Voice low tone to high tone ratio: a potential quantitative index for hypernasal speech

Guoshe Lee b)

b) Institute of Medical Science, Tzu Chi University, Hualien 970, Taiwan and Department of Otolaryngology, Hualien Hospital, Department of Health R.O.C., Hualien 970, Taiwan

e-mail: guosheli@ms12.hinet.net

Ching-Ping Wangc)

c) Department of Otolaryngology, Veterans General Hospital, Taichung 407, Taiwan

Cheryl C. H. Yangd)

d) Department of Physiology, Tzu Chi University, Hualien 970, Taiwan

Terry B. J. Kuoa,e)

e) Institute of Neuroscience, Tzu Chi University, Hualien 970, Taiwan

Running title: VLHR - potential quantitative index for hypernasal speech

a) Address correspondence to: Terry B. J. Kuo, MD, PhD, Institute of Neuroscience, Tzu Chi University, No. 701 Chung Yang Road, Section 3, Hualien 970, Taiwan. Tel.: +886-3-8565301 ext.: 7513; Fax: +886-3-8580639. Electronic mail: tbjkuo@ms33.hinet.net
Hypernasality is associated with various diseases and interferes with speech intelligibility. Measurements of nasality are clinically used for evaluation of treatment outcome. However, some measurements need special devices and some are subjective. We design a quantitative index called voice low tone to high tone ratio (VLHR) to estimate nasalization. The voice spectrum is divided into low-frequency power (LFP) and high-frequency power (HFP) by a specific cut-off frequency (600 Hz). VLHR is defined as the division of LFP into HFP and is expressed in decibels. Voice signals of sustained vowels [a:] and sustained nasalized vowels [ã:] of five healthy subjects were collected for analysis of nasalance using a Nasometer and VLHR. The correlation between nasalance and VLHR was significant ($p < 0.001$, correlation coefficient = 0.89). Simultaneous recording of nasal airflow temperature with a thermistor and voice signals were made in another 8 subjects. Significant increases of nasal airflow temperature and VLHR were both observed in nasalized vowels [ã:]. The correlation between increases of nasal airflow temperature and VLHRs was also significant. We conclude that VLHR is a quantitative index of hypernasal speech and can be applied in either basic or clinical studies regarding diseases of abnormal nasality or speech training.

PACS numbers: 43.70.Aj.
I. INTRODUCTION

Abnormal nasal emissions are frequently associated with hypernasality (Dalston and Warren, 1986). Clinically, there are several diseases related to abnormal nasal emission like velopharyngeal insufficiency, cleft palate, palatal fistula (Pinborough-Zimmerman et al., 1998) and stroke (Yorkston et al., 1989). The associated hypernasality often causes speech intelligibility (Younger and Dickson, 1985) and disturbance of communication.

Clinically, there are various methods to evaluate nasal airflow during phonation such as the mirror-fogging test (Schmelzeisen et al., 1992), pneumotachometer (Trullinger and Emanuel, 1983), nasal anemometer (Hutters and Brondsted, 1992), and aerophonoscope (Devani et al., 1999). The nasalance, which is derived from the equipment Nasometer, is broadly used for the assessment of hypernasality resulting from diseases such as cleft palate (Van Lierde et al., 2003) and velopharyngeal insufficiency (Dalston et al., 1991). However, some of these tests are subjective and some of them need special devices. The development of an accessible, easy and objective tool warrants further investigation.

Many spectral characteristics have been proved to be related to vowel nasalization such as the reduction of the intensity of the first formant, presence of extra-resonance, etc. (Curtis, 1970). To our knowledge, there are no articles describing quantification of hypernasal speech using voice spectral analysis. A recently developed index called voice low tone to high tone ratio (VLHR) has been proven to be closely related to the patency of nasal airways (Lee et al., 2003). The index is defined as a ratio of low frequency power (LFP) to high frequency power (HFP) after arbitrarily dividing the voice spectrum using a specific cut-off frequency. The VLHR increases significantly after nasal obstruction being relieved by nasal decongestion treatment. With this, we forecast a positive relationship between VLHR and hypernasal speech.

VLHR is derived from voice spectral analysis and it is an objective and quantative
index. Measuring VLHR requires a standard microphone and proper software. However, the relationship between VLHR and hypernasal speech or nasalance has not been established yet. Here we designed an experiment using the voice samples of non-nasal vowel [a:] and nasalized vowel [ã:] in healthy subjects to correlate VLHR with nasalance scores and nasal airflow detected using a thermistor. The protocols, results, and applications are discussed.

II. METHODS

Five male healthy subjects were enrolled in the first study for the exploration of the relationship between VLHR and nasalance. The mean age of the subjects was 26.0 ± 2.0 (means ± SD) years. Hearing screening tests using pure tone audiometry at 25 dB were performed prior to the study to ensure normal hearing. Subjects with medical history of dysarticulation, craniofacial anomaly and nasal diseases were excluded. They were requested to produce non-nasal vowels [a:] and nasalized vowels [ã:] for at least 5 seconds for the analysis of nasalance in a quiet room. The voice samples were recorded with a dynamic microphone (JVC model MD480A) which was placed 10 cm in front of the subjects’ mouths. Voice signals were sampled at the rate of 22 kHz and were recorded in a digital format of 16-bit resolution with an IBM personal computer (PC) compatible sound adapter. The nasalance scores were acquired using a Nasometer II (model 6200, Kay Elemetric).

The 0.5-s vocie signals after voice onset were bypassed and the signals of the next 3 seconds were segmented for analysis. The voice spectrum was derived using fast Fourier transform (FFT) of the entire 3-s sound. The voice spectrum was then divided into a low-frequency part and a high-frequency part artificially with the cut-off frequency of 600 Hz. The LFP was defined as the summation of the power of frequency components ranging from 65 to 600 Hz. The HFP was defined as the summation of the power in the range of 600 to 8000 Hz. The VLHR was defined as the division of LFP into HFP and expressed in decibels (dB) like the common expression of sound energy. The equation is list below
VLHR = 10 \times \log_{10}(\text{LFP} / \text{HFP})

Eight healthy volunteers, four men and four women, were enrolled in the second experiment for the exploration of the relationship between nasal airflow and VLHR. The mean age of the subjects was 25.2 ± 6.0 years (means ± SD). The subjects were instructed to produce the sustained vowel [a:] and sustained nasalized vowel [ã:] for at least 5 seconds. The voice samples were recorded in the same conditions as the experiment one.

A temperature sensor (NTC thermistor 5 KΩ, 25 °C, time constant = 1.8 s) was placed rightly in front of the nostril and fixed steadily with a headset-like device to detect the temperature of nasal airflow. The temperature signal was amplified using an amplification circuit. The sampling rate of the thermistor was set to 1000 Hz and recorded in digital format by the analogue to digital device (National Instrument, DAQ-700). Calibration was made before tests. The correlation between temperature signals and voltage signals was found to be linear in the range of 20°C to 40°C (data not shown). Both sound signals and temperature signals were recorded simultaneously. The 0.5-s signals after voice onset were bypassed and the signals of the next 3 seconds were segmented for analysis. The VLHR of each subject was obtained using the algorithm described above.

Correlations among all parameters were accessed using Pearson’s correlation coefficient. We considered a good or strong correlation between two variables at $r \geq 0.7$. Comparisons between two sets of data were performed with the unpaired Student $t$-test. Statistical significance was assumed for $p < 0.01$. Values are expressed as means ± SE.

### III. RESULTS

The F0 of vowels [a:] and nasalized vowels [ã:] were not significantly different and the $p$ value was 0.71. The intensity of these two vowel sounds was also not statistically different and the $p$ value was 0.68.

We illustrated the relationship between VLHRs and nasalance scores in Fig. 1. The
The correlation coefficient was 0.89 and the $p$ value was < 0.001. The lower the nasalance scores, the lower the VLHRs. We also found that the VLHRs of the non-nasal vowels [a:] were in the range of 0 to 5, and the VLHRs of the nasalized vowels [ã:] were in the range of 5 to 10. Thus, a single measurement of VLHR could differentiate the non-nasal vowel [a:] from the nasalized vowels [ã:]. These results provided evidence of the reliability and stability of VLHR.

We demonstrated the temperature of nasal airflow and VLHR of one study subject in Fig. 2. The nasal airflow temperature increased more rapidly in the nasalized vowel [ã:] [Fig. 2(B)] than in the non-nasal vowel [a:] [Fig. 2(A)]. This showed evidence that the vowels phonated by our subjects were nasalized. Meanwhile, we found that the VLHR of the nasalized vowel [ã:] [Fig. 2(D)] was significantly higher than the non-nasal vowel [a:] [Fig. 2(C)]. The higher VLHR for the nasalized vowel [ã:] implied the correlation between VLHR and hypernasal speech.

We plotted the temperature change rates of nasal airflow in Fig. 3(A). Significant increase of the temperature of nasal airflow was observed for the nasalized vowel [ã:] and the $p$ value was < 0.01. We calculated the VLHRs and significant increase of VLHR was also observed for the nasalized vowels [ã:] in all subjects, and the $p$ value was < 0.001 [Fig. 3(B)]. The vowels [ã:] produced by our subjects were nasalized on statistical basis regarding the temperature increase of nasal airflow detected by the thermistor. VLHR reflected the change of nasalization of vowel [a:] in our study.

We found a positive correlation between the VLHR and the rates of temperature change [Fig. 4(A)]. The VLHR increased with the rate of temperature change. The correlation was significant ($p < 0.001$), and the correlation coefficient ($r$) was 0.76. The sound intensity, which was expressed in RMS, very weakly correlated with VLHR [Fig. 4(B)]. VLHR increased independently of sound intensity in the nasalized vowel [ã:].
IV. DISCUSSION

In this study, we collected the voice samples of the vowels [a:] and nasalized vowels [ã:] to analyze the relationship between VLHR and nasalance scores and between VLHR and temperature changes of nasal airflow. The correlations are good ($r > 0.7$) and significant ($p < 0.01$). These indicate that VLHR may reflect the changes in nasalance scores and the airflow resulting from nasal speech.

The voice spectra are different between vowels, thus the VLHR may vary among them. We tested vowels other than [a:] using the methodology described above. The increase of VLHR after vowel nasalization was commonly observed. The difference of the vowel [a:] was very significant according to our results here. The relationship between VLHR and the conventional index nasalance is good ($r = 0.89$) and significant ($p < 0.001$). The promising results warrant future investigations for different vowels, different consonants and disease conditions such as velopharyngeal insufficiency and cleft palate.

Simultaneous recording of voice signals and nasal airflow while phonation is not easy using current measuring methods of nasality. For example, using a pneumotachometer to measure nasal airflow while phonation prevents the voice signals from the nose by the mask. The nasal voice signals are lost and the recorded signals are different from those recorded under ordinary conditions. The thermistors are widely used for detecting airflow in various sleep studies like polysomnography (Akre et al., 2000). Although there is non-linearity of thermistors with regard to the airflow rate, the correlation between temperature change and air velocity has been established. For the thermistors with the response time in the order of 1 second, the voltage response is highly sensitive to relatively low air velocities ($<1$ m/s) (Patra et al., 1986). Thus, we designed a thermistor probe applied in front of the nostril to facilitate synchronous recording of nasal airflow temperature without interference with voice recording. The temperature rates for the nasalized vowel [ã:] in our study were significantly
higher than for the non-nasal vowel [a:]. This revealed that the vowels that our subjects produced were nasalized in certain degrees. Meanwhile, we found significant increases of VLHR in the nasalized vowels [ã:] and a good correlation between VLHRs and the temperature changes of nasal airflow. Although may not be linear, the correlation between VLHR and the nasal airflow during phonation should be significant.

The sound intensity was influenced by the gain setting of recording devices. However, we used the same settings for all of the subjects while recording. In addition, an important idea that we would like to share is that the intensity changes of a voice spectrum appear both in low frequency and in high frequency during various phonation conditions. In certain instances, such as different phonation power and different gain settings of recording, etc., the changes of low frequency and high frequency were essentially symmetrical. The division of LFP to HFP simply countervails the effects. In our previous VLHR study for the evaluation of nasal airway (Lee et al., 2003), the VLHR was weakly related to sound intensity and vocal fundamental frequency regardless of nasal congestion or nasal decongestion. In this study, we found the same relationship between sound intensity and VLHR. The correlation between sound intensity and VLHR was modest as the demonstrated in Fig. 4(B). Consequently, in cases of phonation at comfortable speech levels, the VLHR was weakly influenced by sound intensity.

We surveyed the VLHRs of various cut-off frequencies in the range of 100 to 8000 Hz. The VLHRs of the cut-off frequencies ranging from 400 to 800 Hz made the most significant differences. In order to simplify the calculation of the algorithm, we selected the average value (600 Hz) as the cut-off frequency for analysis.

Several spectral characteristics of voice, such as the reduction of the intensity of the first formant, presence of extra-resonance, and increased bandwidth of formants, have been reported for nasalized speech (Curtis, 1970). The additional acoustic property of shifts in the
center of gravity of the low-frequency spectral prominence (Hawkins and Stevens, 1985) has been reported as well. The spectral characteristics of increased amplitudes of bands between F1 and F2 and decreased amplitude of the band of F2 (Kataoka et al., 2001) have been reported related to hypernasality. In this study, the increase of VLHR in the nasalized vowel [ã:] was compatible with these findings. Moreover, this property was quantified easily using VLHR as we described here.

VLHR is an objective index. In addition, a standard microphone, PC and proper analytic software is adequate for data acquisition and analysis. The voice samples that were previously stored in tapes or disks can be submitted for analysis. With the transmission of digital voice samples via internet resources, VLHR can even be obtained from a distant computer terminal. This approach of nasalization can be applied in either basic or clinical studies regarding diseases of abnormal nasality and speech training.

ACKNOWLEDGMENTS

This study was supported by grants form the National Science Council, ROC (NSC 89-2314-B-320-016).


FIGURE CAPTIONS

FIG. 1. Linear regression plot of voice low tone to high tone ratio (VLHR) with nasalance scores.

FIG. 2. The temperature gradients of nasal airflow and voice low tone to high tone ratio (VLHR) for the sustained vowel [a:] and the sustained nasalized vowel [ã:] of the 3-s sound in one study subject. The temperature increased more rapidly for nasalized vowel [ã:] (B) than non-nasal vowel [a:] (A). The VLHR of the nasalized vowel [ã:] (D) was also significantly higher than the non-nasal vowel [a:] throughout the entire 3-s sound (C).

FIG. 3. The rates of temperature change of nasal airflow (A) and voice low tone to high tone ratio (VLHR) (B) of sustained vowel [a:] and sustained nasalized vowel [ã:]. The rates of temperature change and VLHR were both significantly higher for nasalized vowel [ã:]. Values were expressed in means ± SE.

FIG. 4. Linear regression plot of VLHR with rates of temperature change (A) and sound intensity (B). The two dot lines indicated the range of 95% confident intervals (2 standard deviations). The sound intensity was expressed in RMS unit.
FIG. 1.

$r = 0.89$
$p < 0.001$
FIG. 2.
A

TEMPERATURE CHANGE RATE (°C/s)

0.15

0.10

0.05

0.00

[a:]  [ā]

*p < 0.01  *

B

VLHR (dB)

20

15

10

5

0

[a:]  [ā]

**p < 0.001  **

FIG. 3.