1971

An Investigation of the Airflow Characteristics of Pulmonary Air Expulsion During Esophageal Speech

Alfred S. Lavorato
Portland State University

AN ABSTRACT OF THE THESIS OF Alfred S. Lavorato for the Master of Science in Speech, with emphasis in Speech Pathology and Audiology presented August 7, 1970.

Title: An Investigation of the Airflow Characteristics of Pulmonary Air Expulsion During Esophageal Speech.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

[redacted]

James F. Maurer, Chairman

[redacted]

Robert L. Casteel

[redacted]

Patrick O. Marsh

The general purpose of this investigation was to specify further the activity of the pulmonary tract in esophageal speech. Specifically, the study sought to determine whether pulmonary airflow (PAF) rate varied in continuous speech as a function of manner of production, voicing, syllabic position, and perceived level of stoma noise. PAF rate variation was defined as the frequency and magnitude of changes occurring in association with the variables of this study.

Six esophageal speakers utilizing the inhalation method of air intake were classified as high or low stoma (pulmonary) noise speakers on the basis of ratings by three speech pathologists. The /p, b, s, z/ phonemes were placed in arresting and releasing syllabic positions of single syllable words which were combined with other words to comprise two word phrases. The resulting eight phrases were uttered three times in random order by each speaker, while PAF rate was monitored at the tracheastoma, and recorded simultaneously with the phrases on the graphic printout.

The graphic printout of the PAF rate curves revealed that air flowed from the stoma continuously throughout the phrase for each phrase and each speaker, but showed no fluctuations in rate within phrases for any of the variables of the study. Additionally, it was noted that PAF rate was not associated with perceived level of stoma noise.

TO THE OFFICE OF GRADUATE STUDIES:

The members of the Committee approve the thesis of Alfred S. Lavorato presented August 7, 1970.

[redacted]

James F. Maurer, Chairman

[redacted]

Robert L. Casteel

[redacted]

Patrick O. Marsh

APPROVED:

[redacted]

Robert L. Casteel, Co-chairman, Department of Speech

[redacted]

Frank L. Roberts, Acting Dean of Graduate Studies

August 10, 1970

ACKNOWLEDGMENTS

I would like to express my appreciation to those who have been so very helpful to me in completing this study: Dr. James Shanks and Mr. Charles Dean for their interest and guiding comments; Edith Joyer for her assistance during the recording sessions; the American Cancer Society and members of the New Voice Club for their willing cooperation and assistance, and the Portland State University Research and Publications Committee for financial assistance. The time and equipment made available to me by Dr. Robert H. English, Director of Speech and Hearing at Portland State University, and the time made free for me by my friend and supervising instructor, Mr. Stephen McFarlane, were greatly appreciated.

Special gratitude is expressed to Dr. James Morris and Mr. William (Pete) Temple of the Veterans Hospital of Portland, Oregon for sharing their expertise and technical assistance in the area of pulmonary functions.

Sincere gratitude is expressed to my thesis committee--Dr. James F. Maurer, director; Dr. Robert L. Casteel; and Dr. Patrick O. Marsh for the kind and unremitting guidance and encouragement provided during this investigation.

Finally, with special affection, thanks go to mom and dad, and to my fiancée, Mary Joyer, who neither "appeared" nor "seemed" patient, understanding, encouraging and selfless--they WERE.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

One of the major tasks of the investigator in speech pathology has been to specify and describe normal and abnormal speech in terms of anatomical, physiological, acoustical and psychoacoustical parameters, and the interactions among them. It is recognized that when such specification and description are accomplished the speech process can be better predicted and clinically managed.

In the laryngectomized speaker, isolating the oral cavity from the pulmonary tract by means of laryngeal excision results in marked changes in each of these parameters and their relationships in the process of speech production. Generally, in alaryngeal or esophageal speech, insufflated air is trapped within the superior portion of the esophagus, which functions in esophageal speech as a vicarious lung. This air is propelled into the hypopharynx from the esophageal orifice which is closed off by the pseudoglottis (the area of the pseudoglottis is also called the pharyngo-esophageal segment and the cricopharyngeal sphincter). The folds of the cricopharyngeus function as vibrators in place of the thyroarytenoids in the larynx.

The dynamics of the laryngectomee's unique speech sound-producing mechanism have yet to be fully specified. Studies have been directed toward several areas investigating the mechanico-acoustical transformations of esophageal speech. Anatomically, for example, specifications have accrued pertaining to such parameters as the shape and size of the hypopharynx, composition, length, and cervical level of the pharyngo-esophageal or P-E segment (Diedrich and Youngstrom, 1966).

Physiologically, this modified speech producing system has been examined in terms of the vibratory, phonatory dynamics of the pseudoglottis (Stetson, 1937; Berg, Moolenaar-Bijl, and Damste, 1958) and the excursion and area of constriction of the P-E segment (Diedrich and Youngstrom, 1966). Size changes of the hypopharynx during air intake also have been investigated (Diedrich and Youngstrom, 1966).

Acoustical changes have been found to result from the altered anatomy and physiology of the laryngectomized speaker. Such changes are observed in articulation (DiCarlo, Amster, and Herer, 1955; Hyman, 1955; Diedrich and Youngstrom, 1966), pitch (Snidecor and Curry, 1959; Curry and Snidecor, 1961), loudness (Hyman, 1955), speaking rate (Snidecor and Curry, 1959; Berlin, 1965) and phrasing (DiCarlo et al., 1955).

Finally, psychoacoustical studies, those related to listener responses to esophageal speech, have focused on the intelligibility of alaryngeal speech. Several parameters of speech intelligibility have been used. DiCarlo et al. (1955) specified the number of correct phonemes, and phrasing related to speech intelligibility, while Snidecor and Curry (1959) and Berlin (1965) point out the importance of syllables per air intake. Shipp (1967), on the other hand, studied the "acceptability" of alaryngeal speech in relation to several phonatory variables.

As vital as specification and description of anatomy, physiology, acoustics and psychoacoustics are to the understanding of the laryngectomee's modified speech production, they are insufficient without elucidating the aerodynamics of the speech system. Speech, in part, represents an interplay between air particles and bodily tissues and structures. The laryngectomee, a person whose larynx has been surgically removed, now breathes through the trachea-stoma, an opening located inferiorly and medially in the neck. Consequently, the laryngectomee has two "air systems"--the oral-pharyngeal-esophageal tract and the pulmonary tract.

The following review of the literature pertains to the aerodynamics of esophageal speech in laryngectomees. It discusses methods of esophageal air intake for speech, phrasing, speed of air intake, articulation, air volume, airflow, and air pressure. Additionally, literature pertaining to the aerodynamics of the pulmonary tract is reviewed.

CHAPTER II

REVIEW OF THE LITERATURE

Several approaches have been taken to clarify the altered aerodynamic parameters of the oral-pharyngeal-esophageal tract in the laryngectomee. One approach relates to the process of esophageal air intake. Physiologically, two major types of air intake have been reported. One type is "inhalation" (Seeman, 1958; Diedrich and Youngstrom, 1966) in which the mouth is opened slightly in order to make air available to the esophagus. Thoracic inhalatory action in this type of air intake simulates that of laryngeal speakers. The other type is "injection" (Berg et al., 1958; Diedrich and Youngstrom, 1966). This latter type has been categorized further into two techniques. One technique is the "glossal-press" in which either lingua-alveolar or lingua-palatal contact is made coincidentally with a posterior movement of the tongue dorsum. The other technique is the "glossal-pharyngeal press" in which the same dorsal tongue movement occurs, but with the tongue tip touching the alveolar ridge and the blade articulating with the hard palate and velum. Moolenaar-Bijl (1953) describes the injection method as "insufflating air into the oesophagus [sic] by means of a small, nearly imperceptible movement with the lips on the tongue."

Esophageal air intake has been clarified further, by others who have asserted that the time and area of air intrapment is governed by the sphincteric control of the cricopharyngeus (Hoople and Brewer, 1954). Diedrich and Youngstrom (1966) investigated the factor of hypopharynx size during air intake, and found that speech skill was not significantly correlated with the width of the hypopharynx during air intake.

Diedrich and Youngstrom (1966) studied the temporal dimension of esophageal air intake in relation to phrase units occurring in speech. They observed that air intake could occur during (1) a period of rest, which is the silent time interval preceding an utterance, (2) an interphase pause, which is the silent time interval between words, phrases or sentences, (3) an intraphrase interval, which is the unit of utterance itself, and (4) an intraphrase pause, which is a silence of short duration "contained within the phrase interval due to intervocalic stop consonants and linguistic prosody."

Some studies have dealt with the temporal aspect of air intake as it relates to speed. Snidecor (1962) and Berlin (1963), observing rate of esophageal air inflow, reported for good speakers intake rates ranging from instantaneous to .75 seconds. Snidecor and Curry (1959) and Diedrich and Youngstrom (1966) reported comparable results indicating that air intake speed averages ranged from 1/3 of a second to nearly one full second. Diedrich and Youngstrom (1966) found this to be true almost regardless of the phonemes uttered. Furthermore, correlation of air intake speed with speech skill revealed that better speakers take in air faster than poor speakers, a finding harmonious with that of Snidecor (1962) and Snidecor and Isshiki (1965a). Diedrich and Youngstrom (1966) determined that linguistic context influences speed of air intake. In terms of quickest to slowest, they report the order as being stop consonant intraphrase pauses within an intraphrase interval, inter-phase interval, and rest (0.5 seconds or longer).

Other investigations (Moolenaar-Bijl, 1953; Berg et al., 1958; and Diedrich and Youngstrom, 1966) were concerned with air intake and articulation. The three studies cited above were in agreement with each other in finding that fewer intakes are required for plosive sentences. Additionally, Diedrich and Youngstrom (1966) reported that the phonemes articulated did not influence esophageal air intake when structures were in a rest posture.

Oral-pharyngeal-esophageal tract air volume has been another matter for investigation in the area of aerodynamics. Berg et al. (1958) reported that the capacity of air in the esophagus is between 40 cc and 80 cc. Snidecor and Isshiki (1965b) found a 615 cc capacity in one speaker, but they discovered that the stomach also had been used as an air reservoir. Diedrich and Youngstrom (1966) reported the work of Vrticka and Svoboda which indicated that esophageal air volume increased as rehabilitation progressed. Robe, Moore, Andrews and Holinger (1956) found a sub-pseudoglottic air column in good speakers but not in poor speakers. Snidecor and Isshiki (1965a) studied the air volume parameters of the laryngectomee during the process of esophageal speech. They observed that 5 cc to 16 cc of air was used per syllable. The laryngeal speaker uses about 43 cc per syllable. During the production of words, the better esophageal speakers in their study used 16 cc to 22 cc of air per word; the poor speakers used 7 cc of air per word. In addition, for a 51-word passage, 372 cc to 1, 115 cc of air was used. The laryngeal speaker used 3, 020 cc. The total amount of air insufflated for speech, they reported, was quite variable. Finally, Diedrich and Youngstrom (1966) suggest that the greater the esophageal air capacity, the more likely it will be that there will be better speech.

Another parameter of the oral-pharyngeal-esophageal tract studied has been that of esophageal airflow during speech. Snidecor and Isshiki (1965a) stated that good speakers had faster flow rates than poorer speakers, and that for any given speaker, esophageal air intake flow rate (33 cc/second to 135 cc/second) exceeded exsufflation airflow rate (25 cc/second to 97 cc/second).

A few researchers report results of studies dealing with the parameter of air pressure within the oral cavity. DiCarlo et al. (1955) examined oral air pressure values for /p/ and /b/. Unlike the pressure values for /p/ and /b/ in laryngeal speakers, air pressure for /p/ and /b/ in esophageal speakers was of equal magnitude. Diedrich and Youngstrom (1966) used an oral manometer in an effort to determine whether oral air pressure was related to speech skill. No significant correlations were found.

Very little has been done to specify similar parameters of the pulmonary tract which, due to laryngeal excision and a re-directed breath column, no longer are influenced by laryngeal and supraglottic structures. Yet, pulmonary tract activity seems to coexist with the speech signal in esophageal speech. The major issue pertaining to the activity of this tract has been that of synchrony or dyssynchrony between phonic respiration and pulmonary respiration. Synchronous activity between the two tracts is characterized by concurrent insufflation and pulmonary inhalation, and by concurrent exsufflation and pulmonary exhalation and voicing. This synchronized cycle of inhalation and exhalation is most common for speakers utilizing the inhalation method of air intake. Dyssynchrony exists when phonic and pulmonary respiration activities occur out of phase with each other. Concomitant respiratory activities are supported either by assertions of physiological and psychological economy (Moolenaar-Bijl, 1951) or of increased vocal intensity and number of syllables uttered per exsufflation (Snidecor and Isshiki, 1965b). Proponents of dyssynchrony advocate out of phase respiration and voicing because of the aid to exsufflation provided by a fixed thorax in which abdominal wall and chest muscles serve to exert pressure (Gardner, 1951; Snidecor, 1962).

While earlier studies favored dyssynchrony (Stern in Kallen, 1934; Kallen, 1934; Morrison and Fineman, 1936; Howie, 1947; and Mason, 1950), current investigators tend to advocate synchrony. Robe et al. (1956) addressed themselves to the issue pioneering the use of respiratory equipment which allowed direct monitoring of pulmonary air at the tracheastoma rather than using pneumographs attached to thoracic/abdominal areas. Data were recorded in terms of aerodynamic parameters (air volume) by utilization of oral and pulmonary spirometers. They reported a preponderance of synchronized pulmonary and phonic respiration in all of their subjects. Investigations by Snidecor and Curry (1959) and Snidecor and Isshiki (1965a) also revealed data demonstrating the prevalence of synchronous pulmonary respiration and phonation. Snidecor and Isshiki (1965b) reported in a superior esophageal speaker that 76% of esophageal air intake occurred during pulmonary inhalation, and that 97% of esophageal voicing or phonation occurred on pulmonary exhalation.

In addition to the issue of respiratory-phonatory coordination, pulmonary air volume observations have been made. Robe et al. (1956) found no correlation between pulmonary and oral air volumes, between speech fluency and pulmonary air volumes or oral air volume.

/s/ phoneme. In general, stoma noise along with other noises associated with esophageal speech, complicated the speech signal resulting in "very undesirable effects."

Diedrich and Youngstrom (1966) provide further insight pertaining to pulmonary blowing activity. Differential pulmonary air expulsion in their study was based upon the occurrence of blowing and the stage of an utterance at which onset of blowing occurred. Expulsion of pulmonary air in esophageal speakers was studied using a barium bib and cinefluorography. The speakers, both inhalers and injectors, were instructed to utter /i/, /ɑ/, /u/, /pɑ/, /tɑ/, /kɑ/, /sɑ/, /mɑ/, /nɑ/, and /rɑ/, individually. Outward movement of the bib signified the occurrence of expulsion of pulmonary air from the tracheastoma. Frequency counts of the onset of blowing for each vowel or syllable were made at five different times or stages of the utterance. These stages were: (1) Stage 1--at maximum air intake, (2) Stage 2--during air intake after air was in the esophagus, (3) Stage 3--during pre-phonation, (4) Stage 4--simultaneous with initiation of phonation, and (5) Stage 5--after phonation started. Their data revealed that differential pulmonary blowing did occur between vowels and syllables and that such blowing occurred more frequently for plosives than for fricatives and other phonemes. Furthermore, the greatest difference in blowing frequency between the vowels and syllables occurred at Stage 3, and secondly, at Stage 2. Additionally, the onset of blowing for each vowel and syllable occurred most often during Stage 3, and secondly, during Stage 2. Finally, it was found that blowing occurred for 13 out of 27 speakers, and that of these speakers, 3 of the 7 inhalers exhibited blowing. Pulmonary tract participation in esophageal speech, then, was marked by differential blowing as functions of speaker, phoneme, and stage of utterance.

Further evaluation and specification of pulmonary activity seems highly desirable in view of the participation in and potential depreciation of esophageal speech by such activity. The literature, for example, is unclear as to the degree to which pulmonary air expulsion exists in esophageal speech, and in regard to the factors with which pulmonary air expulsion varies. Possibly, while pulmonary blowing coalesces predominantly with phonic respiration, it also may vary with other factors. The results of the study by Diedrich and Youngstrom (1966) imply fluctuations in pulmonary blowing with the utterance of isolated phonemes and nonsense syllables. It is not known, however, whether phoneme-differentiated blowing is maintained during ongoing, propositional speech. That is, it is unknown whether changing respiratory events occur in Stage 5 specifically for phonemes uttered in a complex phonetic setting. Additionally, the literature is unclear about the relationship between "blowing" and "noise." Consequently, there is a tendency to assume that blowing routinely results in noise. This assumption may be invalid. Esophageal inhalers, for instance, are considered to exhibit pulmonary blowing noises commonly during speech on the exhalation cycle, yet the data of Diedrich and Youngstrom (1966) indicate that injectors also exhale air from the stoma during speech. Pulmonary noise, however, reportedly is not a characteristic of injectors, despite the blowing. Reports in the literature also fail to clarify the components of "blowing," and to relate these components to noise from the stoma. There has been no systematic investigation of pulmonary aerodynamics (1) utilizing the airflow component of pulmonary blowing as a parameter of measurement, and (2) assessing pulmonary air expulsion or blowing as it varies in continuous speech with several phonetic elements.

I. STATEMENT OF THE PROBLEM

The present study was undertaken to specify more fully the participatory and depreciatory aspects of pulmonary blowing in esophageal speech. The general purpose was to determine whether the expulsion of pulmonary air for esophageal inhalers of different perceived levels of stoma noise fluctuated predictably with specific phonetic factors. The specific questions asked were:

1. Does a discrete change in pulmonary airflow (PAF) rate occur with the phonetic factors of manner of production (fricative and plosive), voicing, and syllabic position, when

assessed in the complex phonetic setting of propositional phrases?

2. If so, does the magnitude of the change in PAF rate vary predictably with such phonetic factors?

3. If any of the above relationships exist, are they maintained at different levels of perceived pulmonary (stoma) noise?

CHAPTER III

METHOD

I. APPARATUS

The acoustic recordings were obtained utilizing the Ampex AG 500, Ampex 601, and Wollensak T-1500 tape recorders and speaker systems, and a Shure Unidyne III 545 dynamic microphone. Instrumentation facilitating the recording of pulmonary events consisted of a pneumotachograph, a Statham transducer, model PM 283TC, and a Brush Recorder Mark 240, model 11-6402-02. A block diagram of the components of the system is shown in Figure 1.

The devised equipment allowed simultaneous recording and measuring of both pulmonary air events and speech signals. It operated as follows: A rubber mouthpiece was adapted into a stoma appliance for monitoring the air at the site of the tracheastoma. An air-tight seal at this site was achieved by pasting a 1/8 inch foam rubber sheeting to the skin surrounding the stoma, and then, adhering the stoma appliance to the foam rubber. A 24-inch plastic tube, attached at one end to the stoma appliance, connected at its opposite end to the pneumotachograph (flowmeter), which was mounted on a stand. The flowmeter recorded airflow rate by detecting pressure

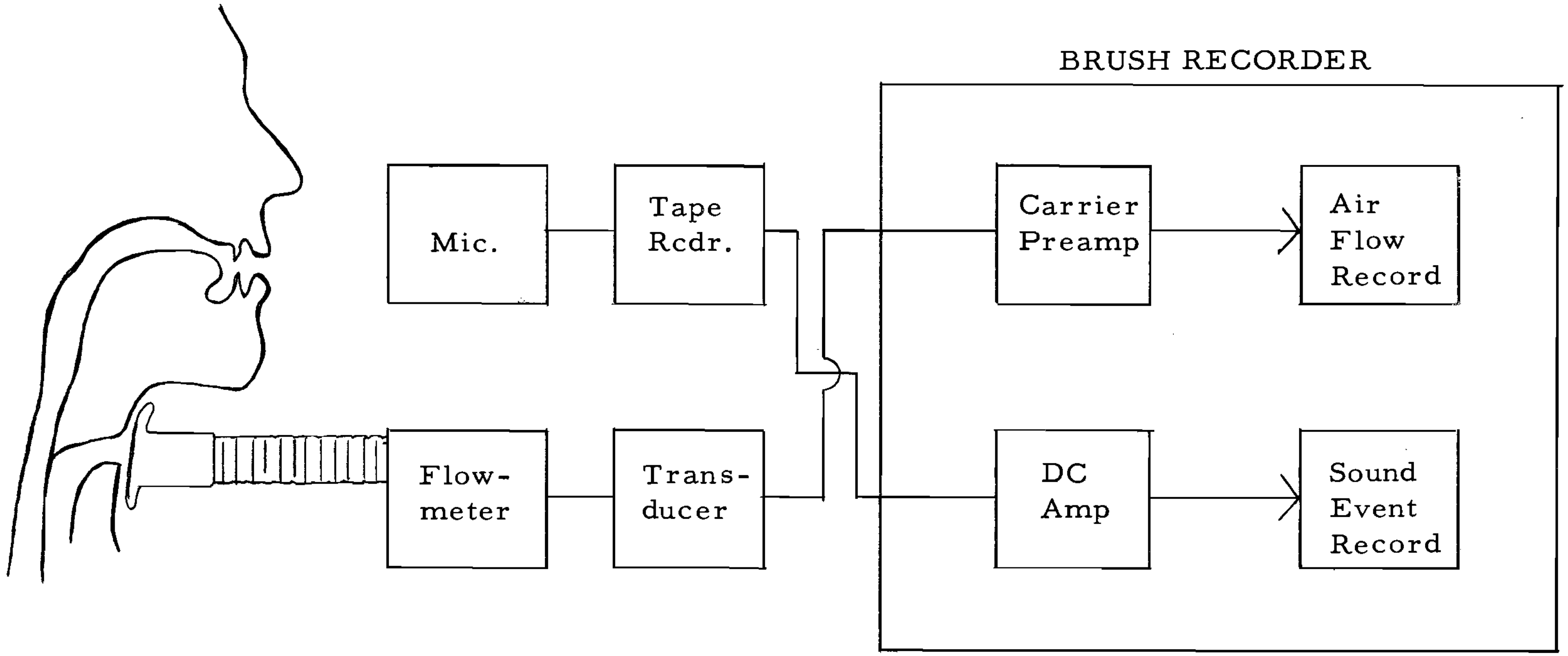

Figure 1. Diagram of apparatus used to record airflow rate and acoustic signals during speech.

differentials across its two fine wire screens. The differentials were converted into electrical signals by a transducer and displayed on the Brush Recorder graph paper. The design of the flowmeter permitted normal respiratory functioning while the apparatus remained attached. The stoma appliance, therefore, required only one fitting for the entire recording session. The microphone was mounted on a stand and positioned three inches in front of the speaker's mouth. A patch cord from the Ampex AG 500 output outlet directed the speech signal to the Brush Recorder, which has a maximum channel frequency response of 100 Hz. The graph paper of the Brush Recorder operated at a speed of 200 millimeters per second (mm/second) during the recording of each utterance.

The instrumentation designed for this study was unique to the study of pulmonary aerodynamics in esophageal speakers, providing sensitive data regarding the occurrence of onset of PAF and variations of PAF during utterance of chains of connected phonemes.

II. SUBJECTS

Six laryngectomee speakers were selected from the New Voice Club of Portland, Oregon, an affiliate of the International Association of Laryngectomees. Each subject was informally judged as predominantly using the inhalation method of air intake. Their first task was to tape record three sentences. On the basis of judgments of three speech pathologists, the speakers were divided into two classes: High Pulmonary Noise Speakers (HPNS) and Low Pulmonary Noise Speakers (LPNS).

Tape Recording the Stimulus Sentences

The microphone, mounted on a stand with a gooseneck attachment, was positioned three inches from the mouth of the speaker. In order to record both the speech signal and any stoma noise, the horizontal level of the microphone was positioned so that the superior rim of the front of the microphone was parallel to the horizontal midpoint of the subject's chin. Intensity of the speech samples to be presented to the judges was controlled prior to recording the sentences by having each subject utter "baseball" and adjusting the Record Level dial on the Ampex AG 500 until the /ɔ/ phoneme in the word peaked -3 on the VU meter of the Ampex. Subjects were instructed to use their typical conversational voice when uttering both "baseball" and the three stimulus speech sentences (see Appendix A) to be taped. Additionally, they were told to read the sentences during the recording consecutively. Once the intensity was established, the subjects recorded the stimulus sentences.

The recorded stimulus sentences were randomly ordered using a table of random numbers, and transferred in the new order to another tape, using the Ampex 601, the Wollensak, and a patchcord.

An additional measure taken to ensure equal intensity of each sentence during the transfer was to monitor the intensity of each sentence on the Ampex 601 VU meter. These precautions were taken to reduce the influence of intensity on the judges' ratings.

The judges were told that the tape would be presented twice and to make no rating on pulmonary stoma noise until the second playing of the tape. At the end of the tape was a comment informing them that the tape was to be replayed and instructing them to listen again to each sentence and rate them on high or low stoma noise according to the instructions on their rating sheet (see Appendix B). The format repeated throughout the tape was as follows: (1) stimulus sentence, (2) a four second time interval, (3) the sample number, and (4) a four second time interval.

Judging Speakers for Pulmonary Noise

Pulmonary noise (stoma noise) judgments were obtained by presenting the stimuli (N = 24) through the Ampex AG 500 speaker system to the three judges sitting at an arbitrarily chosen distance of ten feet from the speaker. The Reproduction Level of the AG 500 was set at "6" while on the speaker system volume gain was set at "7" and Equalization at "0." The first presentation of the randomized sentences was intended to familiarize the judges with the samples and with the range of stoma noise severity to be rated.

During the second presentation of the tape, judges made their ratings of the stoma noise using a 5-point scale in which "1" represented low stoma noise and "5" represented high stoma noise. Instructions to the judges were on a rating sheet. Intrajudge reliability was determined by repeating within the randomized stimulus sentence tape, the first sentence recorded by each esophageal speaker. Each judge, therefore, made two judgments on six sentences (N = 12). The Pearson Product-Moment Correlation Coefficient was used to determine how well the first judgment correlated with the second judgment. The same formula was utilized for assessing agreement between judges. Product-Moment Correlations of intrajudge reliability for the three judges were .98, .97, and .95, indicating high consistency between judgments for each judge. Interjudge reliability coefficients were .91, .78, and .82, indicating high consistency of judgments of stoma noise between judges.

Classification of Speakers

The difference between the three highest and the three lowest raw score ratings of stoma noise were analyzed by a Critical Ratio Z-Test (Thompson, 1965). Statistical analysis of the difference between these two groups revealed differences that were significant beyond the .01 level of confidence (df = 35). The three speakers receiving the three highest stoma noise scores were classed as "High

Pulmonary Noise Speakers" and the three speakers receiving the three lowest stoma noise scores were classed as "Low Pulmonary Noise Speakers."

III. SPEECH MATERIALS

The phonemes /p, b, s, z/ were selected to represent the phonetic variables of plosive and fricative manner of production and voicing. These were placed in initial (releasing) and final (arresting) positions of single syllable words. Several factors prompted the selection of these sounds. DiCarlo et al. (1955) indicated that stop plosives and final consonants are more intelligible than fricatives and initial consonants. Furthermore, voiceless consonants were reported as slightly more intelligible than voice consonants. If differential pulmonary blowing were found to exist, it might be so, in part, with sounds of such opposing distinctions. Ostensibly, unintelligible phonemes are produced with greater difficulty than intelligible ones. Possibly, difficult phonemes elicit pulmonary activity more than easily produced phonemes.

The single syllable words containing the target phoneme were either the first or second word of a two word phrase (see Appendix C). The phonetic environment of the target phonemes was held constant in each phrase. When contained in the first word of the phrase, the target phoneme was immediately preceded by /ɑ/ and followed by /l/ as in the phrase "top lamb." When contained in the second word of a phrase, the target phoneme was immediately preceded by /l/ and followed by /ɑ/ as in "ball pot." The /l/ was used because it is one of the easier phonemes produced by esophageal speakers (DiCarlo et al., 1955). As a control element, /l/ either appeared in the word "ball" or in the word "lamb." The /ɑ/ phoneme always was part of the target phoneme word. When the target phoneme was in the final position of the word, the /ɑ/ sound preceded it; when it was in the initial position of the word, the /ɑ/ sound followed. Stimulus events were limited to two word phrases in order to maintain a certain degree of homogeneic control without relinquishing propositionality. One word, "zombie," was reduced to "zomb" in order to maintain equal length for each utterance. In this instance, some propositionality may have been lost by such a reduction.

The 24 utterances per speaker (8 phrases uttered 3 times each) were randomized, using a table of random numbers. This was done to prevent bias on phonemes due to speaker fatigue.

IV. RECORDING SESSION PROCEDURES

Each subject was provided a list of the phrases to be read aloud. The subjects were instructed to read the material in a normal conversational manner, and to proceed to each successive phrase only when receiving an audible, pre-utterance signal from the experimenter. The design of the equipment was explained to the subject to allay possible fears resulting from wearing the stoma appliance. It was demonstrated that normal, vegetative breathing could continue while the appliance was being worn. The speaker was instructed to say the phrase as one word ("ballpot"), an act that approximated speech production as it might occur in spontaneous speech. The correct manner of utterance was demonstrated by the experimenter and the subject was allowed a short time to study and practice the phrases.

The magnetic tape ran continuously throughout the session. A "Sensitivity" dial on the Brush Recorder was set either at "2" or "5," enabling the graph paper to accommodate the amplitudes of airflow rate unique to each speaker. The Brush Recorder paper speed ran at 2 mm/second until the phrase was recorded at which time it ran at 200 mm/second. Synchronization of the recordings of the acoustic events and the aerodynamic events for later analysis was accomplished utilizing the experimenter's audible, pre-utterance signal prior to each speech utterance. The pre-utterance signal was recorded simultaneously on both the magnetic tape and the graphic printout. The sequence of recording events for each speech utterance was: (1) the pressing of the Event Marker button, (2) passage of a one second interval, (3) an audible pre-utterance signal from the experimenter emitted simultaneously with the switching of the Brush Recorder paper speed to 200 mm/second, and with the pressing of the Event Marker, (4) the utterance by the speaker of the speech material, and (5) the reduction of paper speed to 2 mm/second.

V. GRAPHIC ANALYSIS

Graphic analysis of pulmonary airflow (PAF) rate associated with each target phoneme in a given utterance required definition of the onset and terminal boundaries of each target phoneme, as represented on the acoustic and airflow channels of the Brush Recorder paper, and the assignment of a value of the magnitude of the PAF peak for each phoneme.

Location of the phoneme on the PAF rate curve of the graphic printout was achieved by transferring information from the magnetic tape as follows: The onset of the pre-utterance signal, and the initial and terminal boundaries of the target phoneme on the magnetic tape were determined manually. The onset of the pre-utterance signal, the point at which the phoneme was first audible, and the point at which the phoneme ceased being audible were marked on the magnetic tape with a pen, thus defining the interval from the onset of the recording of a given phrase to the interval representing the target phoneme. These intervals were measured in millimeters and converted into milliseconds. The event mark at the margin of the graph paper aided the location of the various boundaries. The target phoneme boundaries were established on the audio channel of the Brush Recorder graphic printout by extrapolating the intervals, in milliseconds, from the data derived from the magnetic tape. Those sections on the PAF curves in the airflow channel of the printout which were exactly parallel to target phoneme intervals on the audio channel were considered to represent the PAF rates and the PAF peak rates associated with the plosive and fricative phonemes used for this study. The duration of the PAF peaks associated with the target phonemes, including Rise-time, Peak, and Decay-time, probably would approximate an average of 100 milliseconds (20 mm on the graphic printout shown in Figure 2). One hundred milliseconds represents the average length of duration of a phoneme in normal speech (Fairbanks, Everitt, and Jaeger, 1954). A scattering of available research, reported by Dean (1968), indicates that 100 milliseconds may be a reasonable estimate for esophageal speech.

A "Sensitivity" setting of "2" on the Brush Recorder was calibrated through use of a rotometer so that each minor division on the printout represented an airflow rate of 100 milliliters per second (ml/second). A setting of "5" was calibrated so that each minor division represented 250 ml/second. The assignment of the value or magnitude of PAF rate peaks for each target phoneme was made by multiplying K-factor 100 or 250 by the number of divisions from "zero" airflow to the peak on the PAF rate curve. This value represented the peak pulmonary airflow (PPAF) rate in milliliters per second for each utterance by each speaker.

The data under consideration in this study consisted of the frequency of peaks and the PPAF rate magnitudes associated with manner of production, voicing, syllabic position, and perceived level of pulmonary noise. The measures were:

1. Frequency and magnitude of PPAF for plosives and fricatives.
2. Frequency and magnitude of PPAF for voiced and unvoiced phonemes.
3. Frequency and magnitude of PPAF for initial and final plosives, and for initial and final fricatives.
4. Frequency and magnitude of PPAF for HPNS for plosives, fricatives, voiced and unvoiced phonemes.
5. Frequency and magnitude of PPAF for LPNS for plosives, fricatives, voiced and unvoiced phonemes.

CHAPTER IV

RESULTS AND DISCUSSION

I. RESULTS

The exhalation or blowing portion of each PAF curve for each phrase was the focus of interest in this investigation, and may be described by three general phases (see Figure 2): (1) Rise-time--either a gradual or rapid increase in the volume of air flowing per second from the stoma, (2) Major Peak--the greatest exhalatory air-flow rate attained, and often, sustained, and (3) Decay-time--the gradual reduction in airflow rate. The PAF curves are not cumulative; rather, they reflect the volume of air flowing from the stoma at each instance in time throughout the duration of an utterance.

Small flutters of airflow appear occasionally on some of the PAF curves, both for HPNS and LPNS. These may be cardiogenic, the result of speaker tension, or the result of some vibratory disturbance of the equipment. Their average duration is about 5 milliseconds, and their peaks represent 1/10 to 1/20 of the value of the Major Peak for a given phrase of a speaker; consequently, their impact upon the airflow data was considered minimal.

It becomes apparent from visual inspection of the graphic

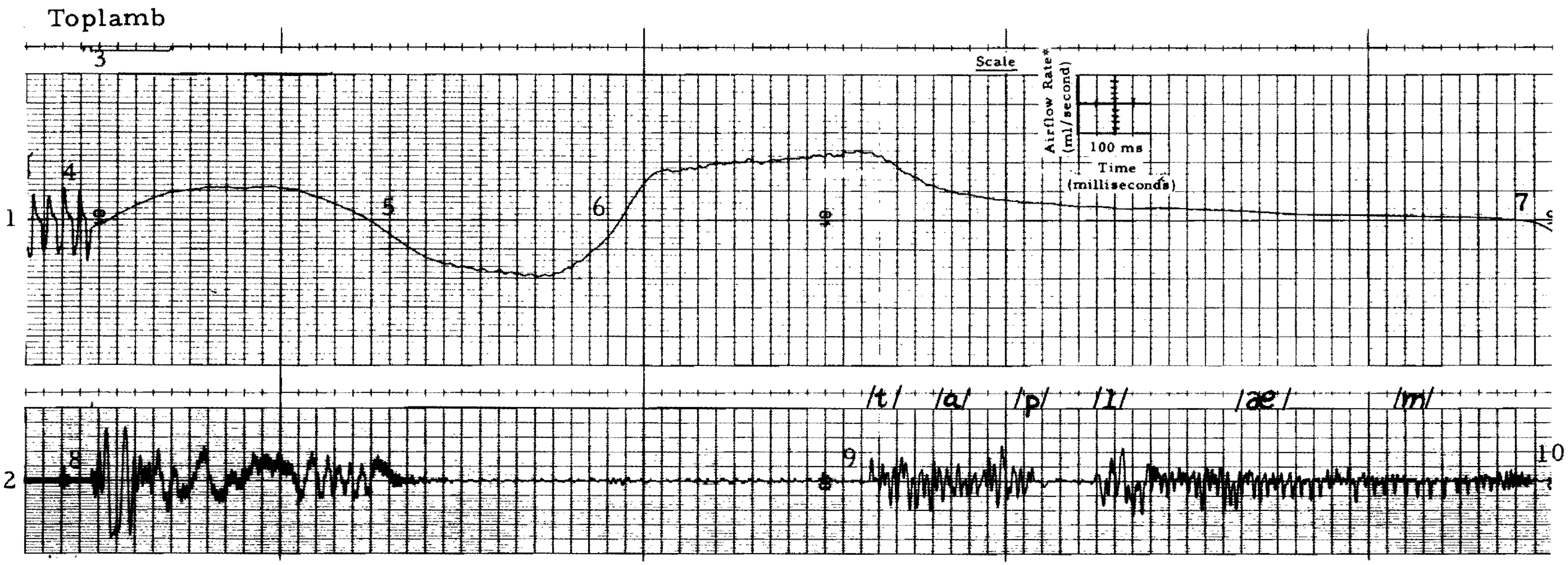

Figure 2. A sample of the synchronized airflow and audio graphic printout data for one phrase of one speaker. (1) PAF channel--airflow rate curve of pulmonary inhalation and exhalation, (2) Audio channel--two word phrases, (3) Event marks, (4) Vegetative respiration--recorded at paper speed of 2 mm/second, (5) Onset of inhalation just prior to phonation, recorded at 200 mm/second, (6) Onset of exhalation coexisting with the utterance of the phrase, (7) Termination of exhalation accompanying phrase, (8) Pre-utterance signal, (9) Onset of the phrase, (10) Termination of the phrase. *The Scale facilitates interpretation of magnitude (milliliters/second) and duration (milliseconds) values. At a sensitivity setting of "2" each interval on the graph represents 100 ml/second; at a sensitivity setting of "5" each interval represents 250 ml/second.

given target phoneme. The onset of blowing occurred mostly in Stages 1 and 2, combined (Table I). No temporal data are available delineating the duration of Stages 1 and 2, thereby preventing further breakdown of these two stages. Table I, however, illustrates that the frequency of onset of blowing in Stages 1 and 2 is 98, which is more than double the frequency of blowing onset in Stage 3, which totals 44. Perusal of Diedrich and Youngstrom's data reveals that the totals of Stages 1 and 2 do not lead to similar results. This

TABLE I

FREQUENCY OF ONSET OF TRACHEASTOMA AIR EXPULSION CLASSIFIED ACCORDING TO PHRASE UTTERED AND STAGE OF UTTERANCE AT WHICH ONSET OCCURRED

	Ballpot	Toplamb	Ballboss	Roblamb*	Ballsob	Bosslamb*	Ballzomb	Gauzelamb	
Stages 1 & 2	13	12	11	10	12	11	14	15	Σ = 98
Stage 3	5	6	7	7	6	6	4	3	Σ = 44
Stage 4	--	--	--	--	--	--	--	--	
Stage 5	--	--	--	--	--	--	--	--	

*One speaker skipped item.

suggests that the differences between data are real, very likely due to the sensitivity of the equipment used in the present study. Design features of these two studies are somewhat incompatible, suggesting caution in generalizing about the differences found. The present study, however, does characterize pulmonary tract participation in esophageal speech differently than Diedrich and Youngstrom in terms of the factors of time, speaker, and phoneme. Their data, based on isolated utterances and gross instrumentation, are not borne out in connected speech assessed by utilizing sensitive equipment.

In terms of speech intelligibility, the depreciatory effects to air-flow rate related to pulmonary participation in a continuous phonetic context appear to be minimal for both classes of speakers with respect to voiced and unvoiced plosives and fricatives in arresting and releasing syllabic positions. Had PAF peaks of sufficient magnitude been phoneme specific for /p, b, s, z/, it might have been suggested that these phonemes, being of relatively low phonetic power, potentially could be masked by large amounts of pulmonary air expulsion. Furthermore, greater air expulsion would have been associated with greater perceived stoma noise had the PAF peaks for HPNS been greater than for LPNS. This noise, in turn, might have been found to have been phoneme specific. The available evidence, however, fails to substantiate these speculations.

It was possible, however, to examine further the data gathered in order to explore airflow as it relates to the possible depreciatory effects of noise on esophageal speech in general. Diedrich and Youngstrom (1966) state:

> In utilizing the inhalation technique, some laryngectomees develop through the tracheal stoma excessive blowing noises which tend to mask the esophageal speech. This usually is caused by the very rapid contraction of the abdominal muscles which results in a rapid expulsion of the pulmonary air out of the trachea.

The possibility exists, therefore, that greater pulmonary airflow rate in general may be associated with greater perceived level of stoma noise. This relationship was analyzed by determining the PAF values of the Major Peaks and the Rise-times for each speaker in each class. Major Peak means and ranges are plotted in Figure 3. The computed Rise-time means and ranges for each speaker in both classes are plotted in Figure 4. This latter measure indicates the pre-Major Peak blowing duration, and to some extent, implies the force with which the column of pulmonary air is emitted from the stoma. Short Rise-times would indicate greater muscular contraction, and therefore, greater volume of molecular air energy within a shorter period of time. These dynamics are associated with greater acoustic intensity.

It is obvious from Figures 3 and 4 that neither Major Peak magnitudes nor Rise-times differentiate the two classes of speakers. Indeed, in Figure 3, the average volume of airflow per second for two

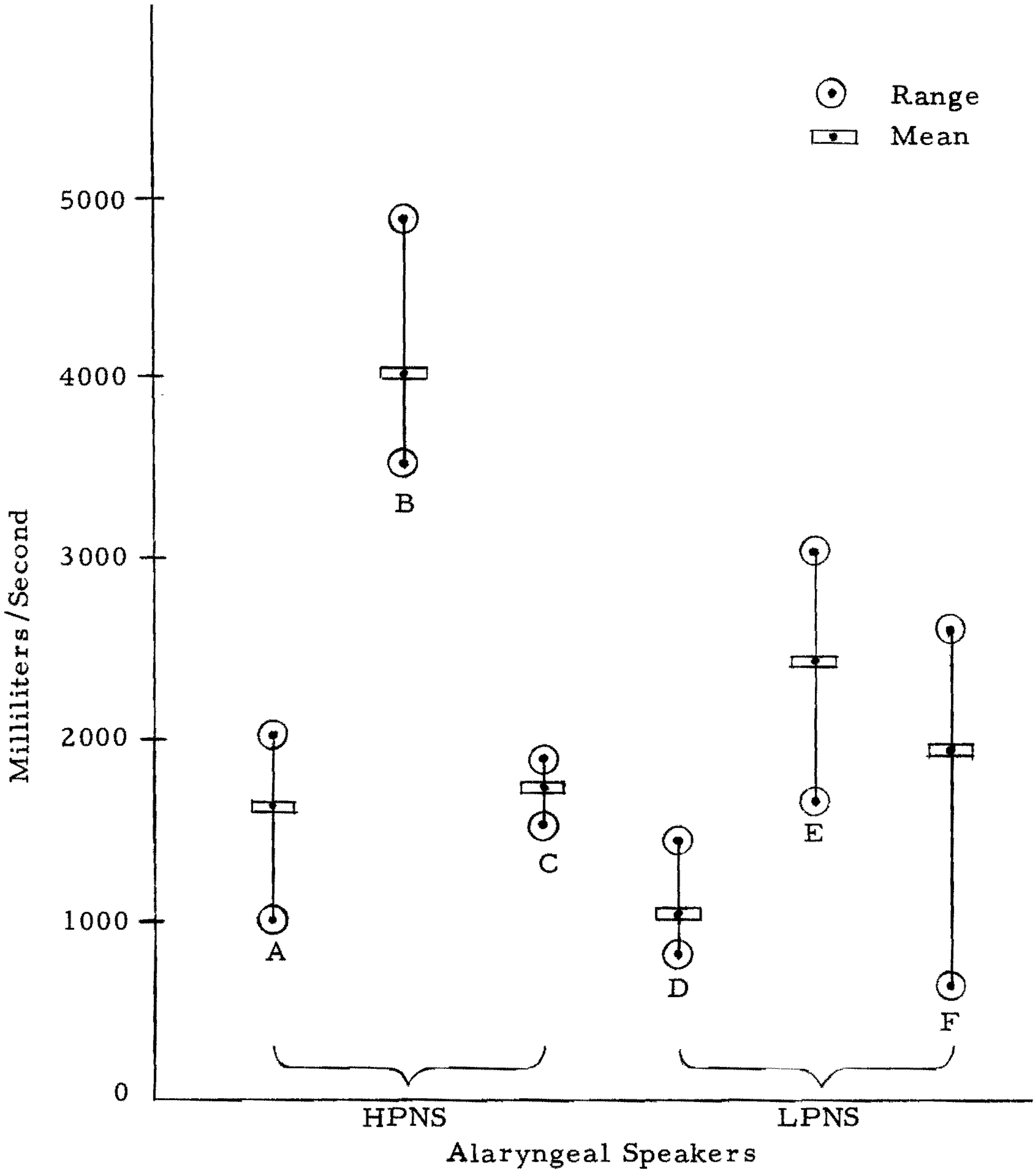

Figure 3. Ranges and means of Major Peak pulmonary airflow (PAF) magnitudes during utterance of phrases for high pulmonary noise speakers (HPNS) and low pulmonary noise speakers (LPNS).

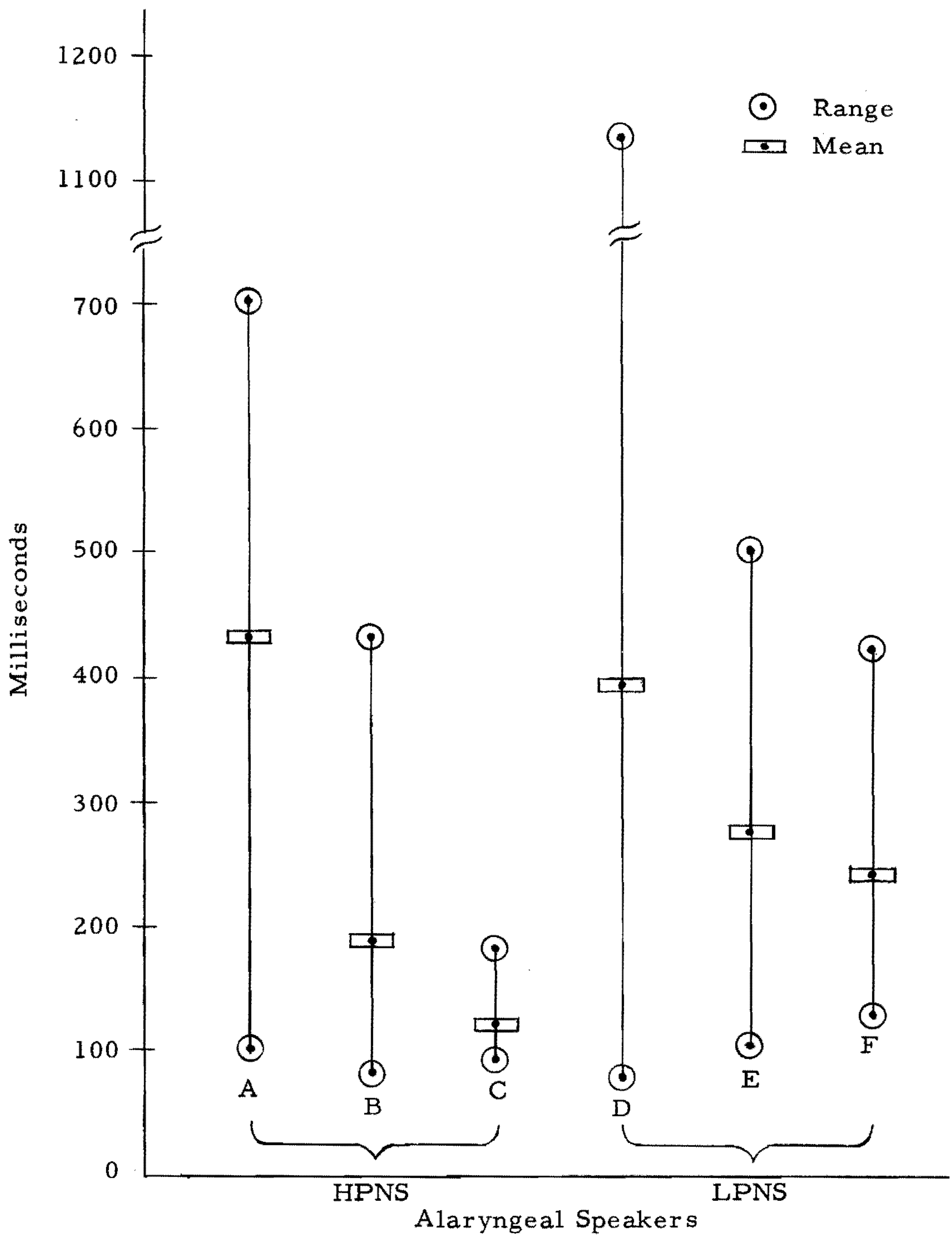

Figure 4. Ranges and means of Major Peak pulmonary airflow (PAF) Rise-times during utterance of phrases for high pulmonary noise speakers (HPNS) and low pulmonary noise speakers (LPNS).

low stoma noise speakers (Subjects E and F) exceeded the averages of two high stoma noise speakers (Subjects A and C). The mean Major Peak value for only one low stoma noise speaker (Subject D) was less than that of the high stoma noise speakers. No clear pattern of airflow rate distinguishes speakers of different levels of perceived stoma noise. Moreover, with respect to Rise-time, no pattern completely distinguishes HPNS (Subjects A, B, C) from LPNS (Subjects D, E, F). It can be seen in Figure 4 that a slight trend exists for more rapid Rise-time means and ranges for HPNS. It is possible that excessive tracheal and tracheastoma tensions are created by forceful thoracic/abdominal contraction. The resulting tense, dense tissue surfaces in the HPNS may generate and resonate frequencies characteristic of stoma noise, and therefore, interact with aerodynamic energy to produce noise.

The observations made on the relationship between noise from the tracheastoma and PAF airflow rate magnitudes and Rise-times can lead to tenuous conclusions only, since the judgments of level of speaker stoma noise were based upon speech materials other than those utilized to derive airflow data.

CHAPTER V

CONCLUSIONS AND IMPLICATIONS

I. SUMMARY

This investigation was aimed at gaining a greater understanding of the pulmonary tract as it functions in the altered speech producing mechanisms of esophageal speakers. The study set out to determine whether the pulmonary tract participated in esophageal speech in a unique manner. Pulmonary airflow rate was monitored and analyzed as a means of indicating participation of the pulmonary tract. The independent variables were manner of production, voicing, syllabic position and perceived level of stoma noise. Six laryngectomees considered as predominantly using the inhalation method of air intake were categorized into high and low pulmonary noise speakers. Each speaker uttered 24 two word propositional phrases in a conversational manner. The phrases contained plosives and fricatives--/p, b, s, z/--which were placed in initial and final syllabic positions. The pulmonary airflow rate was monitored by a pneumotachograph and recorded simultaneously with the audio speech signals on graphic paper.

II. CONCLUSIONS

On the basis of the data gathered, the following conclusions seem warranted:

1. The pulmonary tract does not appear to participate in continuous esophageal speech in a unique manner in terms of frequency or magnitude of PAF peaks as a function of phonemes differing in manner of production, voicing, or syllabic position. While the airflow component of blowing is not phoneme specific, it does coexist generally while chains of phonemes are produced.
2. No differences in frequency or magnitude of PAF peaks seem to exist between HPNS and LPNS with respect to the participation of the pulmonary tract in esophageal speech.
3. Airflow rate in esophageal speakers using the inhalation method of air intake does not distinguish between HPNS and LPNS, and therefore, does not appear to be a component of blowing that is singularly responsible for stoma noise heard in some esophageal speakers. This is true when using PAF Major Peak and Rise-time values as airflow measures.

III. IMPLICATIONS

There are additional airflow measures that future research might explore as possible factors associated with judgments of pulmonary blowing noises which tend to depreciate the speech signal. Such factors are duration of pre-signal blowing, duration of pre-signal blowing of PAF airflow values greater than one-half or three-fourths of the PAF Major Peak values, duration of post-signal onset blowing of airflow values greater than one-half or three-fourths of the PAF Major Peak values, or possibly, a combination of durations of pre- and post-signal blowing of airflow values greater than one-half or three-fourths of the PAF value.

Non-aerodynamic factors worthy of investigation as possible bases for pulmonary noise are stoma size and shape, quality and physiology of stoma tissue, and pulmonary respiration. At this point, it is reasonable to hypothesize that a combination of these factors prompts the noise which contributes to the depreciation of esophageal speech. These factors plus the aerodynamic factors might be compared among esophageal speakers demonstrating varying degrees of pulmonary noise during speech.

The results of this investigation tend to negate the necessity of creating special clinical procedures for inhalers aimed at accommodating unique pulmonary activity for plosive and fricative phonemes in an ongoing phonetic context. Pulmonary and respiratory activity need not be expected to change routinely during connected speech for production of phonemes of varying levels of difficulty. Additionally, if stoma noise coexists with esophageal speech, several factors should be included for clinical consideration as possible causes of such noise.

REFERENCES

Berg, J. van den, Moolenaar-Bijl, Annie, and Damste, P. J., Oesophageal speech. Folia Phoniatr., 10, 65-84 (1958).

Berlin, C. I., Clinical measurement of esophageal speech: I. Methodology and curves of skill acquisition. J. Speech Hearing Dis., 28, 42-51 (1963).

Berlin, C. I., Clinical measurement of esophageal speech: III. Performance of nonbiased groups. J. Speech Hearing Dis., 30, 174-183 (1965).

Curry, E. T., and Snidecor, J. C., Physical measurement and pitch perception in esophageal speech. Laryngoscope, 71, 415-424 (1961).

Dean, C. R., A study of the perceptual and durational characteristics of normal, superior electrolarynx, and superior esophageal speech. An unpublished proposal submitted to the faculty of the Speech and Hearing Sciences Section, Stanford University (October, 1968).

DiCarlo, L. M., Amster, W. W., and Herer, G. R., Speech After Laryngectomy. Syracuse: Syracuse University Press (1955).

Diedrich, W. M., and Youngstrom, K. A., Alaryngeal Speech. Springfield, Ill.: Charles C. Thomas Publishers (1966).

Fairbanks, G., Everitt, W. L., and Jaeger, R. P., Transactions: Institute of Radio Engineers, Professional Group on Acoustics, 7-11 (August, 1954).

Gardner, W., Rehabilitation after laryngectomy. Public Health Nursing, 43, 612-615 (1951).

Greene, Margaret C. L., The Voice and its Disorders. 2nd ed., Philadelphia: J. B. Lippincott Company (1964).

Hoople, G. D., and Brewer, D. W., Voice production in the laryngectomized patient. Ann. Otol. Rhinol. Laryng., 63, 640-650 (1954).

Howie, T. O., Rehabilitation of patients after total laryngectomy. Occupat. Ther., 26, 372-383 (1947).

Hyman, M., An experimental study of artificial larynx and esophageal speech. J. Speech Hearing Dis., 20, 291-299 (1955).

Kallen, L. A., Vicarious vocal mechanisms: The anatomy, physiology and development of speech in laryngectomized persons. A. M. A. Arch. Otolaryng., 20, 460-503 (1934).

Kytta, J., Finnish esophageal speech after laryngectomee. Acta Otolaryng., Suppl. 195 (1964).

Martin, H., Rehabilitation of the laryngectomee. Cancer, 16, 823-841 (1963).

Mason, Marie, The rehabilitation of patients following surgical removal of the larynx. J. Laryng. Otol., 64, 759-770 (1950).

Moolenaar-Bijl, Annie, Some data on speech without larynx. Folia Phoniatr., 3, 7-9 (1951).

Moolenaar-Bijl, Annie, Connection between consonant articulation and the intake of air in oesophageal speech. Folia Phoniatr., 5, 212-215 (1953).

Morrison, W. W., and Fineman, S., The production of the pseudovoice after total laryngectomy. Trans. Am. Acad. Ophthalmol. Otolaryng., 41, 631-634 (1936).

Robe, Evelyn Y., Moore, P., Andrews, A. H., Jr., and Holinger, P. H., A study of the role of certain factors in the development of speech after laryngectomy: 3. Coordination of speech with respiration. Laryngoscope, 66, 481-499 (1956).

Seeman, M., Zur pathologie der Ösophagusstimme. (Contribution to the pathology of the esophageal voice.) Folia Phoniatr., 10, 44-50 (1958).

Shipp, T., Frequency, duration, and perceptual measures in relation to judgments of alaryngeal speech acceptability. J. Speech Hearing Res., 10, 417-427 (1967).

Snidecor, J. C., Speech Rehabilitation of the Laryngectomized. 2nd ed., Springfield, Ill.: Charles C. Thomas (1962).

Snidecor, J. C., and Curry, E. T., Temporal and pitch aspects of superior esophageal speech. Ann. Otol. Rhinol. Laryng., 68, 623-636 (1959).

Snidecor, J. C., and Isshiki, N., Air volume and air flow relationships of six male esophageal speakers. J. Speech Hearing Dis., 30, 205-216 (1965a).

Snidecor, J. C., and Isshiki, N., Vocal and air use characteristics of a superior male esophageal speaker. Folia Phoniatr., 17, 217-232 (1965b).

Stetson, R., Esophageal speech for any laryngectomized patient. Arch. Otolaryng., 26, 132-142 (1937).

Thompson, R. F., Statistical Methods Useful in Medicine. [n. p.] (1965).

Williams, H. N., Speech rehabilitation of the laryngectomized patient. CA - A Cancer Journal for Clinicians, 11, 126-130 (1961).

APPENDICES

APPENDIX A

STIMULUS SPEECH SENTENCES

1. The boy put the newspaper on the step.

2. He found that very few are living on what they have.

3. It's zero and I'm freezing in this breeze.

APPENDIX B

RATING OF STOMA NOISE

Judge: ______________________

INSTRUCTIONS:

You are about to hear 24 recorded samples of alaryngeal speech. Listen to the tape in its entirety. During the replay of the tape, rate only the parameter of stoma noise in terms of its intensity, frequency of occurrence, and the impairment to speech intelligibility by the stoma noise. On a 5-point scale of stoma noise, "1" equals low stoma noise and "5" equals high stoma noise. As an example, low stoma noise means that the noise perceived as being absent or of low intensity, seldom occurs throughout the sample, and speech intelligibility is not judged to be greatly impaired by stoma noise.

SPEECH SAMPLES:

1. _____	13. _____
2. _____	14. _____
3. _____	15. _____
4. _____	16. _____
5. _____	17. _____
6. _____	18. _____
7. _____	19. _____
8. _____	20. _____
9. _____	21. _____
10. _____	22. _____
11. _____	23. _____
12. _____	24. _____

APPENDIX C

TWO WORD PHRASES

1. Toplamb
2. Bosslamb
3. Gauzelamb
4. Toplamb
5. Ballboss
6. Ballzomb
7. Ballboss
8. Ballpot
9. Ballboss
10. Roblamb
11. Ballpot
12. Ballzomb
13. Ballsob
14. Toplamb
15. Roblamb
16. Bosslamb
17. Roblamb
18. Ballpot
19. Gauzelamb
20. Bosslamb
21. Gauzelamb
22. Ballsob
23. Ballzomb
24. Ballsob