A Rating Scale for Evaluation of Video Tape Recorded X-Ray Studies

BETTY J. McWILLIAMS NEELY, Ph.D.
DORIS P. BRADLEY, Ph.D.
Pittsburgh, Pennsylvania

The purpose of this paper is to describe a rating scale technique for use in the analysis of video tape data. At the present time video tape cannot be played at slow speeds or frame by frame. Therefore, the better known methods for reducing cineradiographic data to quantitative results are not appropriate for analysis of video tape recorded information. The rating scale provides one means of evaluating speech structures and their relationship to each other during speech.

History of Technique

Investigation of speech structures by means of x-ray images has been possible since shortly after the discovery of roentgen rays in 1895. The usefulness of the technique was limited by factors such as exposure of patients to radiation, difficulty with methods of analysis, and numerous problems in synchronization of sound with the visually recorded image. In addition, there were problems in obtaining clearly defined x-ray images of soft tissue and bony structures on the same film and both types of tissue were important in studying speech. Early x-ray studies of speech behavior attempted to solve this problem of tissue definition by placing a metal chain along the midline of the tongue to aid in visualizing it. Such procedures, however, were considered to interfere with the production of speech sounds, and the information gained from these studies had to be interpreted with caution.

In spite of such limitations, radiographic techniques have been extremely helpful in observing the structures used for speaking. These techniques made it possible to acquire knowledge that might otherwise have remained inaccessible. Improvements in radiographic techniques have developed rapidly. Today, less than 70 years since x rays were identified, numerous radiographic procedures are in use. Television tape recording of fluoroscopic images is one of the most recent developments. The system employed

Presented at the 1963 Convention of the American Cleft Palate Association, Washington, D. C.

Dr. Neely is Associate Professor and is on the Administrative Committee of the Cleft Palate Research Center, University of Pittsburgh. Dr. Bradley is Director of Speech Clinic, Children's Hospital of Pittsburgh.

This investigation was supported by Research Grant DE-01501, National Institutes of Health, United States Department of Public Health.

at Children's Hospital of Pittsburgh utilizes an Image Orthicon television camera which is about 1,000 times as sensitive to light as other cameras. It scans the fluoroscopic screen at a speed of approximately 525 lines per second. The light images are converted into electron images which subsequently go through a five stage amplifier. According to Paul and Juhl (*8*), this Orthicon tube is capable of increasing screen brightness up to 50,000 times that of an ordinary fluoroscopic screen.

Girdany, Gaither, and Darling (*6*), in cooperation with the X-ray Division of Westinghouse, optically coupled the Image Orthicon television camera to a Philips nine-inch amplifier. The system also employs an electronic magnifier which doubles the size of the television image. This phase of magnification is accomplished by scanning twice the usual number of lines per second. It enhances considerably detail in small areas, and has proven particularly helpful for studying movements of the tongue and soft palate in speech. This equipment allows satisfactory television recording and viewing of speech structures with settings which range from .13 to .5 ma and from 60 to 80 kv.

The advent of recording x-ray images on magnetic video tape provides many advantages. Jutras (*7*) says that it implies possibilities far beyond the restricted capabilities of sensitized emulsions, and acclaims it as an excellent means of studying the effects of motion. Present advantages include remarkable reductions in radiation exposure to the patient, which in turn makes it possible to view longer periods of speech and oral activity. This Televex system also provides other advantages such as simultaneous viewing of the picture being recorded, immediate playback of the recorded images, and synchronization of recorded audio and visual signals. These advantages make this new radiographic technique very desirable for use in studying speech structures and their movements.

Numerous methods for the analysis of data from cineradiographic motion pictures have been devised. DeCastro (*4*, p. 104) points out that the capacity for observation is limited to our sense of vision. He further states that:

> The motion of an organ, although it does not express the totality of a vital function, nevertheless represents the externalization of the motor function, the only one which may be registered on cinematographic film for later observations.

DeCastro recommends several procedures for the analysis of cineradiographic films. These are: (a) subjective observation of motion at regular and slow projection speeds; (b) subjective observation by an individual observer or by groups of observers in frame-by-frame projection; and (c) tracing the outlines of organs on graph paper as observed in the fixed projections. This last method makes possible a mathematical analysis of movements and forms of organs. Variations and combinations of these three types of analyses are widely accepted in studies of speech structures and their movements.

As previously stated, video playback units are not yet practically equipped for slow motion or for still projection. Attempts at developing slow motion projection have resulted in distortion of the images. Brown (*2*) reports several new developments in camera tubes which he recently observed in England. Among these is a new Vidicon Television camera which has the ability to resolve 2,000 lines of picture information per second. Other developments include experimental tubes which store pictures for a number of seconds and thus provide still projection. This latter development holds considerable potential for providing still projection of teleradiographic signals that will be useful in quantifying data. However, all this is in the experimental stage of development and is not yet available for general use.

Currently, regular playback speed must be accepted if the technique of direct analysis of video tape recorded x-ray images is used. The alternative to direct analysis involves the transfer of x-ray images from the video tape to 16 millimeter motion picture film via a kine-recording camera. This process is expensive and presents problems in synchronizing sound with visual signals. The quality of the x-ray images obtained on the 16 millimeter film transfers is often inferior to the quality on video tape. This is particularly true when commercial developing processes are used. The present situation, then, makes it highly desirable to develop some means of evaluating and describing the form, the movement, and the relationships among speech structures by viewing the video tape played at regular speeds.

Most of the previous attempts to evaluate film projected at regular speeds have relied on verbal descriptions of the phenomenon observed. Fletcher (*5*), Truby (*10*), and Powers (*9*), are among those who have used such methods of analysis. Fletcher (*5*) reported, from his observations, certain sequences of action in the pharyngeal wall. Truby (*10*) described the movements of the articulators in the production of consonantal complexes. These descriptions were made after viewing the film projected at various speeds and also frame by frame. Powers (*9*) concluded that his attempts to compare groups of subjects by qualitative assessment of the cinefluorographic films were not productive. Calnan (*3*) devised a descriptive scale for rating the mobility of the soft palate. His scale consisted of five descriptive statements with a numerical rating assigned to each statement. He used it to describe soft palate mobility from standard lateral x-ray films made while the subject phonated a vowel and while speech structures were at rest.

Development of Rating Scale

Since qualitative assessment of video tape appears to be the most feasible method of analysis available at this time, it was decided to attempt to quantify and objectify such analysis. Many hours of observation of video

tapes, together with a review of the literature, resulted in the selection of six major variables which seemed to affect articulation and voice quality either singly or in combination. A rating scale (Figure 1) was developed for evaluating the six variables. Statements, descriptive of observed data, were written for each variable and numerical values were assigned to each statement. Use of the rating scale involved viewing the video tape projected at regular speed and assigning the number of the descriptive statement that best represented the observation. The ratings were made from tapes with subjects performing various tasks and with a variety of head positions (see Table 1).

It might be argued that such an analysis is subjective and gross. In reality the major difference between objective measures and subjective ones is a matter of consistency. In other words, any analysis becomes objective when it is possible to demonstrate that the observer's feelings are not reflected in the judgments, that similar judgments can be made repeatedly by the same observer, and that various observers can agree upon the evaluation of a given set of data.

Training sessions in the use of the rating scale presented in Figure 1 demonstrated that the above criteria for objectivity were achieved. It is the opinion of the authors that training sessions are important and necessary before ratings by different individuals can be considered comparable.

Discussion of disagreements on ratings indicated several sources of error. These were (a) judges rated minimum or average movement instead of maximum movement; (b) judges' conceptions of long, short, and average varied markedly; and (c) the judges' conceptions of the descriptive statements varied. Errors of this nature were minimized by training sessions however. Generally, various judges demonstrated the ability to assign the same rating to the same observation. Information on judges' ability to repeatedly assign the same rating to the same observation is not yet available.

Table 1 provides information concerning agreement between the two authors on ratings for films taken during a variety of conditions. These data were compiled by Bradley (*1*) using 16 individuals whose speech structures had been modified by posterior pharyngeal flap surgery. Pearson product moment correlations between the two judges' ratings for the 19 items on the scale ranged from .51 to .92. The percentage of agreement on ratings ranged from a low of 38% to a high of 88%. Ratings presented in Table 1 are the result of 288 judgments that were made by each author. The authors assigned the same rating to 187 of those judgments. This represents 65% agreement. Only five times did the ratings differ by more than one point on the scale. It is recognized that further investigation is needed regarding intra-judge reliability and test-retest reliability of the scale. It is also desirable to compare judgments from the rating scale to measurements from tracings. High reliability of this nature would support the use of the scale in both research and clinical assessments.

FIGURE 1. Descriptive statements and numerical values assigned those statements for six variables used in the rating scale for video tape–recorded x-ray studies.

Variables	*Values*	*Descriptive Statements*
Palatal pharyngeal contact	1	total blending
	2	partial blending
	3	touch but no blending
	4	close approximation (no touch)
	5	closure not achieved, little
	6	closure not achieved, medium
	7	closure not achieved, great deal
	8	no movement in direction of closure
Soft palate thickness	1	very thick
	2	thick
	3	average
	4	thin
	5	very thin
Soft palate length (variable not used or subjects with pharyngeal flaps)	1	very long
	2	long
	3	average
	4	short
	5	very short
Extent of palatal pharyngeal vertical contact	1	very long
	2	long
	3	average
	4	short
	5	very short
	6	none
Position of tubercle in relation to palatal closure	1	well above
	2	slightly above
	3	in the middle (at level)
	4	slightly below
	5	well below
Position of palatal closure in relation to hard palate	1	well above
	2	slightly above
	3	at the level
	4	slightly below
	5	well below

Discussion

In its present form the rating scale appears to have several advantages over verbal descriptions of observations made from viewing the video tapes. The most important of these is consistency in reporting. As an example, the speech pathologist and radiologist made clinical notes about the six variables being studied at the time that the video tapes were made. Attempts to analyze these descriptions indicated that all variables were not described by either investigator. The following example illustrates this inconsistent reporting.

TABLE 1. Correlation coefficients and percentages of agreement between two judges' ratings of five variables observed during a variety of conditions for 16 subjects.

Variables and Conditions	*r*	*% age of Agreement*
Palatal Pharyngeal Contact		
Connected speech	.92	62
Blowing	.80	75
Syllables, head upright	.80	81
Syllables, head flexed	.70	62
Syllables, head extended	.67	88
Soft Palate Thickness		
Connected speech	.63	44
Palatal Pharyngeal Vertical Contact		
Connected speech	.73	62
Syllables, head upright	.78	62
Syllables, head flexed	.83	38
Syllables, head extended	.78	44
Tubercle in Relation to Palatal Closure		
Connected speech	.78	69
Syllables, head upright	.81	69
Syllables, head flexed	.77	56
Syllables, head extended	.86	69
Palatal Closure in Relation to Hard Palate		
Connected speech	.72	81
Syllables, head upright	.78	81
Syllables, head flexed	.62	62
Syllables, head extended	.51	62

Speech pathologist: Very little movement in the palate. Flap very thin, palate very thick. Closure on blowing. Closure occurs above the tubercle.

Radiologist: Repaired palate with flap. Touch closure on *s* and *g* sounds. Flap thin, palate thick, and short. Good movement of the pharyngeal wall.

Such inconsistency in reporting makes it impossible to compare individuals with regard to specific variables. At the same time, clinical notes are very valuable and often include observations on variables not assessed by the rating scale. A review of clinical notes suggests several more variables that should be measured with the rating scale.

Another advantage of the rating scale is quantification of the observed data. This makes possible comparisons among several variables on the same individual, comparison of one group to another, and numerous other studies.

The rating scale is a much faster method of analysis than the tracing technique. It requires approximately 10 minutes to make all 19 judgments on a subject. This time factor makes the rating scale technique feasible for daily use in clinical assessment of video tapes. More time-consuming methods of analysis can be reserved for use in borderline cases when decisions are not justified from the rating scale alone.

Summary

This paper described a rating scale technique for the analysis of video tape data relating to the form, movement, and relationships among speech structures. Since video playback units are not equipped for practical slow motion or for still projection, the better known methods for reducing data to quantitative results are not satisfactory for analysis of video tape recorded information. The superior quality of the x-ray images on video tape as compared to the 16 millimeter film transfers make it desirable to develop some means of evaluating the data from video tape played at regular speed. The present rating scale consists of descriptive statements applied to six variables. These were palatal pharyngeal contact; soft palate thickness; soft palate length; palatal pharyngeal vertical contact; tubercle in relation to palatal closure; and palatal closure in relation to hard palate. Correlation coefficients between the authors' ratings were reported for each item of the scale and for each of several situations in which ratings were made. Sixteen subjects who had pharyngeal flap procedures were used. The obtained reliability coefficients ranged from .51 to .92. The rating scale technique has considerable value as a method of quantifying data as observed from video tape projected at regular speed. It is especially useful for clinical evaluations and is considered to have an important place in research methodology applied to Televex studies.

Cleft Palate Research Center
University of Pittsburgh
Pittsburgh, Pennsylvania

References

1. Bradley, Doris P., A Teleradiographic Study of Speech Structures of Individuals with Posterior Pharyngeal Flaps. Unpublished Ph.D. dissertation, Univ. of Pittsburgh, 1963.
2. Brown, M., Notes. *Newsletter for Council on Medical Television, 4*, 6, 1963.
3. Calnan, J. C., The mobility of the soft palate: radiological and statistical study. *Brit. J. plastic Surg., 14*, 33–38, 1961.
4. DeCastro, M. J., Fundamental principles in the application of cineroentgenography as an auxiliary method to roentgen diagnosis. *Amer. J. Roentgen, 57*, 103–114, 1947.
5. Fletcher, S. G., A Cinefluorographic Study of the Movements of the Posterior Wall of the Pharynx during Speech and Deglutition. Unpublished M.A. thesis, Univ. of Utah, 1957.
6. Girdany, B. R., Gaither, E. S., and Darling, D. B., Large screen image amplification with closed circuit television employing television tape recorder. *Radiology, 77*, 286–291, 1961.
7. Jutras, A., Teleroentgen diagnosis by means of video tape recording. *Amer. J. Roentgen, 82*, 1099–1102, 1959.
8. Paul, L. W., and Juhl, John H., *The Essentials of Roentgen Interpretation.* New York: Paul B. Hoeber, Inc., 1959.
9. Powers, G. R., Cinefluorographic investigation of articulatory movements of selected individuals with cleft palates. *J. speech hearing Res., 5*, 59–69, 1962.
10. Truby, H. M., Acoustic cineradiographic analysis with considerations special reference to certain consonantal complexes, *Acta Radiol.*, Suppl. 182, 1960.