryngeal and velopharyngeal aerodynamic and physiologic events. Further research may identify specific mechanisms of causality. Finally, as the present study utilized a small sample of children within a restricted age range, additional research is required to confirm the present findings.

Acknowledgements. The authors would like to thank Betty Jane McWilliams, Robert Skwarecki, and Julie Oleh for their assistance during this study.

References

Anastaplo S, Karnell M. (1988). Synchronized videostroboscopic and electroglottographic examination of glottal opening. J Acoust Soc Am 83:1883–1890.

Bernthal JE, Beukelman DR. (1977). The effect of changes in velopharyngeal orifice area on vocal intensity. Cleft Palate J 14:63–77.

Brooks A, Shelton R. (1963). Voice disorders other than nasality in cleft palate children. Cleft Palate J 13:63–64.

Bzoch K. (1964). The effects of a specific pharyngeal flap operation upon the speech of 40 cleft palate persons. J Speech Hear Dis 29:111–120.

Curtis J. (1968). Acoustics of speech production and nasalization. In: Spriesterbach C, Sherman D, eds. Cleft palate and communication. New York: Academic Press:27–60.

D'Antonio LL, Muntz HR, Province MA, Marsh JL. (1988). Laryngeal/voice findings in patients with velopharyngeal dysfunction. Laryngoscope 98:432–438.

Glaze LE, Bless DM, Milenkovic PH. (1988). Acoustic analysis of vowel and intensity differences in children's voice. Presented at the

Annual Convention of the American Speech-Language-Hearing Association, Boston, MA.
HAJI T, HORIGUCHI S, BAER T, GOULD W. (1986). Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. J Acoust Soc Am 80:58–62.
HAMLET S. (1973). Vocal compensation: an ultrasound study of vocal vibration in normal and nasal vowels. Cleft Palate J 10:267–285.
HESS D. (1959). Pitch, intensity, and cleft palate voice quality. J Speech Hear Res 2:113–125.
HOUSE A, STEVENS K. (1956). Analog studies of the nasalization of vowels. J Speech Hear Dis 21:218–232.
LEDER SB, LERMAN JW. (1985). Some acoustic evidence for vocal abuse in adult speakers with repaired cleft palate. Laryngoscope 95:837–840.
MARKS C, BARKER K, TARDY M. (1971). Prevalence of perceived acoustic deviations related to laryngeal function among subjects with palatal anomalies. Cleft Palate J 8:201–210.
McDONALD E, KOEPP-BAKER H. (1951). Cleft palate speech: an integration of research and clinical observation. J Speech Hear Dis 16:9–20.
McWILLIAMS BJ, BLUESTONE C, MUSGRAVE R. (1969). Diagnostic implications of vocal cord nodules in children with cleft palate. Laryngoscope 79:2072–2080.
STATHOPOULOS ET. (1986). Relationship between intraoral air pressure and vocal intensity in children and adults. J Speech Hear Res 29:71–74.
TITZE I, HORII Y, SCHERER R. (1987). Some technical considerations in voice perturbation measurements. J Speech Hear Res 30:252–260.
WARREN D. (1986). Compensatory speech behaviors in cleft palate: a regulation/control phenomenon. Cleft Palate J 23:251–260.
WARREN D, DALSTON R, MORR K, SMITH L. (1988). Respiratory and temporal responses to velopharyngeal inadequacy. Presented at the Annual Convention of the American Speech-Language-Hearing Association, Boston, MA.
WARREN D, WOOD M, BRADLEY D. (1969). Respiratory volumes in normal and cleft palate speech. Cleft Palate J 6:449–460.
WESTLAKE H. (1953). Understanding the cleft palate child. Q J Speech 38:165.
ZAJAC D, LINVILLE R. (1988). Effects of intensity/nasalization on voice perturbations. Presented at the Annual Convention of the American Speech-Language-Hearing Association, Boston, MA.

Commentary

The findings reported by Zajac and Linville are important in that they provide physical evidence from a clinical population that supports the theory that speech motor control may be regulated, at least in part, according to the principles of a system in which speech aerodynamics are monitored. The findings also support Warren's (1986) suggestion that this regulating system may attempt to preserve aerodynamic stability at the expense of speech performance. The notion that individuals with velopharyngeal insufficiency (VPI) invoke laryngeal compensations during speech production is only as new as the finding that glottal stops are frequently used by such individuals. The glottal stop in this context is considered a consonant articulation compensation and, therefore, is different from the type of laryngeal compensation discussed in the accompanying report by Zajac and Linville.

If there is validity in the concept promoted by Warren (1986) that "aerodynamic performance rather than acoustic accuracy (receives) priority in the speech motor control program," (p. 257) it seems likely that this process would be more globally active than previously thought during speech production in individuals with VPI. That is, glottic and supraglottic vocal tract modifications may be invoked in an attempt to equalize aerodynamics during vowel productions as well as during pressure consonant production. It seems logical, therefore, that evidence of abnormal vocal fold vibratory characteristics, such as those reflected in abnormally high electroglottographic jitter and shimmer measures, could follow.

However, these findings must be confirmed in additional studies before they can be completely accepted. Certain details about the methodology described by Zajac and Linville were not discussed. For example, medical evidence confirming that the patients with VPI were determined to be free of vocal pathology would have been useful. Also, no control of loudness was applied during the recording procedure, although the authors acknowledge that differences in speaking loudness could have influenced their results. No data about the loudness levels the speakers voluntarily used are provided, although such data was readily available through the ILS signal processing software employed for signal acquisition and analysis. Signal amplitude is also important for determining the amplitude resolution of the A–D converter. That is, if the amplitude of the signal is considerably lower than the full amplitude range of the 12-bit A–D converter, less than 12 bits of amplitude resolution will be employed. Perturbation measures, shimmer in particular, can be artifically increased if the effective amplitude resolution is below nine bits (Titze et al, 1987).

Studies employing laryngeal perturbation analyses must also carefully consider temporal resolution and length of analysis window. Zajac and Linville used a sampling frequency of 20 kHz. No interpolation was used to increase the effective temporal resolution. It has been suggested (Titze et al, 1987) that no less than 500 samples per cycle are needed to minimize contaminating sampling noise when interpolation between samples is not used. Therefore, when analyzing a 250 Hz signal (the mean F_o for the VPI and control groups reported here was 253 Hz and 258 Hz, respectively), a sampling frequency of at least 125,000 Hz would have been necessary to meet the criteria of Titze et al (1987). Note that sine wave jitter measurements analyzed for calibration purposes in the Zajac and Linville report increased from 0.15 percent at 200 Hz to 1.02 percent at 300 Hz.

Analysis of more than 50 cycles also may be necessary, particularly in nonnormal voices. A recent report (Karnell, 1988) showed that as many as 150 cycles should be analyzed when performing perturbation analyses on nonnormal acoustic speech signals. The optimal analysis window for EGG signals has not yet been determined.

Finally, a clearer distinction should be made between electroglottography and perturbation analyses. It could be inferred from the Zajac and Linville discussion that electroglottographic analysis of voice automatically involves perturbation analysis. This, of course, is not true. Carefully recorded electroglottographic waveforms can provide information about vocal fold vibratory function beyond that reflected in perturbation analyses alone. Although the authors did not provide a rationale for their use of electroglottography, there are some good reasons to perform perturbation analyses in EGG waveforms rather than to perform them on acoustic speech signals, as is most commonly done. For example, aerodynamic turbulence through the velopharyngeal port could conceivably add to measured perturbation in the acoustic speech signal of individuals with VPI, although this seems unlikely during sustained vowels. Use of EGG eliminates the potential effects of such extraneous noise sources on the measures of interest.

The Zajac and Linville report should serve to stimulate additional research regarding vocal processes in patients with velopharyngeal insufficiency. Such research may have important and useful implications for vocal tract modeling and for clinical management. The authors are to be congratulated for their effort.

References

KARNELL MP. (1988). Laryngeal perturbation analysis: minimum length of analysis window. Paper presented at the American Speech-Language-Hearing Association, Boston, MA, November, 1988. Abstract appears in ASHA 30:190.

TITZE I, HORII Y, SCHERER R. (1987). Some technical considerations in voice perturbations measurements. J Speech Hear Res 30:252–260.

WARREN, D. (1986). Compensatory speech behaviors in cleft palate: a regulation/control phenomenon. Cleft Palate J 23:251–260.

Michael P. Karnell, Ph.D.
Assistant Professor of Otolaryngology
University of Chicago