

Clinical Judgment of Hypernasality in Cleft Palate Children

LAWRENCE J. BRADFORD, M.A.

ALTA R. BROOKS, M.A.

RALPH L. SHELTON, Jr., Ph.D.

Kansas City, Kansas

Speech clinicians, dentists, and physicians working in the area of cleft palate habilitation frequently judge whether or not a client's voice demonstrates excessive nasality. The judgment made may influence recommendations for surgical or prosthetic treatment of the palatopharyngeal mechanism or for speech lessons.

Research has indicated that panels of pre-trained judges can rate nasality reliably under experimental conditions which include utilization of recorded speech samples. In a study by Lintz and Sherman (6), 35 speech pathology students rated the voice quality of 10 nasal and 10 non-nasal speakers on a seven-point, equal appearing intervals scale. Judgments were made from tape recordings shortly after a special training session. A correlation of .89 was obtained for median scale values of 100 repeated judgments.

Using the same scaling procedure, Spriestersbach (11) found that pre-trained groups of upperclassmen and graduate students in speech pathology and audiology can reliably rate nasality. Nasality was rated from taped samples, 30 seconds in length, of cleft palate speech when played both backward and forward. The correlations for median scale values of repeated nasality ratings was .96 for forward tape play and .90 for backward play.

In a study designed to investigate the influence of speech context upon perceived nasality, Spriestersbach and Powers (12) asked 30 students in speech pathology to rate vowel and conversational speech samples of cleft palate subjects. Nasality was rated on a seven-point, equal appearing intervals scale. The correlation coefficient for two sets of median scale values for conversational speech ratings was .97; the correlation for two sets of median scale values for vowel production was .81.

One study has indicated that single judges are also reliable in rating nasality under controlled conditions. Weiss (15) used the intraclass correlation procedure to determine if nasality could be scaled reliably by single

Mr. Bradford, Research Assistant, Miss Brooks, Research Associate, and Dr. Shelton, Associate Professor, are affiliated with the Hearing and Speech Department, University of Kansas School of Medicine. This investigation is part of a research program sponsored by the Easter Seal Research Foundation.

judges. Judgments were made from taped samples of speech by the method of paired comparisons after pre-training was given to the judges. The intraclass correlation coefficient for individual judgments was .86.

Although the reliability of nasality judgments has been shown to be adequate in research settings wherein judges with specific and relatively immediate pre-training worked in quiet environments, no studies have been found which consider the reliability of individual clinicians assessing nasality in the clinical setting. Simpson (10), however, reported that she and another teacher of the deaf in training agreed only 52 % of the time in their independent attempts to identify which of 21 deaf children did or did not possess extreme hypernasality. They had no special pre-judgment training. The practicing speech clinician may never receive special pre-training immediately before making judgments of nasality. Tape recordings of voices representing various scale values of hypernasality could be used for such training, but they are not available commercially and the collection and preparation of such material is very difficult. Also, speech clinicians probably make judgments under relatively poor acoustic conditions. Thus, the research work that has been done does not support the generalization that speech specialists working in clinic settings will agree with one another or with themselves about the presence or extent of hypernasality.

This study was designed to investigate the reliability of hypernasality judgments made in a clinical setting without special pre-training. The specific purposes were: a) to determine the reliability of experienced and inexperienced judges, individually and as groups, in rating perceived nasality from a sample of spontaneous speech and during the production of the /a-i/ vowel combination and b) to analyze recommendations for help made by the judges on the basis of the speech samples they obtained. Relative reliability of judgments based on the spontaneous speech sample and on the vowels /a-i/ will be noted.

In performing the /a-i/ test (or task), the client produces the pair of vowels repeatedly while the examiner alternately closes the client's nares by digital pressure and opens them again by a release of the pressure. (It is important to note that this sample involves the alternate production of the two vowels /a/ and /i/ and is not to be confused with the diphthong /ai/.) In a person with hypernasality, the nasal resonance presumably will be more apparent when the nostrils are closed than when they are open. This test has been discussed in some detail by Brodnitz (1, p. 100) and Moser (7). A variation of this test is recommended for use by Johnson, Darley, and Spriestersbach (4, p. 157).

Procedure

Seventeen children, 10 boys and seven girls, attending the diagnostic cleft palate clinic at the Kansas University Medical Center were used as subjects in this study. All children six years old or older who attended two clinic sessions were included as subjects. At the time of testing, the mean

age of children was nine years seven months and the range was from six years to 14 years seven months. No attempt was made to control such variables as articulation, facial grimace, appearance, or type or adequacy of physical management. The children were seen in the clinic for an annual routine check-up. Some children continue to come to the clinic after they have acquired normal speech skills.

Two groups of four judges each were used to rate the subjects. Experienced judges were practicing speech clinicians who held at least M.A. degrees in speech pathology, and who had three or more years of clinical experience, including extensive experience with cleft palate children. Three of these judges held advanced clinical certification with the American Speech and Hearing Association. Inexperienced judges were graduate students or were recent recipients of M.A. degrees with less than three years of clinical experience and with minimal experience with cleft palate children. They had no clinical certification.

The judges were asked to rate each subject on perceived nasality and to indicate whether or not the child should receive help on the basis of a) a spontaneous speech sample and b) the /a-i/ task. Nasality ratings were made on a seven-point scale with *zero* representing no hypernasality and *six* representing extreme hypernasality. The recommendation that the child needed help for his nasal resonance was made by a simple 'yes' or 'no' judgment. The meaning of 'help' was purposely left ambiguous.

The order for testing the children was counterbalanced for each judging group so that one-half of the children were rated first from the /a-i/ task and the remaining children first from the spontaneous speech sample. Thus, two judges from each group rated one-half of the children in the /a-i/ task—spontaneous speech testing order and the remaining children in the spontaneous speech—/a-i/ task testing order. The other two judges rated the same children in the reverse testing order. The subjects were tested by each judge individually, so that each child repeated the entire performance eight times. The first speech sample obtained by a judge, whether spontaneous speech or /a-i/, was rated for nasal resonance and the decision regarding help was made before the second speech sample was elicited. This testing procedure was used in an attempt to cancel a possible order effect of test administration on the judges' ratings.

No special training was given to the judges other than to explain the /a-i/ test since the purpose of this study was to determine reliability of judgments without special pre-training or instructions. It was intended that this procedure approximate the judgment process usually followed by speech clinicians as closely as possible. Judges were given various magazine covers to use in eliciting the spontaneous speech sample if they wished. All testing was done on an individual basis in small rooms in a pediatric clinic. The rooms were not sound treated and sounds of equipment and children were variably present. No attempt was made to observe the judges. Since each judge had to test eight or nine children during a limited period of time

TABLE 1. Intraclass correlation coefficients for judgments of nasality by experienced and inexperienced judges individually and as groups. Results are included for a spontaneous speech sample and for the /a-i/ test.

	<i>Spontaneous Sample</i>	<i>/a-i/ Test</i>
Reliability of Individual Raters:		
Experienced14	.25
Inexperienced25	.33
Reliability of Average Ratings for group:		
Experienced41	.57
Inexperienced57	.67

they could spend little more than approximately five minutes with each child. An assistant helped the judges locate the subjects.

Results

Ebel's intraclass correlation procedure was used for evaluating the reliability of nasality judgments made by individual judges in each group and under each speech condition (3).¹ The intraclass correlation also provided an estimate of the reliability of average ratings by groups of judges for each speech condition. The formula which removes between-raters variance from the error term was used in each calculation. Removal of this variance from the error term prevents distortion of the coefficients by differences between raters. When comparisons between raters are not made in practice, as would occur most frequently with speech clinicians' assessment of hypernasality, the between-raters variance should be removed from the error term in order to determine accurately the ability of individuals. This procedure, therefore, permits a reliable determination of the individual's rating ability which is comparable to the situation facing a practicing speech clinician.

The intraclass correlations for individual experienced judges were .14 for the spontaneous speech sample and .25 for the /a-i/ task. Coefficients for the experienced judges as a group were .41 and .57 for the two voice samples. The inexperienced individual judges obtained intraclass correlation coefficients of .25 for spontaneous speech and .33 for the /a-i/ task. Corresponding coefficients for the inexperienced judges as a group were .57 and .67. These data are summarized in Table 1. It should be noted that while there were no real reliability differences between the two groups, the inexperienced judges did appear to be somewhat more reliable.

The recommendations for help for perceived nasal resonance were compared with the scale values assigned. The results are shown in Figure 1.

¹ For additional material about intraclass correlation techniques, the reader is referred to studies of Sherman and Cullinan (9) and by Stitt and Harrington (13), in which intraclass correlation results are compared with other kinds of correlation coefficients.

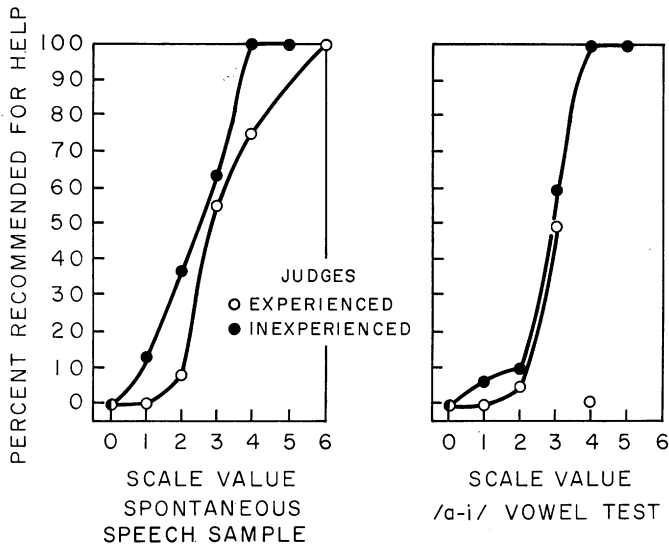


FIGURE 1. Percentages of subjects for whom 'help' was recommended at each scale value. Results are plotted for inexperienced and experienced judges using both spontaneous speech samples and the /a-i/ task. On the /a-i/ task an experienced judge assigned a scale value of four to one subject; help was not recommended.

The inexperienced judges recommended help more frequently than did the experienced judges for both types of speech samples. Both groups of judges tended to recommend help for hypernasality less frequently on the basis of the /a-i/ test than on the spontaneous speech sample. This difference is especially noted in the recommendations of the experienced judges' ratings.

Discussion

In this study, experienced and relatively inexperienced speech clinicians were asked to judge the voices of a group of cleft palate children for hypernasality. Judgments were made in the same rooms that were used in conducting a diagnostic cleft palate clinic in which the judges had participated as staff or students. Instructions to the judges were kept to that minimum which was necessary to elicit the desired judgments. Thus, the judges were expected to behave much as they would in any evaluating situation. The findings of poor inter and intra judge reliability combined with the lack of any evidence from other studies that clinicians agree well in judging hypernasality indicates that management decisions should not be based on judgments of nasality made under conditions similar to those utilized in this study. Also, until other evidence is available, one might generalize from this study to nasality assessments not involving cleft palate samples. We recognize that the clinician will be faced from time to time by individuals with nasal voices and no other speech problem. To the extent that the clinician and client have difficulty discriminating between hypernasal

and normal voice quality, however, they may be expected to have difficulty in pursuing a voice correction program.

Fortunately, articulation is known to be a primary factor in speech intelligibility (5, 8) and articulation testing is known to be reliable (2). Thus, articulation testing can provide a reliable basis for planning speech lessons that will improve intelligibility. Nevertheless, improved techniques of nasality assessment in the clinic should be developed. Perhaps tape recordings of spontaneous speech samples and of the /a-i/ task at various known scale values could be used to 'tune' or 'calibrate' the ear of the clinician shortly before he undertakes to judge or scale nasal resonance. This hypothesis should be put to experimental test.

The low reliability of the clinical judgments of nasality in this study may be partially accounted for by the sample of children used. The scale values assigned were clustered at the lower end of the continuum, near a median of 1.50. Lack of contrast in voice quality among the subjects would make the judgment task more difficult and would reduce reliability. Nevertheless, this is the kind of problem the practicing clinician faces. Since each subject had to repeat his performance eight times, variability in performance would also reduce judge agreement.

No statistical comparison of the results obtained from the spontaneous speech sample with those from the /a-i/ test was made. Each method should be studied further. The spontaneous speech sample is important because it approximates the situation in which human beings converse. The /a-i/ task had the advantage of freedom from contamination by articulation variables which have been shown to influence nasality judgment (11, 14). It is noted that whereas higher reliability was obtained from the /a-i/ task, both groups of judges tended to recommend help less frequently from this kind of speech sample. Greater reliability might have been obtained from the /a-i/ task had the instructions given to the judges been more comprehensive.

The finding that the inexperienced judges more frequently recommended help than the experienced judges reflects a tendency for inexperienced clinicians to err in the direction of caution. If this observation is confirmed in other studies, it should be considered in the training of speech specialists.

Summary

Hypernasality judgments were made by experienced and inexperienced speech clinicians of the speech of 17 cleft palate children six years of age and older. A spontaneous speech sample and the /a-i/ test were used to make the assessment. Two groups of judges, experienced and inexperienced, made ratings of perceived nasality on a seven-point scale and indicated a *yes* or *no* judgment regarding the need for help. These data were analyzed to determine the reliability of individual and average judgments of perceived hypernasality without special pre-training, in a clinical setting.

Neither experienced nor inexperienced judges were able to rate hyper-

nasality reliably within the conditions of this study. Therefore, the conclusion was reached that much caution should be used in making decisions concerning speech lessons or physical management on the basis of hypernasality ratings made in a clinical setting without special pre-training.

Acknowledgements: The authors wish to thank Dr. William B. Arndt, Dr. Kenneth L. Moll, and Mrs. Betty Bellerose for providing statistical advice and editorial assistance.

*University of Kansas Medical Center
Rainbow Boulevard at 39th Street
Kansas City 3, Kansas*

References

1. BRODNITZ, F. S., *Vocal Rehabilitation*. American Academy of Ophthalmology and Otolaryngology, 1961.
2. BYRNE, MARGARET D., SHELTON, R. L., JR., and DIEDRICH, W. M. Articulatory skill, physical management and classification of children with cleft palates. *J. speech hearing Dis.*, 26, 326-333, 1961.
3. EBEL, R. L., Estimation of reliability of ratings. *Psychometrika*, 16, 407-424, 1951.
4. JOHNSON, W., DARLEY, F. L., and SPRIESTERSBACH, D. C., *Diagnostic Methods in Speech Pathology*. New York: Harper and Row, 1963.
5. JORDAN, E. P., Articulation test measures and listener ratings of articulation defectiveness. *J. speech hearing Res.*, 3, 303-319, 1960.
6. LINTZ, LOIS B., and SHERMAN, DOROTHY, Phonetic elements and perception of nasality. *J. speech hearing Res.*, 4, 381-398, 1961.
7. MOSER, H. M., Diagnostic and clinical procedures in rhinolalia. *J. speech hearing Dis.*, 7, 1-4, 1942.
8. SHAMES, G. H., and FISHER, J., The relationship of types of articulation errors to intelligibility of speech. *Speech path. Ther.*, 1, 30-34, 1960.
9. SHERMAN, DOROTHY and CULLINAN, W. L., Several procedures for scaling articulation. *J. speech hearing Res.*, 3, 191-198, 1960.
10. SIMPSON, MARY H., Consonantal nasal pressure in the speech of children with hearing impairments. M. S. thesis, Univ. Kansas, 1963.
11. SPRIESTERSBACH, D. C., Assessing nasal quality in cleft palate speech of children. *J. speech hearing Dis.*, 20, 266-270, 1955.
12. SPRIESTERSBACH, D. C., and POWERS, G. R., Nasality in isolated vowels and connected speech. *J. speech hearing Res.*, 2, 40-45, 1959.
13. STITT, C. L., and HUNTINGTON, DOROTHY A., Reliability of judgments of articulation proficiency. *J. speech hearing Res.*, 6, 49-56, 1963.
14. VAN HATTUM, R. J., Articulation and nasality in cleft palate speakers. *J. speech hearing Res.*, 1, 383-387, 1958.
15. WEISS, A. L., Oral and nasal sound pressure levels as related to judged severity of nasality. Ph.D. dissertation, Purdue Univ., 1954.