Theory and Instrumentation for Quantitative Measurement of Nasality

SAMUEL G. FLETCHER, Ph.D.
Albuquerque, New Mexico

Speech is recognized as the essential criterion for assessing the ultimate results of treatment for children born with clefts of the palate. In spite of this fact, the precision with which speech and underlying oral function may be assessed has developed rather slowly. The main limitation in progress seems to have been that essentially all diagnostic and remedial approaches have depended upon perceptual evaluation as the basis for clinical judgments. This approach has at least two major limitations. First, the standard baselines against which judgments are made are non-stable and thus rather undependable for repeated comparisons over time, or comparisons among different clients. Second, the process of sharing information among clinicians is hampered by the vagueness with which the basic parameters of disability are defined. Thus, the focus of attention has tended to shift to known parameters, such as speech articulation, which can be classified more reliably, rather than to nasality which is more continuously variable and therefore less easily managed by perceptual judgment.

The rapid evolution of electronic technology has made new biomeasurement procedures feasible for more precise, quantitative, and reliable measurement of nasality in persons with operated clefts of the palate. The present paper summarizes the results of a project to develop an instrument for this purpose. The feasibility of this electronic instrumentation designed to measure acoustic characteristics beyond the comparative intensity of sound emitted from the mouth and nose was suggested in 1961 through use of a dual-channel electrophonocardiograph with some capability for frequency selectivity. Formal work to fabricate the presently described instrumentation was begun in 1966.

Acoustic Theory

Three aspects of acoustic theory pertaining to assessment of distortion from nasality are fundamental to electronic instrumentation ap-

Dr. Fletcher is Director of Speech Pathology and Audiology, University of New Mexico.

This project was supported in part by a grant from the Easter Seal Research Foundation of the National Easter Seal Society for Crippled Children and Adults, Inc.

This paper was presented at the 1969 International Congress on Cleft Palate, Houston.

proaches to its measurement. a) Two channels simultaneously generating information contained in the total speech spectrum have the capacity to mask critical characteristics of each other. b) Diversion of acoustic energy from a primary transmission channel may be expected to have deleterious effects upon signals being transmitted through that channel. c) Superimposition of two signals with resultant shifts in acoustic energy will generate unique configurations of the total envelope and these will be reflected in speech distortion. An overriding assumption is that speech components in a number of frequency bands contribute independently to speech intelligibility.

The importance of masking to perception of sound signals has long been recognized in acoustic theory. In essence, masking refers to the reduction in audibility of one signal when it coexists with another (*7*). The frequency content of one sound dictates its acoustic masking properties with respect to another. For example, when a pure tone is sounded in the presence of a random noise, a narrow band of frequencies extending to each side of the tone is capable of covering it from a listener's perception. These frequencies capable of masking out a particular tone are known as its "critical band" (*5*). As the frequency ranges of two competing signals are expanded and have greater overlap, the potential for masking is seen to be increased.

Acoustic theory concerning masking seems to have considerable relevancy to hypernasality. In such a condition, sound is not neatly and alternatively separated into nasal and oral routes of emission by the precise palatal valving typical of normal speech. Rather, the phonic stream is bifurcated to pass into the two routes simultaneously. Since each of these channels functions as a variable filter tuned by the resonating characteristics of its cavity, the outputs from the nose and from the mouth emerge with unique acoustic attributes capable of cross masking. Our main concern is of course nasal masking of the oral signal since this is a principal source of distortion in cleft palate speech.

The second type of acoustic distortion affecting intelligibility of cleft palate speech is postulated to be loss of energy in critical components or elements within speech. Fragments of acoustic energy may be lost to the nasal cavity at critical junctures of the speech event. Although they may not reach the threshold of audibility and therefore would not be a source of acoustic masking, they could nevertheless represent a disruption in the normal balance of acoustic energy as it is distributed among the phonemes. In this context, those elements which have the greatest demands for pressure buildup within the oral cavity would be most liable to disruption from distortion from inadequate palatopharyngeal valving. Thus, plosives which demand a period of pseudo-silence followed by a sharp noise burst and sibilants which have high frequency components from sustained air flow through narrow channels would be expected to have relatively high levels of acoustic distortion from fragmentation of the primary speech stream through the mouth.

The third type of acoustic distortion postulated as basic to reduction in cleft palate speech intelligibility is a shift in relative energy among frequency bands within the total acoustic envelope. When two waves meet each other in the same medium, the instantaneous displacement of the medium is given by the algebraic sum of the instantaneous displacements of the individual waves. Thus, sound from the nasal cavity superimposed upon sound from the oral cavity may increase the prominence of previously inconsequential frequency bands. This reshaping of the acoustic spectrum within the total envelope may bring certain frequency bands from a state of recessiveness into prominence with resultant distortion and varying disruption of speech intelligibility.

Instrument Criteria

With the above theoretical framework in mind, instrumentation was fabricated to meet the following criteria: a) The instrumentation must be capable of separating the acoustic outputs from the oral and nasal cavities during speech. b) It must include separate filter networks for the two channels with identical bandwidths and frequency ranges which are simultaneously tunable by manual control or through the use of a common sweep voltage. c) All frequencies between 50 and 20,000 Hz should be potentially analyzable without distortion. d) Ratios between sound levels in each of the channels and within selectable frequency bands must be able to be instantaneously computed, displayed and quantified. Information concerning the magnitude of sound within each channel must also be separately displayed. e) The data must be yielded in such a form that it can be interpreted by clinicians.

Description of Tonar Instrumentation

Since the principal purpose of this system is to measure ratios between acoustic outputs from the nose and the mouth, the instrument is called TONAR (The Oral Nasal Acoustic Ratio). Figure 1 is a block diagram showing the general components of the system developed. The circuit was designed to accept an input from live voice or from any dual-channel tape recorder with high impedance output.

Separation of sound being emitted from the oral and nasal cavities is initially accomplished by two lead chambers cut to fit the facial contours. Microphones suspended in fiberglass packings are contained in each chamber. An intermicrophone attenuation greater than 40 dB is realized in the sound separator.

Signals from the oral and nasal microphones are amplified, then fed to Quantech tracking wave and spectrum analyzers. These analyzers are connected together to form a dual-channel, phase locked, master-slave

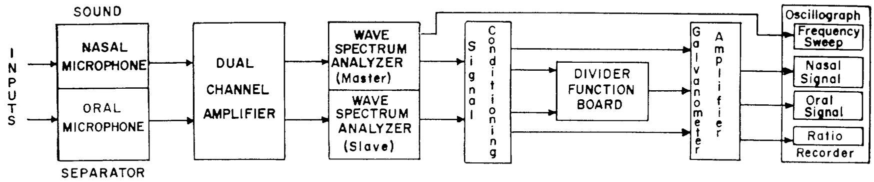

FIGURE 1. Diagram of the Oral Nasal Acoustic Ratio (TONAR) System.

system with rather unusual scanning flexibility. One may manually tune the master spectrum analyzer using a 5000 division linear scale to measure phase and amplitude simultaneously from the oral and nasal channels in broadband from 10 Hz to 65 K Hz or in passbands of 10 Hz, 100 Hz, 1000 Hz with 24 dB rolloff. The system can also be placed in a sweep mode so that signals with a predetermined frequency band of 0.5, 1, 5, 10, or 50 K Hz are scanned automatically in time increments of 0.5, 5, 50, or 500 seconds. Some limitations are imposed in sweep rate with regard to bandwidths which can be accepted without loss of accuracy.

The d-c voltages emerging from the wave analyzers are divided. One set of signals is led to a CEC type 1-162A galvanometer-driver amplifier and thence to a CEC 5-124 oscillograph recorder for display. The equivalent set of signals is led to an analog ratio computer system for further processing before recording them on the oscillograph.

The ratio computer system developed for Tonar has an ISR 1041A divider function printed circuit board as its central component. Operating as an electronic slide rule, the ratio computer automatically provides a continuous ratio of the oral and nasal voltages from the wave analyzers. The comparative voltage from the two channels emerges from the ratio computer in analog form and is passed through a galvanometer-driver amplifier to the oscillograph. This analog ratio is then converted to digital form through calibration of the oscillographic recorder such that baseline represents a reading of zero and a ten-line deflection of the oscillographic beam represents a ratio of 1.0. Each line of the graph thus represents a ratio change of 0.1. Further interpolation may be done as desired since the total deflection is essentially linear.

In early stages of instrument development, the ratio generated by the computer was felt to have questionable linearity beyond the 0-1.4 ratio range; therefore a switching circuit was placed in the line so that either channel could be alternatively selected as numerator in the ratio. Thus, in normal subjects or in subjects with mild to moderate hypernasality, the circuit preceding the ratio computer is switched so that the nasal signal (E_1) is the numerator and the oral signal (E_2) is the denominator. In this situation, the ratio E_1/E_2 is directly proportional to nasality. Numbers near zero indicate normal nasal resonance. In instances of severe hypernasality, the voltage comparison may be switched to E_2/E_1. In this setting, the ratio is inversely proportional to nasality. That is, the numerical value of the ratio will become smaller if either the voltage from

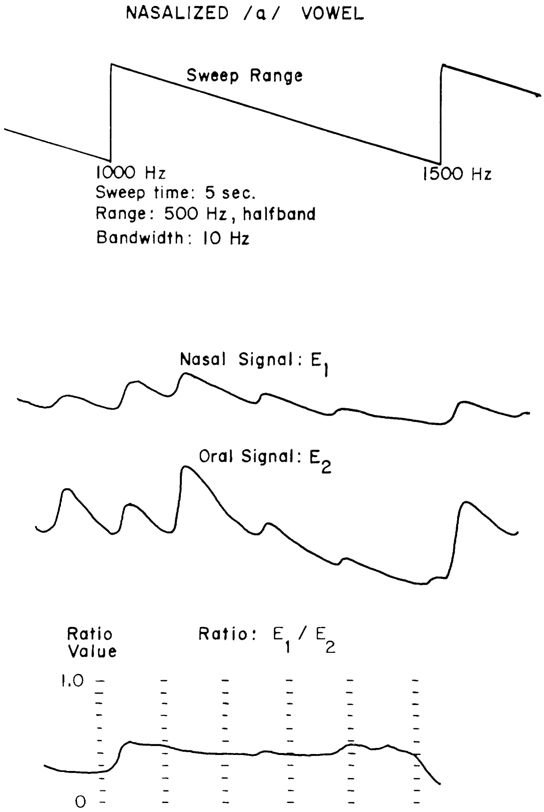

FIGURE 2. Oscillographic display of the four-signal output from Tonar.

the oral channel increases or the voltage from the nasal channel decreases. Numbers approaching zero would represent extremely high nasal sound levels.

Experimental work to date indicates that residual hypernasality of the vast majority of subjects with repaired palatal clefts can be adequately measured with the E_1/E_2 setting.

An additional signal was built into the system to facilitate frequency analysis. The saw tooth line at the top of Figure 2 is from an automatic sweep voltage generator built into the master spectrum analyzer. This allows the spectrum analyzers to be swept through a particular frequency range and to reflect continuously the intensities of sounds throughout the range. The rate at which this sweep is accomplished may be varied from 0.5 to 5.0, 50.0, or 500.0 seconds. Each time the sweep is initiated at the frequency dial setting and progresses through a specified frequency range of 0.5 K, 5.0 K, 10.0 K, or 50.0 K Hz. As the sweep is made, the spectrum analyzer generates a parallel change in voltage, starting at zero and progressing in linear fashion to 4.5 volts.

Identification of any instantaneous frequency being scanned in time may be accomplished in a rather straightforward manner. For example, if the master analyzer is set at 1.00 K Hz and the sweep increment is set at 0.5 K Hz with a five-second sweep, both spectrum analyzers will simultaneously scan the speech input between 1000 and 1500 Hz in consecutive five-second intervals. If one then wishes to know the amplitude of any signal at a particular frequency point such as 1200 Hz, he simply measures the fractional distance of that frequency between the 1000 and 1500 limits, marks that point on the sweep line, and drops a vertical from this point through the other trace marks displayed on the readout.

The second and third trace marks on the oscillographic display in Figure 2 reflect the rectified and amplified voltages from the oral and nasal channels as they emerge from the wave analyzers. It may be seen that both of these traces have considerable variation by frequency.

The final trace on the bottom of Figure 2 represents the instantaneously computed ratios of the voltages from sound emitted from the oral and nasal chambers. This trace reflects the comparative intensity of sound according to frequency band and the sounds being uttered. It has a totally different configuration from any of the other traces since it represents comparative nasality. When this analog data is converted to numerical form, it may be used readily for clinical interpretation.

Discussion

Objective measurement of nasality is fundamental to theoretical constructs relating speech competency and structural impairment in cleft palate speakers. Achieving such objectivity has been particularly problematical. Psychophysical techniques using rating scales have been plagued by "halo" effects whereby secondary characteristics such as differences in pitch, speech articulation and level of linguistic skills contaminate judgment (*12*). Playing the recorded sequences of speech backwards has been suggested as a means to remove the halo effects mentioned. This procedure, however, introduces its own variables such as disrupted rhythm patterns which may generate a new set of halo effects (*6*). Most psychophysical procedures are also prohibitively time consuming. More direct approaches are needed.

Sound spectrography gave much promise as an objective tool for direct measurement of nasality (*1*). High variability in acoustic characteristics of nasality from person to person has obstructed use of this approach in general studies of cleft palate speech.

Clearly, the most direct method for objectively estimating nasality would be to measure and compare sound pressure levels of sound emitted from the mouth and the nose. This approach was used in a published study by Shelton and associates (*11*) and in unpublished studies by Weiss (*13*), Low (*8*), Pierce (*9*), Bryan (*2*) and Coleman (*3*). In each instance, judged ratings of nasality were compared to oral and nasal sound pressure levels measured by condensor microphones. The results among these studies were highly variable.

In the Weiss study, perceived nasality was established by fourteen judges using paired comparisons. The frequencies in which each sample was more nasal than its pair were transformed into proportions and then normalized by arcsin transformations. Pearson correlation coefficients, obtained between judged ratings and sound level measures were: nasal pressure, .74; oral pressure, .57; oral pressure minus nasal pressure, .94; and nasal pressure divided by oral pressure, .91. These findings suggest high validity in the sound level approach to measurement of nasality. Later studies have not confirmed such high correlations.

Pierce compared judged ratings of nasality with oral and nasal sound level differences measured on five subjects. Kendall coefficients of correlation derived were in the general range of .47 to .73 although four of fourteen comparisons made reached a level of 1.00. Pierce concluded that while "the coefficients do not show completely consistent relationships between the mean SPL (sound pressure levels) differences and the expert listener's judgment of perceived nasality . . . the two methods of measurement appear to have yielded information that is clearly related, particularly for subjects A, B and C". He also noted that speech characteristics other than loudness of nasal cavity resonance may well have influenced listener interpretation of nasality.

In 1955, Low devised an ingenious system whereby an "acoustic separator" was used to isolate oral from nasal sound and thereby help make sound level measurements of either source of emission which were relatively uncontaminated by each other. Interchannel attenuation by the sound separator was not formally established. He examined 64 to 86 young adults in different phases of his study. All but four of the subjects were free of palatopharyngeal incompetency. Preliminary correlations between sound level measurements and listener ratings of perceived nasality by four judges were low (.19). When subjects with "excessive" vocal stridency and those with denasality were excluded, correlations were significant but not high (.50).

The high correlations found in the Weiss study are seen to stand in marked contrast to studies by other investigators. This discrepancy motivated Shelton and associates (*11*) to conduct a careful investigation which replicated the study by Weiss. In the Shelton study, 31 subjects representing a wide range in nasal resonance were examined. Perceived nasality was established by ratings of twelve judges, using a normal voice and an "extremely hypernasal voice" as anchors in their nasality ratings. Nasal sound level was measured by means of a microphone suspended five and one-half inches from the speaker's lips with a probe attachment placed "just inside the subject's naris". Oral sound pressure was obtained with a microphone. Pearson correlation coefficients computed between judged nasality and sound level measurements were: nasal pressure, .47; oral pressure, .01; nasal minus oral, .41; and nasal divided by oral, .37. While the correlations of .37 and above are significant, they again suggest that the correlations in the Weiss study were spuriously high.

Perhaps the discrepancy noted between listener ratings of perceived nasality and previous electronic instrumental evaluation of nasality should not be interpreted as too disconcerting. As noted, contamination arises from a variety of factors which can and do influence judgments of nasality. Furthermore the electronic instrumentation cited was programmed to compare only the intensity of the sound emitted from the nasal and oral channels. Frequency and time relationships were not systematically considered. The importance of these aspects of nasality is suggested in the preliminary study reported in the accompanying article (*6*). The function of frequency as an attribute of nasality is also emphasized in recent discussions by Curtis (*4*) and by Schwartz (*10*).

In view of the complexity of psychophysical and acoustic specification of nasality it would seem that the level of agreement found by previous investigators between psychophysical scaling and instrumental measurements of nasality should be interpreted as rather encouraging.

Summary

Methods for objective measurement of nasality are briefly reviewed. Refinement in electronic instrumentation for this purpose is suggested as a needed development. Underlying acoustic theory is presented and an instrument is described (TONAR) which was fabricated for electronic analysis of nasality and instantaneous derivation of an oral-nasal intensity ratio within specified frequency bands. The results can be readily converted to digital form.

reprints: *Samuel G. Fletcher, Ph. D.*
Director, Speech Pathology and Audiology
University of New Mexico
Albuquerque, New Mexico 87106

References

1. Bloomer, H., and G. Peterson, A spectrographic study of hypernasality. *Cleft Palate Bull., 5,* 5–6, 1955.
2. Bryan, G. A., Relationships among nasal and 'oral' sound pressures and ratings of nasality in cleft palate speech. Ph.D. dissertation, University of Oklahoma, 1963.
3. Coleman, R. O., The effect of changes in width of velopharyngeal aperture on acoustic and perceptual properties of nasalized vowels. Ph.D. dissertation, Northwestern University, 1963.
4. Curtis, J. F., Acoustics of speech production and nasalization. In *Cleft Palate and Communication,* pp. 27–60. New York: Academic Press, Inc., 1968.
5. Fletcher, H., Auditory patterns. *Rev. mod. Physics, 12,* 47–65, 1940.
6. Fletcher, S. G., and M. E. Bishop, Measurement of nasality by TONAR. *Cleft Palate J., 7,* 610–621, 1970.

7. HIRSCH, I. J., Auditory masking and fatigue. In *The Measurement of Hearing*, pp. 154–187. New York: McGraw-Hill, Inc., 1954.
8. LOW, G. M., An objective index of nasality. Ph.D. dissertation, University of Minnesota, 1955.
9. PIERCE, B. R., Nasal resonance differences resulting from speech appliance modification in cleft palate adults. Ph.D. dissertation, Northwestern University, 1962.
10. SCHWARTZ, M. F., The acoustics of normal and nasal vowel production. *Cleft Palate J., 5,* 125–140, 1968.
11. SHELTON, R. L., A. W. KNOX, W. B. ARNDT, and M. ELBERT, The relationship between nasality score values and oral and nasal sound pressure level. *J. speech hearing Res., 10,* 542–548, 1967.
12. SHERMAN, DOROTHY, The merits of backward playing of connected speech in the scaling of voice quality disorders. *J. speech hearing Dis., 19,* 312–321, 1954.
13. WEISS, A. L., Oral and nasal sound pressure levels as related to judged severity of nasality. Ph.D. dissertation, Purdue University, 1954.