Measurement of Nasality with Tonar

SAMUEL G. FLETCHER, Ph.D.
MILO E. BISHOP, M.A.
Albuquerque, New Mexico

Objective measurement of nasality is fundamental to clinical considerations of speech competency in speakers with repaired clefts of the palate. Achieving objectivity has been particularly problematic. Psychophysical techniques using rating scales have been plagued by "halo" effects whereby secondary characteristics such as differences in pitch, level of linguistic skill, and proficiency in speech articulation tend to contaminate judgments (*16*). Most psychophysical procedures are also prohibitively time-consuming and difficult to use in routine clinical application. More direct approaches are needed.

Nasalization of speech may be most easily interpreted as the variable coupling of a nasal shunt onto the laryngeal-pharyngeal-oral vocal tract (*9*). The effects of this coupling are reflected in the acoustic output; therefore sound spectrography gave much promise as an objective tool for direct measurement of nasality (*3*). The usefulness of this approach in general studies of cleft palate speech was limited by the high variability in the general acoustic characteristics of speech containing imbedded nasality.

Since nasality is contingent upon physiological intermixing of phonic energy between the nasal and oral channels, the most direct method to estimate nasality objectively would seem to be to separate and compare the sound energy from the mouth and the nose. This approach has been used in a published study by Shelton and associates (*15*) and in unpublished studies by Weiss (*18*), Pierce (*14*), Bryan (*5*), and Coleman (*6*). 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 but did demonstrate significant agreement between the sound pressure levels obtained and listener judgments of nasality.

Fletcher (*8*) recently reported on an instrumental system, Tonar, which is similar to instruments used by previous investigators in that it makes use of separated oral and nasal signals to quantitize nasality.

Dr. Fletcher is Director, Speech Pathology and Audiology, University of New Mexico. Mr. Bishop, now at Purdue, was formerly a graduate student at the 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.

Tonar, however, has the additional capability to measure the intensity of the emerging signals within specific frequency ranges, and it provides electronically computed ratio readouts that can be readily converted to digital form.

Before any electronic instrument such as Tonar can be used to gain meaningful information about the acoustic nature of clinical nasality, it must be shown to analyze parameters which correspond with perceived nasality. The present study was designed to examine perceptual and acoustical factors thought to be subsumed in nasality measurements by Tonar. The study was designed to answer the following types of questions:

> What are the relationships between Tonar ratios and listener ratings of nasality? Are such relationships constant regardless of how perceived nasality is rated? Do Tonar ratios vary systematically as a function of frequency? Is the frequency distribution within the oral and nasal channels such that acoustic masking could be a factor in perceptual judgments? Is Tonar a potentially valuable clinical tool?

Procedures

Subjects. Thirty-five subjects were chosen randomly from the files of 250 children with surgically repaired palatal clefts at the Rehabilitation Center, Inc., Albuquerque, New Mexico. Of these 35, eleven could not be contacted or were not available during the period of the study and four were excluded because of upper respiratory infections or nasal congestion during the test period. The data herein reported was drawn from the twenty children remaining. They ranged in age from five to nineteen years with a median of eight and a mean of ten years.

Data Collection. A 70-word, phonemically balanced, non-nasal passage prepared by the first author of this paper was used as the speech sample for the study. Two recordings were made of each subject, one with the passage spoken in a sound field and the other with the subject speaking the paragraph into the sound separator of Tonar and with oral and nasal channels recorded separately. All recordings were made on a Sony model 600 stereo tape recorder with the subject, microphones, and sound separator inside an IAC model 1202 sound control room. This instrumentation was connected via a jack panel to an external recorder, amplifier, spectrum analyzers, and oscillographic recorder outside the chamber.

For sound field recordings, the subject was positioned in the sound control room with his lips four to five inches from the microphone. Each subject was then individually instructed to read the passage until he was familiar with all of the words before the recording session was initiated. The younger subjects who were unable to read were familiarized with the material by having it read to them. They then repeated the sentences following earphone instructions.

Tape recordings for Tonar analyses were obtained by positioning the subject to speak into the sound separator. The subject's head was postured to assure a tight seal between the oral and nasal cylinders without restriction of the nares by the lead separation plate. This could be accomplished by adjustment of the height and angle of the separator without discomfort to the subject. The remainder of the recording procedure was similar to that described in sound field recording.

Calibration of Tonar. Calibration of Tonar was accomplished by placing the sound separator eight inches in front of a ten-inch speaker driven by a Hewlett-Packard model 200 ABR audio oscillator set to generate a 1000 Hz tone at 65 dB SPL. The calibration tone was recorded at equal VU levels on both channels of the tape recorder. This tone was then used to calibrate the remaining instrumentation so that a baseline on the oscillographic readout represented zero and each of ten progressive line deflections reflected a voltage ratio change of 0.1 in relative nasal to oral intensity. During the recording and analyzing sessions, calibration checks were made on the system after every three speakers or three hours, whichever occurred first.

Psychophysical Scaling of Nasality. Ten observers were used to scale perceived nasality. Eight observers were graduate students enrolled in a cleft palate seminar at the time of the scaling sessions. The reamining two observers consisted of one of the authors of the present paper (S.G.F.) and another staff member with considerable experience working with cleft palate children.

An initial attempt was made to scale the nasality in backward play mode through use of a paired comparison procedure. The resultant data showed inconsistencies that were felt to be too great for adequate comparison with the electronic measurements; therefore the scaling procedure was restructured to use a modified direct magnitude estimation procedure. The initial scaling experience with paired comparisons then served as a practice session to orient the observers to backward played speech.

All ten observers were seated in a large room and given instructions as to the procedures involved in the direct magnitude scaling procedures to be followed. They were asked to assign values between 0 and 100 to the recorded stimuli using the first stimulus as the standard for further rating. A total of thirty stimuli presentations was given. The subjects were then given a ten-minute rest period, and the procedure was repeated with the speech in forward play mode. A randomized presentation was used in each mode to remove any order effect or inter-rating contamination.

As a check for test-retest reliability of the observers, ten of the thirty stimulus presentations in each mode were replicas. In each instance at least fifteen other stimuli were interposed between the first stimulus and its replica. The ratings of the first stimulus in each mode of presentation were used to calculate an average score for each of the twenty subjects. These average values were employed later in test-retest comparisons of the ten subjects with replication presentations and with findings from Tonar measurements.

TONAR ANALYSIS OF NASALITY. As described in detail in an accompanying article (*8*), Tonar consists of a dual-chamber sound separator with nasal and oral microphones, two interlocked spectrum analyzers, an analog ratio computer, and an oscillographic recorder for display and digital conversion of the oral and nasal signals and a computed ratio between them.

The first stage of Tonar analysis, namely separation of the oral and nasal signals, has already been described. The tape-recorded speech samples from all twenty subjects were fed into the spectrum analyzers at central frequencies of 1250 Hz and 1750 Hz and a 1.0 K Hz half bandwidth. A time constant of 10 seconds was used to smooth out the general response characteristics. The ratio computer was placed on a nasal/oral setting (E_1/E_2) and the oscillographic time marker was set at one mark per second with a paper speed of 0.5 ips. Ratio values were obtained by counting the line deflections at each second time marker and calculating their mean value. To check the nasal/oral ratio computer, voltages were obtained by direct measurement of the oral and nasal tracings on the calibrated oscillographic record.

Results

PSYCHOPHYSICAL SCALING. The averaged scores assigned each stimulus during backward and forward play modes of stimulus presentation and the scores from Tonar analysis are listed in Table 1. The scores from backward play ratings ranged from 25.0 to 76.5 with a mean of 58.9. Corresponding scores from forward play ratings ranged from 2.9 to 87.0 with a mean of 35.4. Inspection of the data revealed that values assigned the stimuli when the speech was played backward were generally higher than those in forward playback mode, and the total spread was less with only two exceptions.

In general the differences in absolute value between the two sets of scores were rather large. In contrast to when the scores were arranged by the relative rank of the subjects along the nasality continuum, the two sets of scaling data were very similar. This consistency is reflected in a Spearman rho correlation coefficient of 0.84 obtained between the two rank orders. This correlation is significant beyond the .01 level.

Observers were also found to be relatively consistent in the rating of different speakers when the absolute magnitude scores were rank ordered. Using a procedure outlined by Guilford (*10*), average rank order correlations among listener ratings of the responses were calculated for both forward and backward play modes. The correlation coefficients obtained were 0.83 and 0.82 for forward and backward playback respectively.

TABLE 1. Mean psychophysical scaling values, Tonar ratios, and rank orders from forward and backward playback procedures and Tonar analyses.

subject	*mean scores, backward play*	*rank order*	*mean scores, forward play*	*rank order*	*Tonar ratio*	*rank order*
1	75.1	19	87.0	20	1.341	20
2	57.0	8	13.2	4	.006	4
3	60.9	11	25.6	11	.096	5
4	56.4	7	8.9	2	.100	6
5	76.5	20	33.3	13	.491	12
6	55.1	6	19.5	9	.110	2
7	63.5	14	39.4	14	.918	17
8	61.5	12	15.8	5	.272	10
9	53.6	4	18.3	7	.823	15
10	60.5	10	21.0	10	.362	11
11	70.5	17	48.3	15	.937	18
12	54.9	5	18.4	8	.663	13
13	40.5	3	16.3	6	.213	7
14	34.2	2	12.1	3	.001	2
15	25.0	1	2.9	1	.001	2
16	71.6	18	84.7	19	.788	14
17	60.0	9	29.0	12	.241	9
18	67.0	15	70.5	17	1.245	19
19	67.5	16	60.2	16	.235	8
20	63.3	13	82.8	18	.825	16

The observers were somewhat less consistent in rerating the same subjects when the speech was presented in the backward play mode, even when the absolute magnitude scores were converted to rank order before a comparison was made. The Spearman rho correlation coefficient computed between the rank ordered listener ratings of the ten subjects with replications in backward play was 0.60, which is significant at the .05 level but seems to show rather poor agreement.

Conversely, listeners were highly consistent in their rerating of responses heard in forward play. This is indicated by a rho coefficient of 0.94, significant beyond the .01 level.

TONAR MEASUREMENT AND PSYCHOPHYSICAL SCALING COMPARISONS. The acoustic ratios computed by Tonar, shown in Table 1, ranged from 0.001 to 1.341 with a mean of 0.431. Six of the twenty speech samples had acoustic ratios of 0.1 or less. Forward play values assigned to these same responses of the six subjects ranged from 2.9 to 25.6. The scores for backward play scaling of these same responses ranged from 25.0 to 60.9. These last ratings were surprisingly high.

When the three sets of data were rank ordered, somewhat better agreement was found. A rho correlation coefficient of 0.70 was obtained between backward play scaling and Tonar scores and of 0.74 between forward play and Tonar. These correlations were significant beyond the 0.01 level.

TABLE 2. Comparison of oral and nasal voltage intensities, calculated ratios, and Tonar-computed ratios (250–2250 Hz bandwidth). (Calculated ratios are calculated from instantaneous moments at one-second intervals; Tonar-computed ratios are computed over one-second time averaging.)

subject	*voltage intensities in millivolts*		*calculated ratios*	*Tonar-computed ratios*
	oral	*nasal*		
5	47	35	.72	.75
8	140	75	.54	.53
4	79	10	.13	.12
15	94	3	.03	.03
9	105	105	1.00	1.00
10	106	33	.31	.31

To scrutinize the relationships between the absolute values assigned by the listeners in their judgments of nasality and Tonar measurements, Pearson product moment correlations were computed. The backward play vs. Tonar coefficient derived was 0.40. This was not significant. For forward play vs. Tonar scores the resultant coefficient was 0.62 which was significant beyond the 0.01 level.

TONAR MEASUREMENT AS A FUNCTION OF FREQUENCY AND INTENSITY. To examine frequency factors in nasality, speech samples of six subjects displaying ratios spread along the continuum were selected for additional analyses. The phrase *we hear that straw covers the floor of cages to keep the chill away . . .* was drawn from the recorded paragraph for use as the criterion stimulus in an intensity by frequency analysis.

Two frequency distribution sets of data were generated. These consisted of comparisons among nasal/oral intensities and ratios within a) a wideband frequency range of 250–2250 Hz, and b) twelve 200 Hz narrow bandwidths between 150 and 3350 Hz.

Table 2 displays the data from analysis of the 250–2250 Hz bandwidth by within-channel voltage intensity, calculated ratios between the two channels, and Tonar-computed ratios between the two channels. It is important to note that neither high nasal nor high oral intensities may in and of themselves result in high ratio scores. Rather, the ratio scores generated apparently resulted from variations in *either* the nasal *or* the oral channel. Furthermore, although the voltage intensities shown were calculated from instantaneous moments at one-second intervals, they agree rather closely with Tonar-computed scores derived over one-second time averaging.

The results of the narrow band frequency analysis by Tonar are displayed in Table 3 and shown graphically in Figure 1. The tabular data display no results between 250 and 950 Hz because the intensity outputs at the higher frequencies were too low to produce measurable oscillographic deflections. Ratio comparisons shown in graphic form were

TABLE 3. Comparison of oral and nasal voltage intensities, calculated ratios (calculated from instantaneous moments at one-second intervals), and Tonar-computed ratios (200 Hz bandwidth) (computed over one-second time averaging).

subject	*central frequency*	*voltage intensities in millivolts*		*calculated ratios*	*Tonar-computed ratios*
		oral	*nasal*		
5	250	—	—	—	—
	350	230	84	.37	.38
	550	6	9	1.50	1.23
	750	24	14	.58	.54
	950	21	26	1.24	1.23
8	250	—	—	—	—
	350	195	34	.17	.08
	550	9	10	1.10	.88
	750	13	11	.85	.82
	950	20	9	.45	.42
4	250	21	7	.33	.27
	350	70	5	.07	.00
	550	11	5	.45	.20
	750	3	2	.67	.18
	950	1	1	1.00	.20
15	250	250	10	.40	.00
	350	170	4	.24	.00
	550	30	3	.10	.05
	750	6	1	.17	.04
	950	4	—	—	.00
9	250	70	15	.22	.22
	350	250	100	.40	.19
	550	54	80	1.43	1.28
	750	13	25	1.92	1.78
	950	27	27	1.00	1.00
10	250	40	5	.12	.05
	350	300	60	.20	.13
	550	30	6	.20	.19
	750	24	15	.63	.66
	950	18	8	.44	.45

able to be extended to 2550 Hz. Caution should be exercised in comparing the absolute intensity values among individuals because the record level of the tape recorder was adjusted for some speakers.

Greater variability over time for the narrow band ratios is suggested by the poorer agreement between the calculated and computed ratios than was found in the wide band data.

Table 4 contrasts the Tonar computed ratios for the broadband range

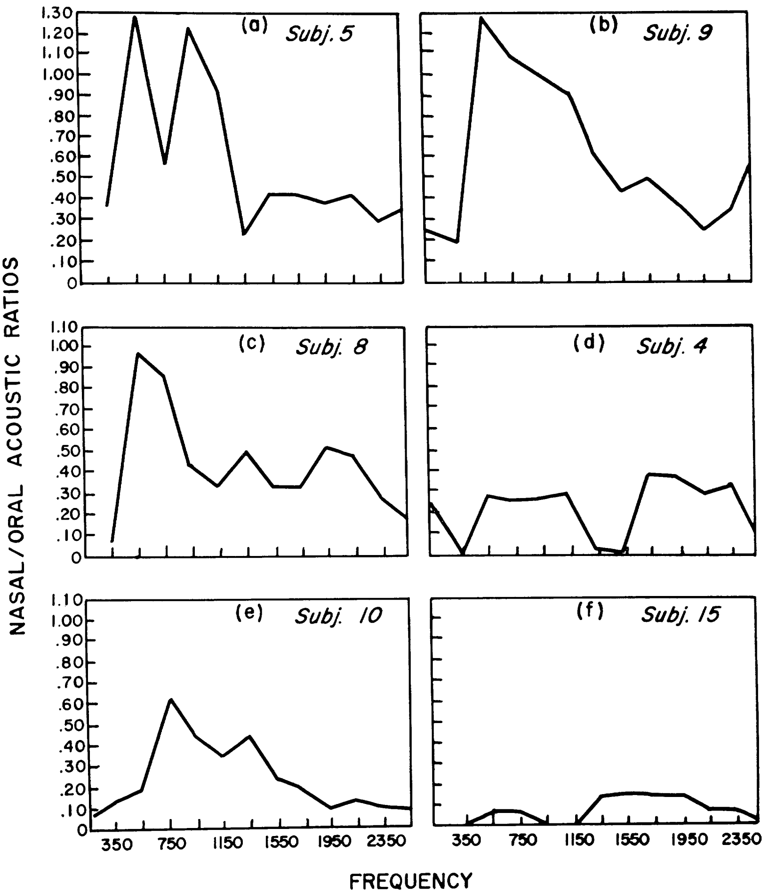

FIGURE 1. Nasal/oral acoustic ratios expressed as a function of frequency.

TABLE 4. Comparison of three-frequency ratio coverage with broadband ratio average.

subject	*350, 550, 750 Hz av.*	*250–2250 Hz average*	*difference*
15	0.030	0.030	—
4	0.127	0.121	0.007
10	0.323	0.314	0.009
8	0.592	0.533	0.059
5	0.713	0.747	0.034
9	1.085	1.000	0.015

and a mean ratio calculated from the intensities at 350 Hz, 550 Hz, and 750 Hz. A maximum difference of 0.059 was found between the two resultant values. This suggests that the 250–850 Hz frequency range is probably the principal contributor to the broadband ratio scores.

Discussion

Nasalization of vowels and other non-nasal speech sounds is the result of a bifurcation of the phonic stream such that some of the energy typically egressing through the mouth is shunted through the nasal channel. The relative distribution of the phonic energy between the nasal and oral channels is contingent upon physiological factors such as palatopharyngeal orifice size, tongue position in the oral cavity, and openness of the mouth (*4, 13*). Thus, the degree of nasal flow is regulated by the relative impedance of the oral and nasal channels.

The resonating characteristics of the oral and nasal passageways are determined by the configuration of the cavities. The presence of some nasal flow during the production of oral phonemes seems to have little effect on the final acoustic signal (*17*). However, with increasing nasal coupling the final signal undergoes acoustic changes which may be perceptually detrimental to the intelligibility and acceptability of the message the speaker is attempting to transmit (*2, 11, 12*). Evaluation of this phenomenon requires adequate instrumentation. Such instrumentation must be able to plot certain specified changes in the relative levels of nasality in the speech output and should agree in general with listener judgments of nasalization. These two facets of instrumental validation were under scrutiny in the present study of Tonar measurements and listener ratings of perceived nasality.

The findings of the study revealed that rank-ordered listener judgments of nasality in speech played in both forward and backward mode had highly significant correlations with rank-ordered scores from Tonar. This suggests that Tonar does in fact analyze parameters which affect listener ratings of nasality.

One of the interesting results of the present study was the apparent divergent findings in perceived nasality. Although a high rank-order correlation was obtained between the forward and backward play modes of speech presentation, the absolute magnitudes of the resultant ratings were markedly divergent. When heard in the forward play mode, the listeners tended to rate many of the speakers as having very little nasality. Tonar scores showed a general coincident trend toward the lower end of the continuum for these speakers. Conversely, when the same listeners rated the same speakers with the speech presented in backward play mode, all of the scores tended to cluster in higher ratings. Furthermore, comparisons between scores from rating speech nasality in backward play mode and Tonar measurements showed significant relationships to the rank order of the subjects but not between the absolute magnitudes of the scores assigned by Tonar and those by the listener. These data suggest that the conclusion by Sherman (*16*) that nasality judgments rating speech in the backward play mode are "more valid" than those for speech in the forward mode is certainly open to question.

Colton and Cooker (*7*), using backward play procedures, reported:

"...all of the normal speakers were judged to be more nasal when they spoke in a slower tempo than when they spoke in their normal tempo". They interpreted this apparent increase in nasality as the result of changes in palatopharyngeal valving linked with the reduction in tempo. An alternative explanation may well be that the preparatory set to judge "nasality" caused the observers to interpret the differences in tempo as differences in nasalization. The backward play mode of speech presentation removes the speech signal from its normal linguistic environment and disrupts its inherent distinctive features. This may precipitate distorted judgments of the perceptual field by bringing background features into the foreground. The observer is then asked to make decisions concerning a new constellation of perceptual experiences with his accumulated past experience denied him. In this setting it seems quite possible that phenomena that are relevant and those that are irrelevant to the task may be "supersummated" (*1*) into a new whole with its own new set of undefined halo effects.

Acoustical masking is another area of interest in considerations of nasality. The data revealed considerable overlap in the frequency ranges of signals emerging from the nasal and oral passages. The most prominent differential pattern evident among the narrow band data was a spectral peak which was identified around 500 Hz in the three speakers with relatively high levels of nasality. Conversely, those subjects with lower levels of nasality tended to have nasal/oral ratios that were relatively flat across the entire frequency band. This trend must be viewed as rather tenuous because of the small sample size.

The data clearly demonstrated marked frequency overlap between the two channels. Distortion in nasalized speech could thus be generated by variable nasal masking of an underlying, information-bearing oral signal. Conversely, nasality could be present but undetected if the oral signal were of sufficient strength and frequency content to mask out the effects of a nasal shunt. In this context, however, if the 500 Hz peak found in the more detailed frequency analysis of subjects with relatively higher levels of nasality were representative of all such persons, the nasalization could rise to prominence rather sharply as the level of nasal turbulence increased. Identification and classification of specific configurations of ratio curves as a function of specific poles and zeros along the frequency continuum generated by different classes of speakers await further investigation.

Our final conclusion is that Tonar seems to be a promising system for quantitative measurement of nasality. The apparent agreement between listener ratings of nasality and Tonar measurement is felt to be sufficiently close to be encouraging. It appears that the central trends of both the psychophysical and instrumental data were sufficiently congruent to make the results somewhat interchangeable. If this is so, the ease of measurement and the stability of the results would make use of the electronic system advantageous in clinical as well as in experimental contexts.

Summary

The validity of Tonar (The Oral-Nasal Acoustic Ratio) was examined by psychophysical and acoustical approaches. For psychophysical comparisons, ten trained observers judged nasality in forward and backward play modes from tape-recorded responses of twenty subjects with surgically repaired clefts of the palate. Results of acoustical analyses by Tonar were then compared with the listener ratings. Other acoustical studies were also conducted to establish tentative relationships between levels of nasality and frequency distribution of the acoustic energy. The data revealed the following. a) Significant correlations were obtained between rank-ordered listener judgments of nasality and Tonar measurements. b) Significant correlations were obtained between Tonar measurements and the absolute magnitudes of the ratings of speech in forward play but not in backward play modes. c) Nasal/oral ratios vary as a function of frequency and peak around 500 Hz as the relative level of nasality increases. d) Narrowband ratios for central frequencies of 350 Hz, 550 Hz, and 750 Hz, when averaged together, show close agreement with ratio scores from wideband analyses. e) Tonar has promising clinical possibilities.

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

References

1. Allport, F. H., *Theories of Perception and the Concept of Structure,* p. 125. New York: John Wiley & Sons, Inc., 1955.
2. Andrew, J. R., A study of the contributions of nasally emitted sounds to the perception of nasality. Presented at the national convention of the American Speech and Hearing Association, Chicago, 1967.
3. Bloomer, H., and G. Peterson, A spectrographic study of hypernasality. *Cleft Palate Bull., 5,* 5–6, 1955.
4. Bosma, J. F., and S. G. Fletcher, The upper pharynx: a review. Part II, physiology. *Annals Otolaryng., Rhinol., Laryngol., 72,* 134–157, 1962.
5. Bryan, G. A., Relationships among nasal and oral sound pressure and ratings of nasality in cleft palate speech. Ph.D. dissertation, University of Oklahoma, 1963.
6. Coleman, R. O., Jr., The effect of changes in width of velopharyngeal aperture on acoustic and perceptual properties of nasalized vowels. Ph.D. dissertation, Northwestern University, 1963.
7. Colton, H., and H. S. Cooker, Perceived nasality in the speech of the deaf. *J. speech hearing Res., 2,* 553–559, 1968.
8. Fletcher, S. G., Theory and instrumentation for quantitative measurement of nasality. *Cleft Palate J., 7,* 601–609, 1970.
9. Fletcher, S. G., Acoustic phonetics. In *The Hard of Hearing Child.* New York: Grune and Stratton, Inc., (in press).
10. Guilford, J. P., *Psychometric Methods,* p. 395. New York: McGraw-Hill Book Co., 1954.

11. Hattori, S., K. Yamamoto, and O. Fujimura, Nasalization of vowels in relation to nasals. *J. acoust. soc. Amer.*, *30*, 267–274, 1958.
12. House, A. S., and K. M. Stevens, Analog studies of the nasalization of vowels. *J. speech hearing Dis.*, *21*, 218–232, 1956.
13. Isshiki, N., I. Honjow, and M. Morimoto, Effects of velopharyngeal incompetence upon speech. *Cleft Palate J.*, *5*, 297–310, 1958.
14. Pierce, B. R., Nasal resonance differences resulting from speech appliance modifications in cleft palate adults. Ph.D. dissertation, Northwestern University, 1962.
15. Shelton, R. L., Jr., A. W. Knox, W. B. Arndt, Jr., and M. Elbert, The relationship between nasality score values and oral and nasal sound pressure level. *J. speech hearing Res.*, *10*, 549–557, 1967.
16. 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.
17. Warren, D. W., and F. A. Hoffman, A cinefluorographic study of velopharyngeal closure. *Plastic reconstr. Surg.*, *28*, 663–674, 1961.
18. Weiss, A. L., Oral and nasal SPL as related to judged severity of nasality. Ph.D. dissertation, Purdue University, 1954.