The Effect of Filming Rate and Frame Selection in Cinefluorographic Velopharyngeal Analysis

SARA S. SPARROW, M.A.
BYRON G. BROGDON, M.D.
KENNETH R. BZOCH, Ph.D.
Gainesville, Florida

Two commonly used radiological procedures, lateral-head radiography and laminagraphy, have provided useful information in the investigation of the speech mechanism. However, since previous studies (*4, 6*) have shown that the physiological characteristics of speech sounds are dependent upon their phonetic environment and that the transition between sounds is an important variable in the movement of the speech mechanism, the static nature of these techniques has imposed limitations on their use. In recent years cinefluorography has become a fundamental investigative technique since this method permits study of the articulatory mechanism while it is moving.

Initially, cinefluorographic utilization was confined primarily to qualitative comparisons of the motion pictures obtained and little attempt was made to analyze individual frames or to measure the movements of the various structures (*3, 4, 5, 9*). Recently, the need for studying the cinefluorographic film by a quantitative process has been recognized (*1, 2, 6, 8, 10*). Powers (*10*) found that: '... most of the differences between subjects obtained from measurements of the single frame tracings were not apparent from gross evaluation of the films or inspection of individual frames.'

The major problem of quantitative analysis of cinefluorographic film is the large amount of information derived from only a few seconds of filming (*6*). It has been noted that the large number of frames to be traced and measured places practical limitations on the number of subjects who may be studied and the size of the speech sample which may be used (*2, 10*).

A related problem in the utilization of cinefluorography is selection

---

Mrs. Sparrow is Coordinator of Speech Therapy, Orange County Schools, Orlando, Florida; Dr. Brogdon is Associate Professor of Radiology, Johns Hopkins University School of Medicine and Radiologist-in-charge, Division of Diagnostic Roentgenology, The Johns Hopkins Hospital, Baltimore, Maryland; and Dr. Bzoch is Associate Professor of Speech and Head, Department of Communicative Disorders, College of Health Related Services, University of Florida, Gainesville, Florida. This study was supported in part by Research Grant D-1389 from the National Institutes of Health.

of the proper filming rate (*6, 10*). Obviously, faster filming rates not only increase the amount of time necessary to analyze the film (since more frames are acquired per second), but also substantially increase the amount of radiation exposure sustained by the subject.

The present study was designed to investigate the possibility of shortening the time necessary in quantifying cinefluorograms of velar function in two ways: a) determination of the minimum number of frames that must be analyzed in order to produce reliable data, and b) comparison of data reduction from two filming rates (30 fps and 60 fps).

Material and Method

Subjects. A group of 11 men and nine women served as initial subjects in this study. All were either students or staff members at the University of Florida, ranging in age from 19 to 43 years. The subjects were screened by two speech pathologists and were judged to exhibit no abnormalities in their speech structures or in their manner of speaking. No subjects were included who were pregnant or who had undergone extensive x-ray diagnostic procedures or radiation therapy.

Since this study was concerned only with methodology and was not concerned with investigation of certain features or deviations of behavior, massive amounts of duplicative data were not considered essential. Consequently, films of only three subjects were selected for detailed analysis. Selection was on the basis of film quality, those chosen giving the best definition of speech structures.

Speech Sample. The speech sample consisted of the following vowel-consonant-vowel (VCV) combinations: /ɑpɑ/, /ɑbɑ/, /ɑmɑ/, /ɑsɑ/, /ɑnɑ/, and /ɑtɑ/. These combinations were spoken after the identification sentence *My name is* . . . . , in which each subject used his own name. The six disyllables were said on one continuous expiration and each set of six was repeated four times. The first and last VCV combinations, /ɑpɑ/ and /ɑtɑ/, were included only for control purposes, so that all analyzed disyllables would be preceded and followed by phonation. The middle four disyllables were visually identified on the film. The sounds selected were those in which the back of the tongue is low in the mouth, thereby preventing the obscuring of the inferior surface of the velum by the tongue shadow. Nasal consonants were alternated with those containing non-nasals in order to assure maximum variations in velar movement. Subjects were instructed to speak the VCV combinations at a rate approximating normal connective speech. Attainment of normal speech was facilitated by utilization of the carrier sentence.

All subjects tended to use a neutral vowel as the first element in all disyllables and therefore to have a second vowel component of a slightly longer duration. Thus, as actually spoken, the six VCV combinations tended to be: /əpɑ/, /əbɑ/, /əmɑ/, /əsɑ/, /ənɑ/, and /ətɑ/.

Recording Apparatus. The cinefluorographic apparatus used in this

investigation is located in the Department of Radiology of the University of Florida Teaching Hospital. It consists of a Picker fluoroscopic unit with an eight-inch image intensifier, mounting a modified 16 mm Eastman Cine II camera. The equipment is synchronized so that during cineradiography radiation bursts are synchronized with the camera shutter, thereby reducing radiation dosage. A 3.5 mm filter provided additional radiation reduction. Filming speeds of 30 and 60 frames per second were utilized and these have corresponding radiation exposure times of 1/60 and 1/200 seconds respectively. Kodak Linagraph Shellburst film was utilized as it seemed to provide the best balance of film speed and definition. The cine films were processed in a Smith-Picker automatic processor. The entire cinefluorographic procedure was monitored by a Magnacord, Model M90-A, tape recorder with a standing directional Electro Voice microphone, model 644, six inches from the subject's mouth (Figure 1). Positioning and immobilization of the subjects were achieved by an adjustable lucite head positioner equipped with ear rods and a nasion bar, all movable (Figure 2). (This device was a modification of a head positioner devised in the Division of Radiotherapy at the University of Michigan Hospital.)

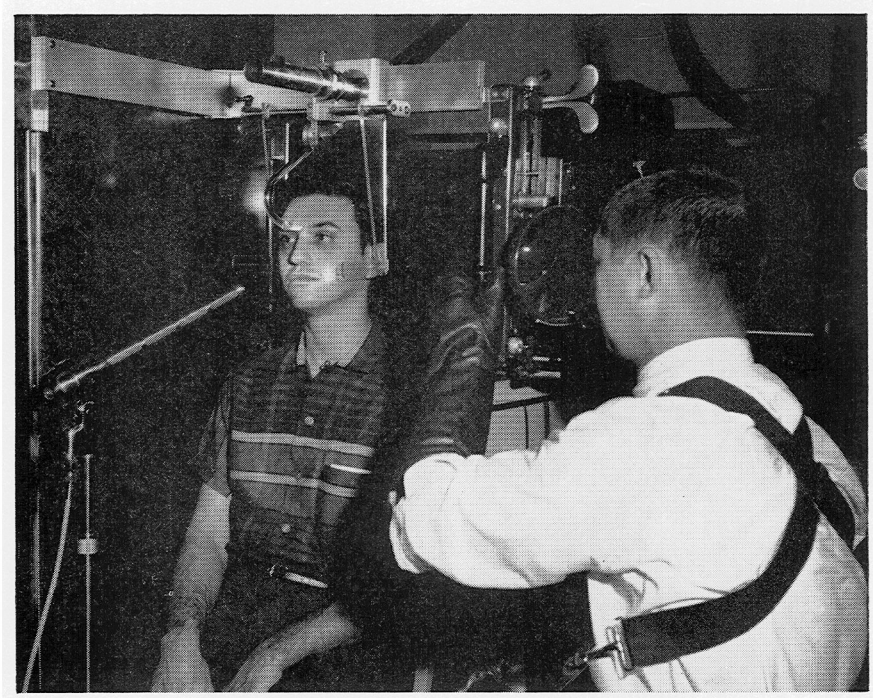

FIGURE 1. Subject and equipment set-up for simultaneous tape recording and cinefluorography of speech sample. The microphone is at left. The subject is seated between the vertical fluoroscopy table and the image intensifier. The subject's head is in a lucite positioner which is adjustable. The fluoroscopist is in position for a mirror-viewing of the intensified fluoroscopic image during filming.

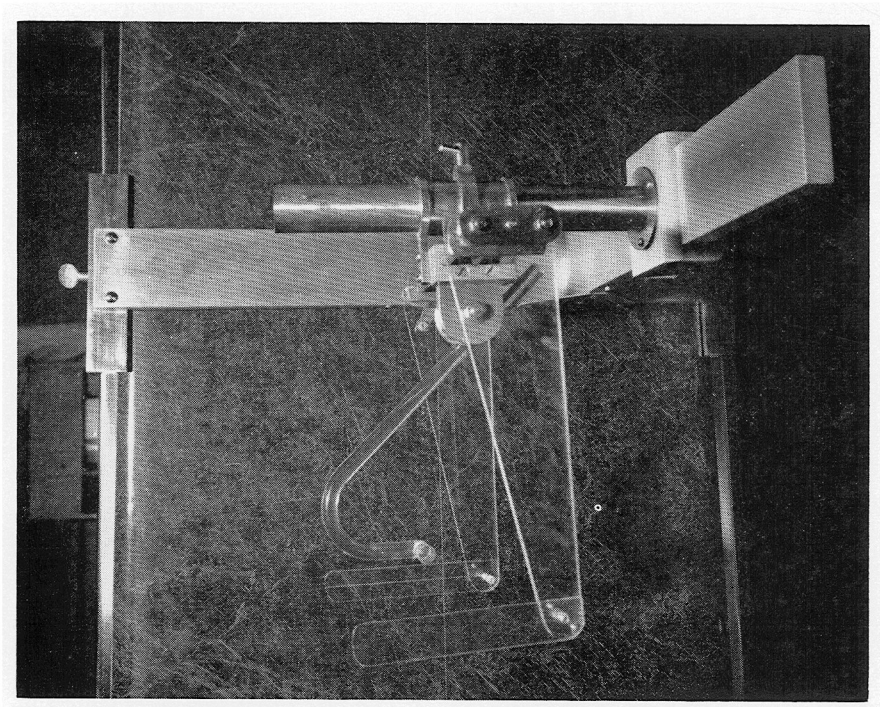

FIGURE 2. Close-up view of lucite head positioner. The positioner is adjustable at both vertical and horizontal dimensions on the fluoroscopic table. The separation between the ear rods is adjustable, as is the position of the nasion bar.

TRACING APPARATUS. The primary unit in the tracing apparatus (Figure 3) was a 16 mm L and W Photo Analyzer which features a remote control attachment that allows forward and reverse slow motion drive plus single frame projection. The analyzer was mounted at 30 degrees from the horizontal, above and behind the tracer, and the image was projected over the tracer's shoulder onto the tracing surface. The tracing surface was mounted at 60 degrees from the horizontal to allow perpendicular projection of the image at a comfortable tracing angle. The tracing surface was covered with green Ariel cover paper ※65-130 which, after experimenting with various materials, was found to give the best image resolution and contrast. The distance between the tracing surface and the projector was adjustable; therefore, the image size could be altered by moving the tracing surface either toward or away from the projector.

CINEFLUOROGRAPHIC PROCEDURE. The subject was seated on a standard chair between the fluoroscopic table and the image amplifier so that the x-ray beam was transmitted horizontally through the vertical head. The head positioner was applied as firmly as possible without producing discomfort. The subject was made to practice reading the speech sample at normal rate, pitch, and intensity. Immediately prior to filming, a stainless

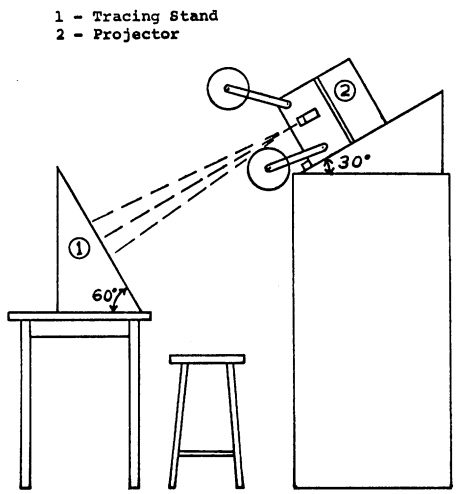

FIGURE 3. Diagramatic representation of equipment for tracing cinefluorographic frames: 1) tracing stand and 2) projector.

steel ruler perforated at centimeter intervals was placed in the mouth along the midline of the tongue, and a short cinefluorographic sequence was obtained (Figure 4). This permitted correction of image enlargement for projection and tracing. Cinefluorograms were made while the subjects read the carrier sentence and the two sequences of VCV combinations at 30 fps (frames per second). Subsequently, two sequences were read at a filming rate of 60 fps.

After the first four subjects were studied some modifications in technical factors were made, and subsequent subjects had their nasopharynx sprayed with a ¼% neosynephrine solution to shrink the mucosa and thereby improve soft tissue definition. The mucosal shrinkage was not found to alter voice quality or speech pattern. The filming time for the three selected subjects averaged 22.86 seconds. Radiation dosage to the subjects was extrapolated from measurements made on a masonite phantom (Table 1). The average exposure to a small field over the posterior nasopharynx was 1.84 r. Dosages at 30 fps were lower than those at 60 fps.

TRACING AND MEASURING. Prior to tracing, the projected cinefluorographic image was reduced to actual size by registration of the perforations in the centimeter rule. Subsequently the velum, posterior pharyngeal wall, and the back of the tongue were traced from each frame for the three selected subjects. After tracing, measurements were made for extent of velopharyngeal closure and for the minimum distance between the velum and the posterior pharyngeal wall. However, since it has been shown that the measurement of extent of closure is of doubtful reliability (*6*, *7*), this measurement was not included in the analyses performed in this study.

A reliability analysis (using phi coefficient and Pearson correlation techniques) was conducted to determine the extent of agreement between

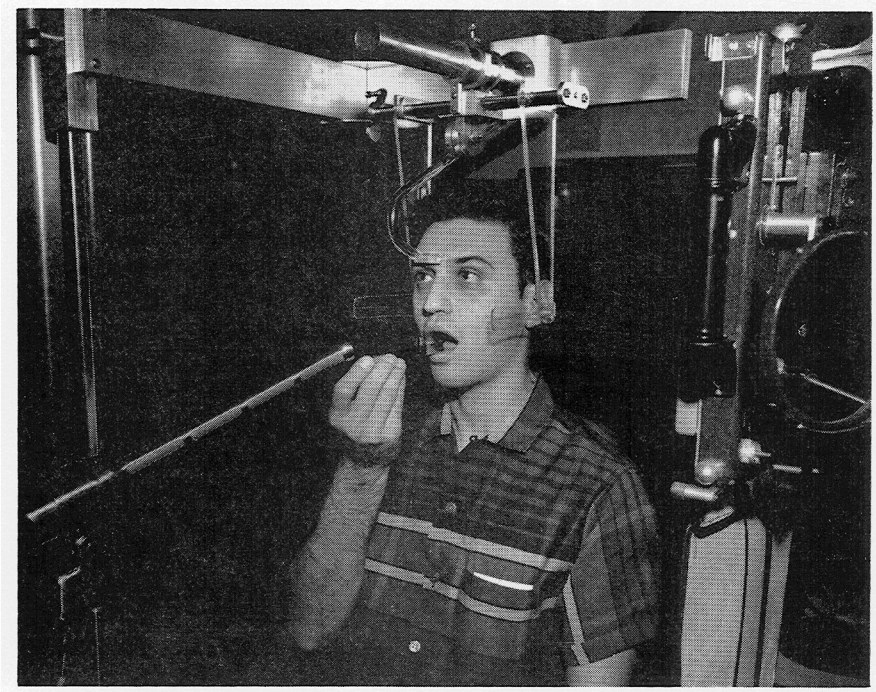

FIGURE 4: Positioning of steel ruler in the midline of the mouth for short filming sequence prior to the speech sample. The ruler is perforated at centimeter intervals to allow for correction of image magnification upon projection and tracing.

two tracers when tracing and measuring the same frames, and to determine agreement when the same tracer measured the same frames after an interval of time. Inter- and intra-tracer reliability proved to be high (Table 2).

Analysis and Results

DETERMINATION OF PRESENCE OR ABSENCE OF VELOPHARYNGEAL SEAL. When the velopharyngeal port opens or closes, it remains in that state for some time. Therefore, at the filming rates utilized, presence or ab-

TABLE 1. Calculated x-ray dosage to nasopharyngeal field for three subjects during filming of speech sequences at 30 and 60 fps. The measurements for 30 fps include the identification sentence.

Subject	*60 fps*		*30 fps*	
	duration in seconds	*calculated dose in r*	*duration in seconds*	*calculated dose in r*
1	11.0	.98	7.2	.18
2	13.0	1.66	12.0	.67
3	12.2	1.46	13.2	.74

TABLE 2. Summary of inter-tracer and intra-tracer agreement on tracings and measurements of velar function on the same cinefluorographic frames.

Tracer agreement	*No. of frames*	*Pearson correlation for agreement of measurement of velar opening*	*Phi coefficient for agreement of presence or absence of velar opening*
Inter	40	.95	.80
Intra	58	.94	.77

sence of velopharyngeal seal may be observed readily on many frames. Consequently, determination of the interval between frames analyzed depends upon the minimum number of consecutive frames which show the velopharyngeal port to be open or closed. If the number of omitted frames between the analyzed frames exceeds the number in this critical area, then an entire open or closed sequence may be missed. In the three subjects analyzed, the minimum number of consecutively open or closed frames at 60 fps was 10 frames and at 30 fps was eight frames. Thus, intervals of every eleventh frame at 60 fps and every ninth frame at 30 fps might omit an entire open or closed velopharyngeal sequence.

DETERMINATION OF EXTENT OF VELOPHARYNGEAL OPENING. To determine how analysis of consecutive frames compares with analysis of frames selected at varying intervals in the investigation of velar opening, auto-correlations as described by Senders (*11*) were run on the data derived from the films of the three subjects at both filming rates. The first order auto-correlation compared the measurement of velar opening on each frame with the measurement obtained from the frames immediately following it. The second, third, and fourth order were similarly done, with the second order comparing each frame to the second frame following it, the third order comparing each frame with the frame three frames after it, etc. The results of the auto-correlations are found in Table 3. As would be expected, the greater the number of frames between compared frames, the lower the correlation.

TABLE 3. Summary of auto-correlations for measurements of velar opening from frames exposed at 60 fps and 30 fps. Means were calculated from Fisher's r to z_r transformation.

Order of correlation	*60 fps Subject*				*30 fps Subject*			
	1	*2*	*3*	*Mean*	*1*	*2*	*3*	*Mean*
1	.91	.81	.93	.89	.82	.79	.83	.82
2	.77	.71	.85	.77	.57	.85	.69	.73
3	.57	.65	.76	.67	.31	.72	.58	.56
4	.40	.55	.62	.53	.08	.66	.39	.41

TABLE 4. Z scores for testing the significance of the difference between means of auto-correlations for frames exposed at 60 fps and 30 fps. Asterisked z scores are significant at the 5% level.

Orders of Auto-correlations compared	*z scores*	
	60 fps	*30 fps*
1–2	.88	.50
1–3	1.61	1.45
1–4	2.64*	2.27*
2–3	.73	.94
2–4	1.76	1.18
3–4	1.02	.83

The drop in correlation size as frames are omitted represents greater magnitude and rate of velar movement. Therefore, Subject 2, who showed a relatively slower velar movement and a small velopharyngeal opening, exhibited a much more gradual descent in correlations than did Subject 1 whose velum moved more quickly and through a greater excursion. Thus, individual variations in velar movement would influence the selection of an interval of analysis—the slower, less extensively moving, velum allowing larger intervals between frames without loss of information.

Table 4 shows the result of z tests which were used to evaluate the significance of the differences between the means of correlations. The only significant differences found at either filming rate were between the first and fourth order correlations. Shifting from a first to second or third order correlation did not produce a significant decrement in the degree of correlation.

Comparison of Filming Rates. Filming rates of 30 frames per second and 60 frames per second were compared in two ways: a) a comparison of the amount of movement measurable between adjacent frames; and b) a graphic analysis of the data obtained by measuring extent of opening.

It was predicted that if the slower filming rate were less adequate for quantitative analysis, the amount of movement between frames would be significantly greater at the slower rate because of the shift of the velum during the period when the shutter was closed. As shown in Table 5, the difference in the mean amount of movement between adjacent

TABLE 5. Mean amount of velar movement in millimeters, between adjacent cinefluorographic frames, exposed at 30 fps and 60 fps. Significant *t* values are asterisked.

Subject	*60 fps*	*30 fps*	*Difference*	*t*
1	.73	.63	.10	.53
2	.45	.58	.13	1.20
3	.61	1.0	.39	2.41*

frames for the two filming rates is not significant for Subjects 1 and 2. Subject 3 exhibited a mean range of movement of greater degree which indicates that adequacy of filming rates may vary with individual differences.

Comparison of the two filming rates in determining extent of velar opening is graphically presented in Figure 5. This figure compares the graphs of the mean of the measurements for the three subjects at percentage intervals. Per cent intervals were used to correct for the differences in speaking rates. The total open sequence (which included the last closed frame before velar opening and the first closed frame after the period of velar opening) is represented as 100%. These graphs present the pattern of velar movement occurring during the opening of velar seal for phonating the /m/ and /n/ sounds. The resulting curves show marked similarity, revealing few differences. The 30 fps film compares favorably with the 60 fps film and appears to be of sufficient speed to adequately study velar function.

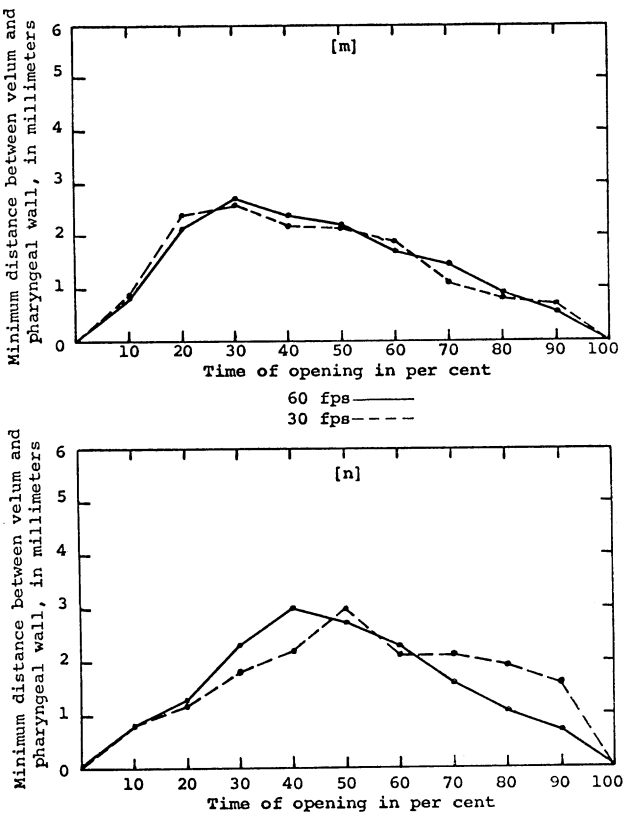

FIGURE 5. Graphs showing the mean measurement of velar opening for subjects 1, 2, and 3 on/m/ and /n/ at 60 and 30 fps.

Discussion and Conclusions

Data reduction and analysis in our subjects indicate that for cinefluorographic investigation of presence or absence of velar seal and extent of velar opening, analysis of all frames is unnecessary. Presence or absence of velopharyngeal seal can be determined for our three subjects by analysis of every eighth frame. It should be noted that the research design provided that every other non-syllabic element was a nasal. This insured maximum function of velopharyngeal activity for the segment of articulation studied. If this procedure for this sample revealed as much information as frame-by-frame analysis, it should be equally adequate for any sample of speech.

For most investigations, an interval of analysis as large as every third frame (at either 30 or 60 fps) is adequate for determination of extent of velar opening. The use of a smaller interval or frame-by-frame analysis did not yield a significantly greater amount of information in our subjects.

When compared with 60 frames per second, filming rates of 30 frames per second did not compromise the analysis of presence or absence of velar seal or extent of velar opening. The subject radiation exposure for a given speech sequence at 30 frames per second is less than at a filming rate of 60 frames per second. Additionally, with our equipment, the quality of films obtained at 30 frames per second, in terms of contrast and detail, was uniformly better than that obtained on the same subject at the faster filming rate.

Summary

A study was undertaken to determine the feasibility of facilitating data reduction in cinefluorographic investigation of velar function by interval frame selection and variation of filming rate.

Under controlled conditions, cinefluorographic filming sequences at 30 and 60 frames per second were obtained from 20 subjects during phonation of six disyllables. Subsequently the films of three subjects were chosen for frame-by-frame tracing and measuring. The presence or absence of velar seal and the extent of velar opening were analyzed.

Statistical analysis of the data obtained indicates that frame-by-frame analysis is unnecessary when investigating velar function. Films exposed at a rate of 60 frames per second do not appear to be significantly superior to those exposed at 30 frames per second in terms of data availability and reliability. Film quality is improved and radiation exposure is reduced when sequences at 30 frames per second are compared with those exposed at the faster filming rate.

Forrest Park School
1600 Silver Star Road
Orlando, Florida

Acknowledgements: This paper is based on an M. A. thesis completed by the senior author at the University of Florida under the direction of Professor G. Paul Moore.

References

1. Berry, H. M., Jr., and Hofmann, F. A., Cineradiographic observations of temporomandibular joint function. *J. prosth. Dent., 9*, 21–33, 1959.
2. Bjork, L., Velopharyngeal function in connected speech. *Acta Radiol.*, Suppl. *202*, 1961.
3. Carrell, J. J., A cinefluorographic technique for the study of velopharyngeal closure. *J. speech hearing Dis., 17*, 24–228, 1952.
4. Cooper, H. K., Cinefluorography with image intensification as an aid to the treatment planning for some cleft lip or cleft palate cases. *Amer. J. Orthod., 42*, 814–826, 1956.
5. Kirkpatrick, J. A., and Olmstead, R. W., Cinefluorographic study of pharyngeal function related to speech. *Radiology, 73*, 557–559, 1959.
6. Moll, K. L., Use of cinefluorography in speech research: a methodological study. Ph.D. dissertation, Univ. Iowa, 1960.
7. Moll, K. L., Velopharyngeal closure on vowels. *J. speech hearing Res., 5*, 30–37, 1962.
8. Nylen, B. O., Cleft palate and speech. *Acta Radiol.*, Suppl. *203*, 1961.
9. Nylen, B. O., A preliminary study by x-ray movie of cleft palate function pre- and post-operatively. In Tard Skoog (Ed.), *Translations of the International Society of Plastic Surgeons* (First Congress), 220–224, Baltimore: The Williams and Wilkins Co., 1955.
10. Powers, G. R., A cinefluorographic study of the articulatory movements of selected individuals with cleft palates. Ph.D. dissertation, Univ. Iowa, 1960.
11. Senders, V. L., Further analysis of response sequences in the setting of a psychophysical experiment. *Amer. J. Psych., 66*, 215–228, 1953.