1977

A Quantitative Study Based on a Sonographic Examination of Four Vowel Sounds in Alaryngeal Speech

Cheryl Ann Schultz
Portland State University

AN ABSTRACT OF THE THESIS OF Cheryl Ann Schultz for the Master of Science in Speech Communication: Emphasis in Speech Pathology/Audiology presented May 3, 1977.
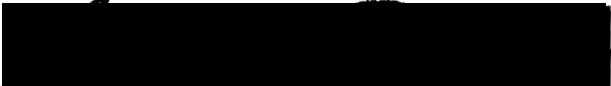
Title: A Quantitative Study Based on a Sonographic Examination of Four Vowel Sounds in Alaryngeal Speech.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

Mary E. Gordon, Chairman

Robert H. English, Ph.D.

Robert L. Casteel, Ph.D.

Ronald E. Smith, Ph.D.

Laryngectomy, as a treatment for malignant laryngeal lesions, requires the patient to seek a substitute method of producing speech. Three types of alaryngeal speech were described: esophageal, Asai, and artificial larynx. One consideration in deciding which mode of speech is best for the patient is how closely each type of alaryngeal speech approximates normal.

This investigation was an objective examination of

esophageal, Asai and artificial larynx speech as compared with normal in males and females. Comparisons were based on production of four vowels /i/, /æ/, /ɑ/, /u/. Eight subjects were selected, one male and one female in each type of alaryngeal speech. All speakers were recorded producing the four vowels three times. Each production was analyzed sonographically to determine fundamental frequency and formants 1, 2, and 3.

The following information was obtained to answer the questions posed:

1. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by esophageal speakers and normal speakers?

Fundamental frequencies produced by male and female esophageal speakers were 60 and 70 Hz lower than their normal counterparts, respectively. Formant frequency values for both esophageal speakers showed little relationship to normal, varying both higher and lower in no predictable pattern.

2. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by Asai speakers and normal speakers?

The male Asai speaker produced a fundamental frequency approximately 30 Hz higher than the normal male speaker. Formant 1 was found to be lower than normal; however, this trend did not continue in formants 2 and 3.

Female Asai fundamental frequency values, only 20 Hz

lower than normal, suggest that the Asai technique may provide the female laryngectomee with the closest approximation to normal.

3. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by speakers using electronic vibratory devices and normal speakers?

Laryngectomees using an artificial larynx select the fundamental frequency they desire in operating the mechanical device. These chosen values are considerably lower than normal. As with the other alaryngeal speakers, formant values were scattered above and below normal formant frequencies.

4. What is the difference between male and female production of these vowels in each of the four modes of speech?

Male alaryngeal speakers showed much greater consistency in the production of fundamental frequency over the range of the four vowels. Female alaryngeal speakers varied as much as 53 Hz from one vowel to another. Artificial larynx speakers are not considered because of the mechanical production of fundamental frequency.

5. What is the difference between esophageal, Asai and electronic vibratory device speakers for the four vowels examined?

Esophageal speakers produce a low fundamental frequency when compared to Asai speakers. Artificial larynx also produces a low fundamental frequency, but this is at the

discretion of the user. Data on formant frequencies were inconsistent and provided minimal information.

Conclusions derived from this investigation must be guarded due to the small sample population. If the data on these subjects is representative of larger populations, it may be inferred that the fundamental frequency of female Asai speakers is considerably closer to normal than the other two types of alaryngeal speech studied. Male alaryngeal speakers appear to have a broader choice of fundamental frequency ranging above and below normal values.

Further study is suggested on the quantitative aspects of alaryngeal speech using a considerably larger population of subjects. Extreme variations noted in some instances in formant relationships only clarify the need to minimize individual differences through a larger sample.

A QUANTITATIVE STUDY BASED ON A SONOGRAPHIC

EXAMINATION OF FOUR VOWEL SOUNDS

IN ALARYNGEAL SPEECH

by

CHERYL ANN SCHULTZ

A thesis submitted in partial fulfillment of the
requirements for the degree of
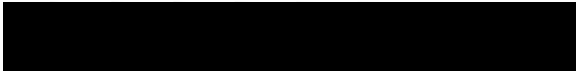
MASTER OF SCIENCE
in
SPEECH COMMUNICATION: EMPHASIS IN SPEECH PATHOLOGY/AUDIOLOGY

Portland State University
1977

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Cheryl Ann Schultz presented May 3, 1977.

Mary E. Gordon, Chairman

Robert H. English, Ph.D.

Robert L. Casteel, Ph.D.

Ronald E. Smith, Ph.D.

APPROVED:

Robert W. Vogelsang, Head, Department of Speech Communication

Stanley E. Rauch, Dean of Graduate Studies and Research

ACKNOWLEDGMENTS

Reflecting back over the events which have ultimately resulted in the completion of this project, certain personalities stand out as being invaluable contributors to the effort.

First and foremost, I must express profound gratitude to my husband, Gary Schultz, for bearing up under what must have been trying periods for him throughout the pr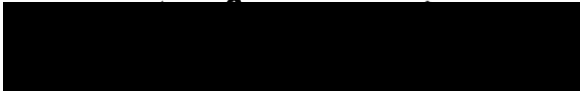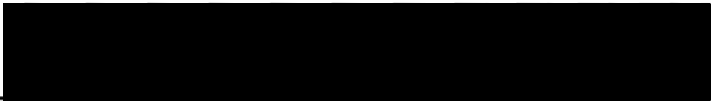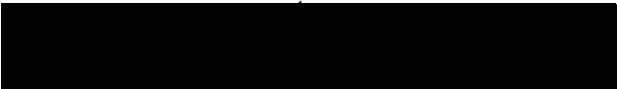ocess of making this project a finished product. He was a sounding board for ideas and frustrations, a traveling companion and equipment bearer during the data collection phase, and most of all a "rock" when I needed support.

Mary Gordon will never know how appreciated her efforts in this endeavor have been. Besides guidance and assistance as thesis Chairperson, she has given friendship during trying times, been understanding of frustrations, and calmed frazzled nerves.

Dr. Robert English must be thanked for allowing himself to be "hassled" all too frequently. His contributions to this project and to my life cannot be expressed.

Dr. Robert Casteel allowed himself to become a sounding board for trying out the ideas which led to this project. His guidance and help in making it into a workable plan; his

insights, thoughtfulness, and assistance have been invalu-
able.

Finally, I must acknowledge gratefully each of the
individuals who allowed themselves to be used as subjects
for this project. Each offered his time, his hospitality,
and many personal insights into the daily life problems and
feelings of laryngectomees which cannot be learned from
reading research studies or textbooks. They were each warm,
willing, eager subjects without whom this project would never
have been completed.

To all who participated in this project go my sincerest
thanks, and profound gratitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

INTRODUCTION

In the presence of malignant laryngeal lesions, the
patient's health and preservation of life become the para-
mount considerations in medical intervention, and to these
ends the ability to produce voice is sacrificed. According
to Perkins (1971), "a total laryngectomy involves the exci-
sion of the entire larynx from the trachea to the base of
the tongue, often including the hyoid bone." With the elim-
ination of voice, the patient's main task becomes the devel-
opment of a substitute communication method as a replacement
for his excised vocal capabilities. Numerous vocal substi-
tutes are possible for the laryngectomee, including esopha-
geal speech, electronic vibratory devices commonly referred
to as artificial larynx, and surgical laryngoplasty resulting
in the use of Asai speech.

Boone (1971) reported that while laryngeal cancer may
befall an individual of any age, its most frequent victims
are in their fifties and sixties. Incidence of laryngectomy
among males is much higher than among females, with males
accounting for 18,600 of the 23,000 cases as reported by
Snidecor (1971). Laryngectomy surgery is performed with a

survival rate of about 3,000 per year, adding substantially to those 23,000 cases reported alive in the United States in 1968 (Snidecor, 1971). According to the National Cancer Institute, there has been a 75 percent increase in laryngeal cancer in the past decade, but the mortality rate has risen 42 percent (Diedrich and Youngstrom, 1966). With survival, comes the need to develop a new means of communication.

Approximately 60 to 70 percent of the laryngectomy patients attempting to learn esophageal speech meet with success, according to Snidecor (1971). While these figures can be encouraging to the prospective laryngectomee, what of the 30 to 40 percent who do not acquire the skills of esophageal speech? For them, there is a great need to find an alternative to esophageal speech. Two possibilities are surgical laryngoplasty and the electronic vibratory device.

When considering alternative modes of speech, both objective and subjective elements must be explored by patient and physician. Much has been written about the subjective aspects of alaryngeal speech (Snidecor, 1968, 1971), but little research has been reported on objective characteristics.

STATEMENT OF PURPOSE

The purpose of this investigation was to identify quantitatively those modes of alaryngeal speech which most closely approximate "normal" speech in the voice parameters measured.

This investigation was undertaken to objectively assess certain quantitative aspects of three alaryngeal speech methods. The four cardinal vowels /a, æ, u, i/ produced by esophageal speech; Asai speech following surgical laryngoplasty; an electronic vibratory device; and normal speech were examined sonographically to identify the fundamental frequency of each mode of voice production, as well as the identification of frequencies of formants 1, 2, and 3.

The specific questions this investigation sought to answer were:

1. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by esophageal speakers and normal speakers?

2. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by Asai speakers and normal speakers?

3. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by speakers using electronic vibratory devices and normal speakers?

4. What is the difference between male and female production of these vowels in each of the four modes of speech?

5. What is the difference between esophageal, Asai and electronic vibratory device speakers for the four vowels examined?

CHAPTER II

REVIEW OF THE LITERATURE

When patients undergo a total laryngectomy, no laryn-
geal structure remains; thus, breathing must be accomplished
via a tracheostoma formed at the top of the trachea and pro-
truding out through the lower neck at the midline. Following
surgery the patient is left without the capability of pro-
ducing voice. Selection of a substitute voice for the laryn-
gectomee will take into consideration the following:

THREE TYPES OF ALARYNGEAL SPEECH

Esophageal Speech

Esophageal speech has become the preferred method of
alaryngeal speech in approximately 64 percent of 3,366 laryn-
gectomized patients surveyed by the American Cancer Society
(Diedrich and Youngstrom, 1966). While three separate meth-
ods exist for producing esophageal speech, the majority of
esophageal speakers employ the "injection method" (Boone,
1971). This method requires the speaker to use his tongue
as a driving mechanism, trapping air between tongue, hard
and soft palates, and pharynx. The compressed air is forced
backward into the hypopharynx and into the open esophagus
where it is entrapped for phonation (Boone, 1971). Air

trapped in the esophagus then is expelled, activating vibra-
tory movement of the pharyngo-esophageal junction, which is
thought to be largely comprised of the cricopharyngeus
muscle; there sound is produced (Perkins, 1971). Using the
normal processes of resonation and articulation, the esopha-
geal speaker is able to produce intelligible speech in this
manner.

Asai Speech

Asai speech was named for the Japanese laryngologist,
Ryozo Asai, who pioneered the surgical technique. It was
first reported to the medical world in October, 1965, at the
Eighth International Congress of Otorhinolaryngology in
Tokyo, Japan.

Asai laryngoplasty is usually accomplished in three
stages or operations. The entire technique is designed to
create a dermal tube buried under the skin at the anterior
midline of the neck. This tube reaches from the upper end
of the remaining tracheal tube into the hypopharynx (see
Figure 1). With surgery completed, the patient is able to
occlude the permanent tracheostoma used for breathing,
thereby shunting expired air up through the dermal tube
producing sound as the passing air sets up a vibration of
the walls of the dermal tube. From the tube, sound proceeds
into the oral and nasal cavities for articulation and resona-
tion, thus creating speech. Immediately following the third
stage of surgery to imbed the dermal tube, this speech produc-
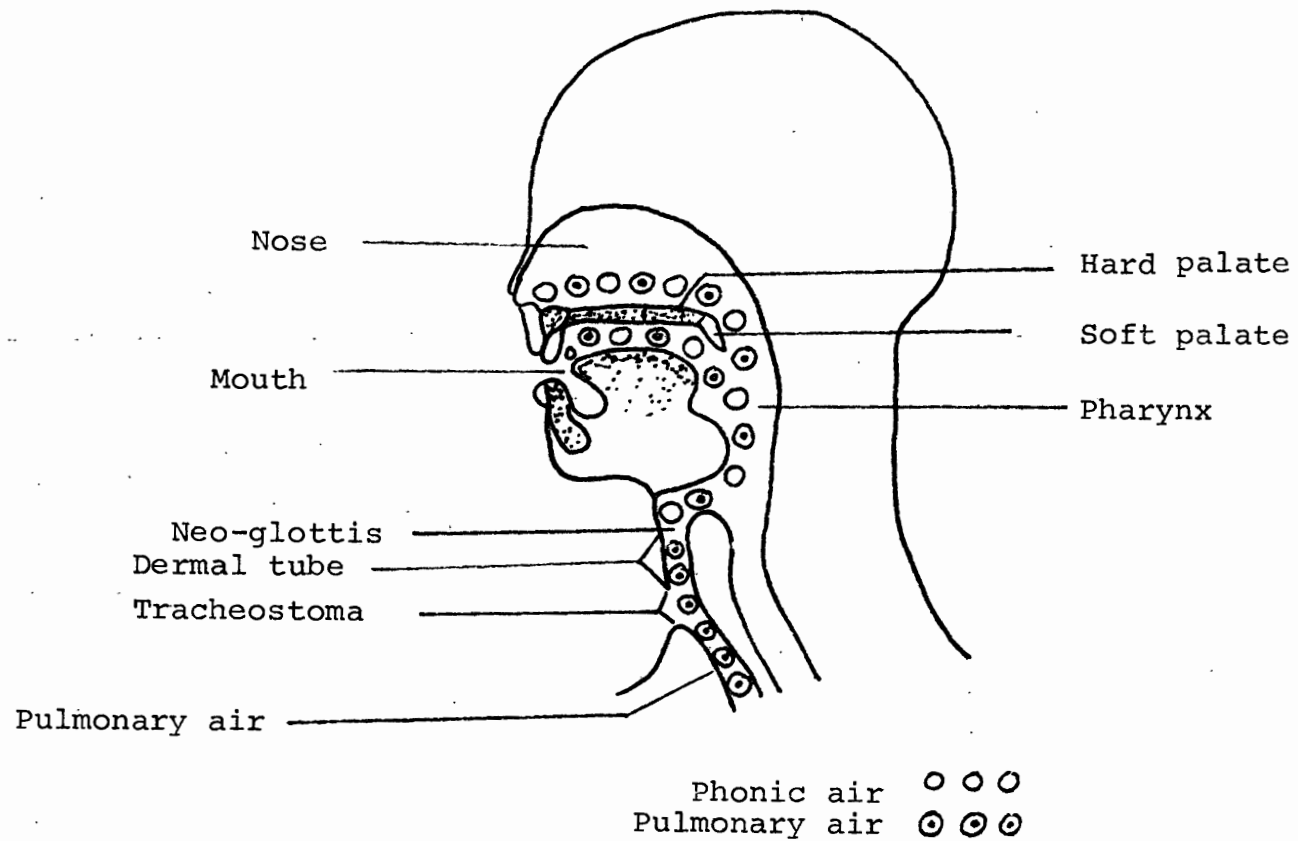
Figure 1. The head and neck after Asai surgery (modified from Snidecor, 1971).

tion process may be performed by the patient as Asai speech does not require a training period (Miller, 1971). Using expelled pulmonary air and directing it through the dermal tube with digital occlusion of the tracheostoma, a substitute voice is available to the patient without delay.

Artificial Larynx Speech

Though several models of artificial larynges are available to the patient, the Western Electric Electrolarynx is widely used. This handheld device, with its vibrator at one end, is placed firmly against the neck in the area of the hypopharynx. A pushbutton control operated by the thumb results in sound production. These sound vibrations penetrate the hypopharyngeal skin to enter the vocal tract for resonation and articulation for speech (Boone, 1971).

A training period is necessary before the patient will be adept at speaking with an artificial larynx. He must learn proper placement of the instrument for his own best sound, as this varies from patient to patient. Additionally, he must learn to articulate consonants with no pulmonary airflow, thus, necessitating some minor changes in his articulatory patterns. Finally, he must learn to coordinate his operation of the on-off switch with the flow of speech to avoid excesses of noise, i.e., sound production between phrases or utterances (Boone, 1971).

Following the training process the patient has speech available without the need for breath support. Thus, many

esophageal patients use the artificial larynx as a back-up during training or when fatigued. Boone (1971) and Diedrich and Youngstrom (1966) view esophageal speech and artificial larynx speech as compatible to provide the laryngectomee with a ready speech source for all occasions.

QUANTITATIVE ASSESSMENT OF NORMAL PHONATION

Vocal fold vibration in the normal larynx may produce a wide variety of tones. Fairbanks (1960) indicated relaxed, normal phonation may vary in one person, covering a range of nearly two octaves. Normally, however, an individual's pitch is distributed in such a manner as to indicate a definite central tendency (Zemlin, 1968).

When expressed in cycles per second (hereafter called Hertz and expressed Hz), the fundamental frequency may be derived as a direct result of the number of times per second the vocal folds vibrate (Kaplan, 1971). Kaplan (1971) has indicated the usual range of fundamental frequency is from 100 to 400 Hz, with variations being the result of variations in the length (Hollien, 1960) and thickness (Boone, 1971; Hollien, 1962; Hollien and Curtis, 1960) of each individual's vocal folds.

With vocal fold length of approximately 18 mm (Boone, 1971) to between 23 mm and 27 mm (Gray and Wise, 1959), the adult male fundamental frequency is reported to be approximately 125 Hz to 130 Hz (Boone, 1971; Zemlin, 1968). Peter-

son and Barney (1952) found average fundamental frequencies for two vowel sounds to range from 124 Hz on /ɘ/ to 141 Hz on /u/ in adult males. Mysak (1959) found, in addition to slight variations in fundamental frequency depending on whether the task was reading or impromptu speaking, there was evidence of male pitch rising with age. His results indicated that in the middle-aged group (32 to 62 years, mean age 47.9 years) the average fundamental frequency was 107.9 Hz. In the elderly group (65 years to 79 years, mean age 73.3 years), the average fundamental frequency had risen to 120.0 Hz. In the very elderly group (80 years to 92 years, mean age 85.0 years), the average fundamental frequency had risen further to 137.1 Hz. Each of these values is for impromptu speaking, with reading values found to be 3 to 5 Hz higher. Mysak (1959) attributed the noted changes to multiple factors, primarily changes in the laryngeal structure.

Female vocal folds are considerably shorter than those of the male. Boone (1971) reported them to be 10 mm in length, while Gray and Wise (1959) found them to be five-sevenths the length of the male folds, or approximately 16 to 19 mm. The shorter female vocal folds result in a higher frequency tone production, ranging from 210 Hz on /æ/ to 235 Hz on /i/ in the Peterson and Barney (1952) study. Zemlin (1968) reported 260 Hz as contrasted with Boone's (1971) finding of 200 Hz.

Accompanying the production of the fundamental frequency are overtones produced by segmental vibration of the vocal folds which occur simultaneously with the production of the fundamental frequency (Hanley and Thurman, 1962). Thus, the vocal tone at the level of the larynx is a complex tone.

As the tone is transmitted through the pharynx and oral cavities resonation occurs, damping some frequencies while enhancing others. Articulatory movements create a particular set of resonating surfaces and cavities to produce a given phoneme. The articulators are capable of creating ". . . a whole series of resonant frequencies (Bogert and Peterson, 1957)." It is the standardization of these articulatory movements to specific cavity sizes and shapes which provide the concentrations of energy at certain frequencies to make each phoneme unique (Bogert and Peterson, 1957; Hanley and Thurman, 1962).

To quantitatively assess voice production, the spectrograph was designed to analyze speech in its acoustical form (Gray and Wise, 1959). A graphic illustration is obtained of the formants present in a sample of speech. (Appendix A illustrates sonographic representations of the vowel /æ/.) Gray and Wise (1959) define the term formant after Fletcher as consisting of:

> . . . two (sometimes three, seldom one) frequency regions of prominence in the speech sound as observed in the outside air. The frequency locations of these regions of accentuated intensity, known as formants,

are an important characteristic lending individual-
ity to the various periodic sounds.

The acoustic representation of a formant is tied to
the physiologic structures directly, stated Gray and Wise
(1959):

> . . . it is the existence in the vocal tract of
> these two major resonating cavities (pharyngeal and
> oral), together with the narrow passages connecting
> them with each other and with the external air,
> that produces the 'frequency regions of prominence'
> in the speech sounds, and differentiates them one
> from another.

They further stated ". . . each of these cavities thus cre-
ated, the throat and the mouth, produces one formant, the
coupling between them probably producing another (Gray and
Wise, 1959). Potter and Steinberg (1950) supported this
view stating that the ". . . formant frequencies for a given
vowel are probably directly related to the dimensions of the
vocal cavities."

Froeschels (1948) noted that the theory of Helmholtz
(and Stumpf) indicates ". . . each vowel consists of a basic
tone and several harmonic partial tones, the so-called over-
tones." Further, Stumpf is reported to state ". . . some of
the over-tones are especially loud (formants) and that these
have to be present in the tone mixture in order that the
vowel in question will be recognized (Froeschels, 1948)."
Denes and Pinson (1963) indicated that ". . . usually the
first three or four formant frequencies are adequate for
recognition."

Thus, for each individual phoneme one can expect to

see, via spectrographic analysis, a unique and rather con-
stant set of formants based on the standard articulatory
movements required to make that phoneme identifiable.

With individual differences in cavity shape and size,
it is difficult for two different vocal tracts to produce
matching sounds. Potter and Peterson (1948) suggested
formant relationships, rather than exact frequency match,
may be responsible for determining vowel sounds.

Formant frequencies of vowel sounds produced by 76
speakers were studied by Peterson and Barney (1952). For
purposes of this investigation, the data in Table I excerpted
from that study are of significance.

It may be noted from the data that the ratios among
formants appear to remain ". . . fairly constant over the
range of speakers (Peterson and Barney, 1952)."

TABLE I

FUNDAMENTAL FREQUENCIES AND FORMANT FREQUENCIES
FOUR CARDINAL VOWELS
PETERSON AND BARNEY STUDY
1952

Adult Males	/i/	/æ/	/ɑ/	/u/
Fundamental Frequency	136	127	124	141
Formant 1	270	660	730	300
Formant 2	2290	1720	1090	870
Formant 3	3010	2410	2440	2240
Adult Females	/i/	/æ/	/ɑ/	/u/
Fundamental Frequency	235	210	212	231
Formant 1	310	860	850	370
Formant 2	2790	2050	1220	950
Formant 3	3310	2850	2810	2670

QUANTITATIVE ASSESSMENT OF ALARYNGEAL SPEECH

Quantitative Assessment of Esophageal Speech

With the excision of the larynx and the implementation of esophageal speech, gross changes in frequency values would be expected. Probably due to the statistical preponderance of male laryngectomees (Snidecor, 1968), most studies have dealt with male esophageal speakers. Thus, specific data on female voices are limited.

Fundamental frequency values characteristically shift downward in esophageal speech. Snidecor and Curry (1961) reported a mean fundamental frequency for six esophageal speakers of 62.8 Hz. Rollin, in an unpublished dissertation cited by Curry (1960), found a mean fundamental frequency of 65.6 Hz for ten speakers. Shipp (1967), in a study of the acceptability of esophageal speech, found the mean fundamental frequency of what he termed "six best speakers" of thirty-three subjects to be 94.38 Hz. His group judged to be of lower acceptability (N=16) were found to have a mean fundamental frequency of 64.74 Hz, in contrast to a group rated as more acceptable (N=17) whose mean fundamental frequency was 84.4 Hz.

The Shipp data may not be directly comparable to the data previously cited due to a difference in measurement technique. Shipp used a wave-by-wave assessment technique; whereas, Snidecor, Curry and the others employed an averaging technique. This discrepancy results in slight variations in

the final outcome (Weinberg and Bennett, 1972). Regardless of the method of measurement employed, the general trend downward in fundamental frequency is definitely present.

Weinberg and Bennett (1972) divided their sample population of esophageal speakers into categories of male and female to determine mean fundamental frequency. Results indicated a difference between male and female, with eighteen males averaging 58 Hz, and fifteen females averaging 87 Hz. Their combined mean fundamental frequency was 69 Hz. Again, this study employed a wave-by-wave rather than averaging technique in measurement. Thus, these results are not totally comparable to those of Curry and Snidecor or Rollin.

The Weinberg and Bennett (1972) data suggest a significant difference in the esophageal speech production of males and females. Snidecor and Curry (1960) attempted to have a panel of judges identify female esophageal voices from a mixed group of recorded male and female esophageal voices. No success was found in the subjective task; hence, Snidecor and Curry concluded there was a great similarity between male and female esophageal voices.

While most acoustic data relative to esophageal speech have dealt with the fundamental frequency, Sisty and Weinberg (1972) investigated vowel formant frequencies in fourteen male and thirteen female esophageal speakers. Pertinent to the current study, their results for selected vowels are as follows:

TABLE II

FORMANT FREQUENCIES 1, 2, AND 3
FOUR CARDINAL VOWELS
MALE AND FEMALE ESOPHAGEAL SPEAKERS
SISTY AND WEINBERG - 1972

Males	/i/	/æ/	/ɑ/	/u/
Formant 1	401	791	984	459
Formant 2	2684	2059	1357	1213
Formant 3	3067	2762	2830	2666
Females	/i/	/æ/	/ɑ/	/u/
Formant 1	390	908	1031	435
Formant 2	2925	2412	1432	1134
Formant 3	3627	3106	3012	2835

When these values are compared with the Peterson and Barney data (1952), the formant frequencies for esophageal speakers are consistently higher. Sisty and Weinberg (1972) found an average increase in males on F_1 of 122 Hz, F_2 of 325 Hz, and F_3 of 321 Hz when compared to Peterson and Barney's (1952) normal group. Females in the study also demonstrated a systematic rise in formant frequency. Female average increases were 76 Hz on F_1, 211 Hz on F_2, and 208 Hz on F_3, again as compared with Peterson and Barney (1952).

From these comparisons, it becomes apparent that excision of the larynx does indeed alter the transmissive characteristics of the vocal tract. While Sisty and Weinberg (1972) do not feel this change can be fully explained on the basis of altered tongue positioning or mouth opening, they do suggest the possibility of these alterations being explained on the basis of a ". . . reduction in the effective

length of the vocal tract as a result of total laryngectomy surgery." The altered vocal tract apparently serves to elevate the formant frequencies. This position can be supported by the Diedrich and Youngstrom (1966) data which cinefluorographically illustrated the post-operative reduction in the vocal tract length as compared to its pre-operative state.

Evidence collected on esophageal speech suggests that when the larynx is excised and esophageal speech substituted for normal voice, variations will be noted in fundamental frequency and formant frequencies. The fundamental frequency appears to be lowered, due to the mass and size of the neo-glottis while the trend for formant values is upward due to altered size and shape of the resonating cavities.

Quantitative Assessment of Asai Speech

The phonatory capabilities available to the patient who has undergone Asai reconstruction following laryngectomy have not been fully investigated for trends in fundamental and formant frequencies. Curry and Snidecor (1973) noted several of the characteristics of Asai speech included higher pitch, more rapid, and breathier than esophageal speech. To study fundamental frequency, five Japanese Asai speakers, speaking Colloquial Japanese, were selected as subjects. They were asked to read a passage which was very closely compared with American dialect in language and voiced/voiceless structure, and, therefore, felt to be compatible for making cross-language comparisons. Curry and Snidecor (1973)

determined the average fundamental frequency of five sub-
jects was 97.08 Hz. Contrasted with the reported data on
esophageal speakers, this represents a trend toward the
mean fundamental frequency of normal speakers. Three males
in the group demonstrated mean fundamental frequencies of
96 Hz, 124 Hz, and 80 Hz respectively. Note that one of the
three male Asai speakers produced a mean fundamental fre-
quency comparable to the mean fundamental frequency of normal
male speakers in the Peterson and Barney study (1952). The
two female speakers using Asai voice exhibited mean funda-
mental frequencies of 150 and 81 Hz respectively. These
values are widely variant, one falling above the mean for
normal male speakers, the other falling below the normal
male mean. Neither of the female Asai speakers produced a
mean fundamental frequency comparable to the Peterson and
Barney (1952) normal female speakers.

Formant frequencies for Asai speakers have not been
reported in the literature. From the data available on
fundamental frequency, it appears that Asai speakers more
closely approximate normal fundamental frequency than do
esophageal speakers. Much further research is needed in the
area of Asai speech before conclusions can be made.

Quantitative Assessment of
Artificial Larynx Speakers

Little investigation of voice production via the use
of an artificial larynx has been done. While attitudes are

changing, the artificial larynx was not considered a desirable substitute for normal voice until recently (Boone, 1971). Control of the fundamental frequency is accomplished by a manual setting on the instrument, thereby making investigation of that parameter of artificial larynx speech unnecessary.

Coleman (1971), in a study employing normal speaking adult subjects, 10 male and 10 female, spectrographically investigated the effect of using the artificial larynx (specifically the Western Electric Company, Model 5, Electrolarynx). Fundamental frequency was set at 85 Hz $\pm$ 3 Hz for all subjects. Formant frequencies were obtained for vowels /i/ and /u/. Results are shown in Table III.

In addition, Coleman (1971) obtained listener judgments on maleness and femaleness. He found judges were able to correctly identify male and female voices 88 percent of the time. He suggested the cue to these distinctions lies in the vocal tract resonances of the speaker.

TABLE III

FORMANT FREQUENCIES 1, 2, AND 3
NORMAL SPEAKERS USING ARTIFICIAL LARYNX
COLEMAN - 1971

Male	/i/	/u/	Female	/i/	/u/
Formant 1	290	356	Formant 1	343	408
Formant 2	2400	1075	Formant 2	2707	1019
Formant 3	3270	2639	Formant 3	3454	2932

Since this research was conducted using normal laryn-geal speakers, caution must be exercised in directly trans-lating the results to compare with alaryngeal speakers whose vocal tract has been altered considerably.

CHAPTER III

METHODS AND PROCEDURES

GENERAL PLAN

This study investigated the variations in fundamental and formant frequencies 1, 2, and 3, for three types of alaryngeal speakers, and a control group of normal speakers. A spectrographic analysis of four vowel sounds (/æ, a, u, i/) produced by each speaker was performed from data recorded on a high-fidelity tape recorder.

SUBJECTS

Subjects for this investigation included the following:

1. Two adults (one male and one female) who had undergone Asai reconstructive surgery following laryngectomy and used Asai speech as their principle form of oral communication. The male Asai speaker was aged 62 at the time of recording and had used Asai speech for four months. The female Asai speaker was 56 years of age and had used Asai speech eight years prior to recording. Both subjects lived in the Los Angeles area.

2. Two adults (one male and one female) who had undergone total laryngectomy and used esophageal speech as their principle form of oral communication. The male esophageal

speaker was 52 years old at the time of recording, and had used esophageal speech for ten years. The female esophageal speaker was 56 years old and the exact length of time she had used esophageal speech is not known. The male esophageal speaker lived in Los Angeles, and the female esophageal speaker lived in Portland, Oregon.

3. Two adults (one male and one female) who had undergone total laryngectomy and used an artificial larynx as their principle form of oral communication. The male artificial larynx speaker was 72 years of age and it is not known how long he had used artificial larynx speech prior to recording. The female artificial larynx speaker was 52 years old and had used that form of communication for 19 years. Both artificial larynx speakers lived in the Los Angeles area.

4. Two adults (one male and one female) who are normal laryngeal speakers. They exhibited no current voice or speech abnormalities, nor had any abnormalities in their history. Both normal speakers were 63 years of age at the time of recording.

<center>INSTRUMENTATION</center>

A Uher tape recorder, model 4000 Report-L (see specifications in Appendix B-2), with accompanying microphone, was used to record each voice sample on Ampex tape, type 651. The recorder was operated at 3 3/4 inches per second.

This machine was selected for recording because of its port-
ability, as several subjects were recorded in Los Angeles
and others in Portland, Oregon.

The taped vowel productions were played into the Kay
Sonograph, Model 6061B (specifications are reported in
Appendix B-1) using a Type A display illustrating frequency
plotted on the ordinate and time on the abscissa. A Type B
display was subsequently overlayed on the graphic display
to indicate the relative intensity of each frequency repre-
sented in the Type A display.

PROCEDURES

Each of the following monosyllabic words were produced
three times by each subject: /tip/, /tup/, /tæp/, /tɑp/.
These vowel sounds were chosen because they are the corner
vowels on the standard vowel chart. The vowel portion of
each monosyllabic word was elongated as demonstrated by this
investigator. Consonants /t/ and /p/ were selected as the
acoustic environment for the vowels because both are voice-
less and would isolate the vowel formants. In addition /t/
is easy for alaryngeal speakers to produce in initiating
phonation.

The microphone was held 6 - 8 inches in front of the
mouth at the level of the chin. All recordings were made in
the subjects' homes with efforts made to eliminate background
noise.

The specific instructions given verbally to the subjects were:

> I am going to tape record your voice as you say these words: /tip/, /tup/, /tæp/, /tɑp/. I want you to say each word three times in a row and say them just the way I do.

A model was then presented demonstrating each monosyllabic word, elongating the vowel portion approximately one second. The model was as follows: "/tip/, /tip/, /tip/." The model was followed by "Let's practice them before we record."

Recordings were made after the subject demonstrated he could elongate the vowel to approximately one second in length.

ANALYSIS OF DATA

All vowel productions were analyzed spectrographically to determine fundamental frequency and the first, second, and third formant frequencies. Three productions of each vowel were analyzed for each speaker, the values averaged and expressed as the "mean fundamental frequency" and "mean formant frequencies."

After mean values for each vowel produced by each speaker were found these values were compared to answer questions posed in Chapter I. The mean values for each vowel were graphically compared among the four groups.

CHAPTER IV

RESULTS AND DISCUSSION

RESULTS

The purpose of this investigation was to objectively assess certain quantitative aspects of alaryngeal speech. Four vowels produced by esophageal speech, Asai speech and electronic vibratory device (artificial larynx) speech were examined to determine their relationship to normal speech production in both fundamental frequency and the frequencies of formants 1, 2, and 3. In addition, the variability of these parameters within the three modes of alaryngeal speech was examined.

To determine fundamental frequencies and frequencies of formants 1, 2, and 3 for each speaker, a taped sequence of vowels was analyzed sonographically. Each vowel was within the context of a monosyllable word (e.g.: /tip/) and each word was produced three times to provide mean values for fundamental, as well as formant frequencies, for each speaker. The four vowels represented in this investigation were the four corner (cardinal) vowels on the standard vowel chart; thus, extremes of production from high front to low back focus within the oral cavity were present.

To respond to each of the questions posed in Chapter I,

data were separated into productions by male and female
subjects in order to eliminate intrinsic sex-linked differ-
ences in voice production. No conclusions may be drawn
from the data gathered that would reflect characteristics
of voice production for a general population of alaryngeal
speakers or for normal speakers because the size of the
sample population was not sufficient to minimize individual
differences in voice production. These data reflect char-
acteristics of voice only for the eight subjects examined.

The resulting mean values of the fundamental frequency
and formants 1, 2, and 3 are shown in Table IV for each
vowel by each speaker. The data provide comparative infor-
mation on each form of alaryngeal speech with normal speak-
ers participating in this study. All data are expressed in
Hertz (Hz) for fundamental or formant frequency indicated.
Spaces in the charts marked with an asterisk (*) indicate
that no data were collected on these items. The same data
on fundamental frequencies is illustrated in Figure 2.
Figures 3 through 6 illustrate relationships of male and
female alaryngeal speakers to normal performance on each
vowel. Variability of each speaker among three productions
of the same vowel is shown on Tables V and VI. It can be
seen that all speakers varied in the formant frequencies of
vowel productions; however, with a few notable exceptions,
this variation seemed to be minimal.

Comparison of male alaryngeal speech to normal male

TABLE IV

FUNDAMENTAL FREQUENCY AND FORMANT FREQUENCIES 1, 2, AND 3
MEAN VALUES FOR THREE PRODUCTIONS OF FOUR CARDINAL VOWELS

MALE SPEAKERS					FEMALE SPEAKERS				
SPEAKER	F_0	F_1	F_2	F_3	SPEAKER	F_0	F_1	F_2	F_3
Vowel /i/									
Esophageal	65	333	1016	2133	Esophageal	112	250	3125	*
Asai	160	550	1366	2033	Asai	192	408	3183	3550
A. L.	104	325	900	2466	A. L.	80	400	2733	3366
Normal	138	275	2358	3100	Normal	221	350	2700	3216
Vowel /æ/									
Esophageal	72	1000	2066	3250	Esophageal	165	800	2441	3600
Asai	168	916	1500	2333	Asai	168	783	2316	3283
A. L.	98	800	1433	2200	A. L.	80	858	2313	3216
Normal	128	741	1850	2733	Normal	189	833	2200	2800
Vowel /ɑ/									
Esophageal	74	450	1100	1683	Esophageal	125	333	1166	1575
Asai	170	916	1566	*	Asai	162	850	1783	3300
A. L.	98	866	1533	2150	A. L.	80	616	1433	3300
Normal	120	716	1300	2600	Normal	192	933	1416	*
Vowel /u/									
Esophageal	72	383	1116	2450	Esophageal	128	408	1108	2141
Asai	157	700	1600	*	Asai	197	300	1166	1983
A. L.	104	483	1150	2641	A. L.	80	533	983	3033
Normal	130	300	1300	2250	Normal	208	383	1466	2783

Note: A. L. denotes speaker using artificial larynx
       * denotes information unavailable for that frequency
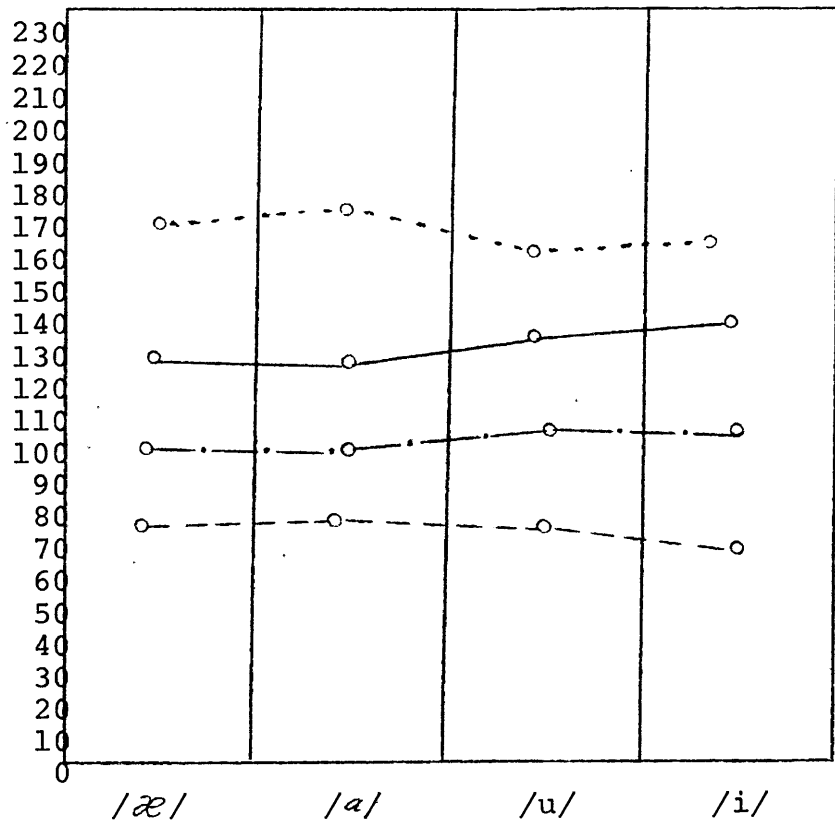
F$_0$ - ALL VOWELS - MALE SPEAKERS

F$_0$ - ALL VOWELS - FEMALE SPEAKERS

Legend: ——————— Normal Speakers
————— Esophageal Speakers
· – · – · – · Asai Speakers
—·—·—·— Artificial Larynx Speakers

Figure 2. Fundamental Frequencies of Vowels /æ/, /a/, /u/, /i/
Esophageal, Asai, Artificial Larynx and Normal Speakers
Male and Female
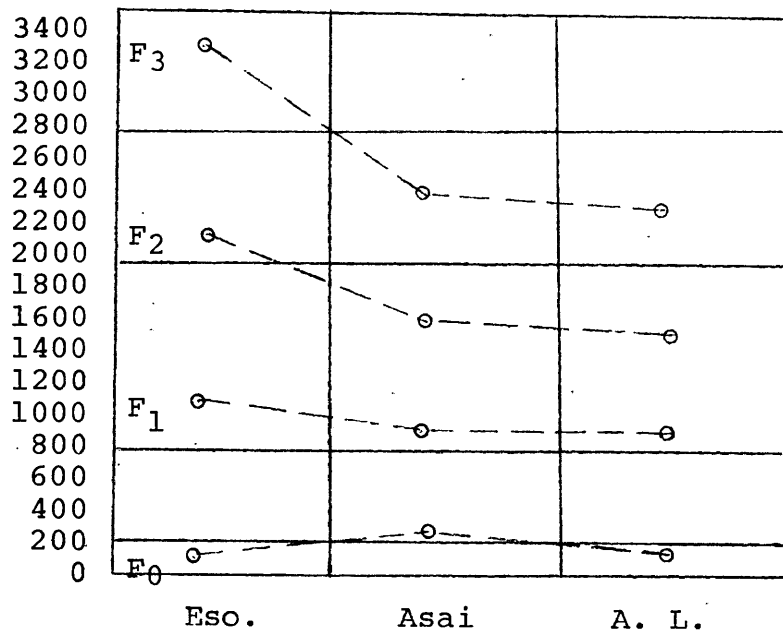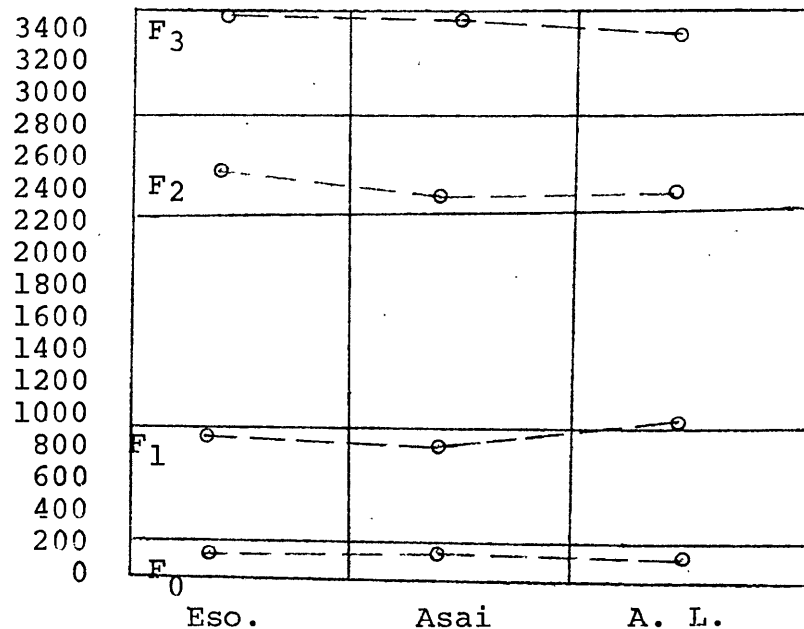
Note: Connecting lines denote normal speaker's frequency for F_0, F_1, F_2, F_3

Figure 3. Comparison of Alaryngeal Speakers to Normal Speakers on Fundamental and Formant Frequencies, Vowel /æ/.
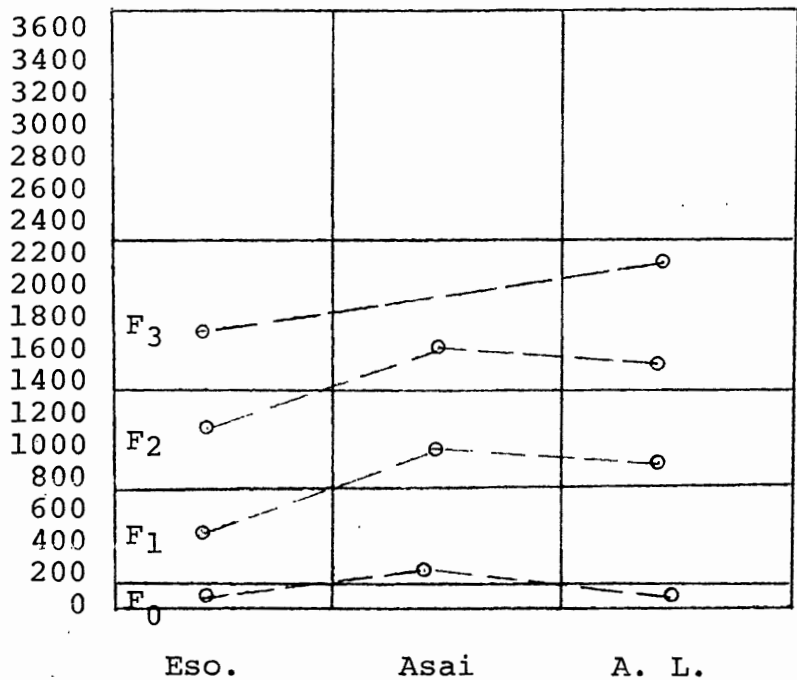
MALE SPEAKERS

F_0, F_1, F_2, F_3 - /a/

FEMALE SPEAKERS

F_0, F_1, F_2, F_3 - /a/

Note: Connecting lines denote normal speaker's frequency for F_0, F_1, F_2, F_3

Figure 4. Comparison of Alaryngeal Speakers to Normal Speakers on Fundamental and Formant Frequencies, Vowel /a/.

*No information available for F_3 for normal female and Asai male subjects.
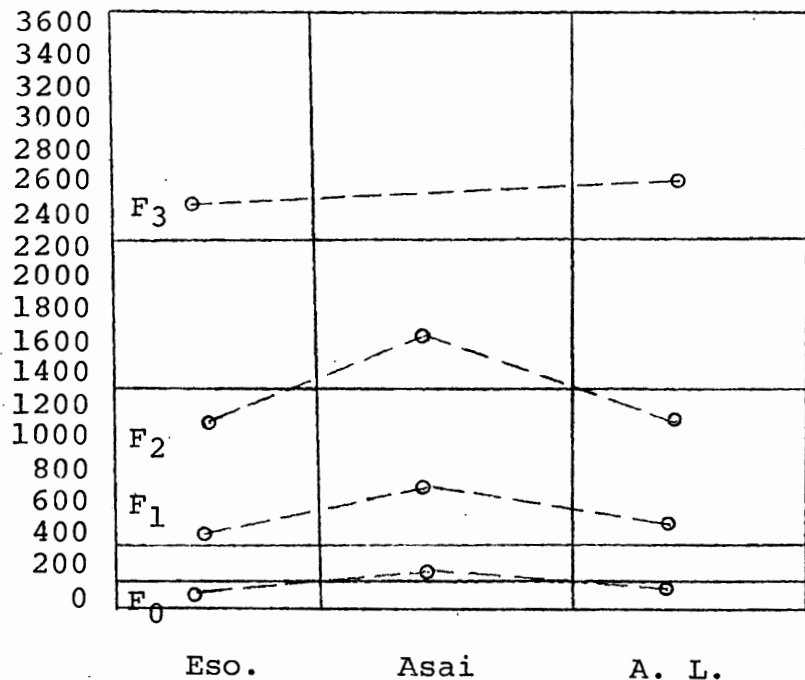
MALE SPEAKERS

F_0, F_1, F_2, F_3 - /u/
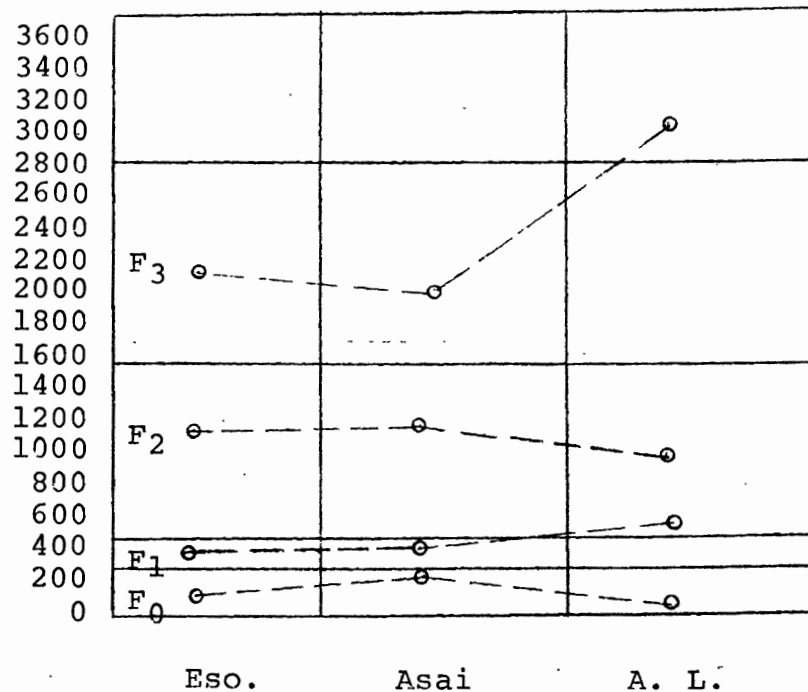
FEMALE SPEAKERS

F_0, F_1, F_2, F_3 - /u/

Note: Connecting lines denote normal speaker's frequency for F_0, F_1, F_2, F_3

Figure 5.   Comparison of Alaryngeal Speakers to Normal Speakers on Fundamental and Formant Frequencies, Vowel /u/.

*No information available for F_3 for Asai male speaker.
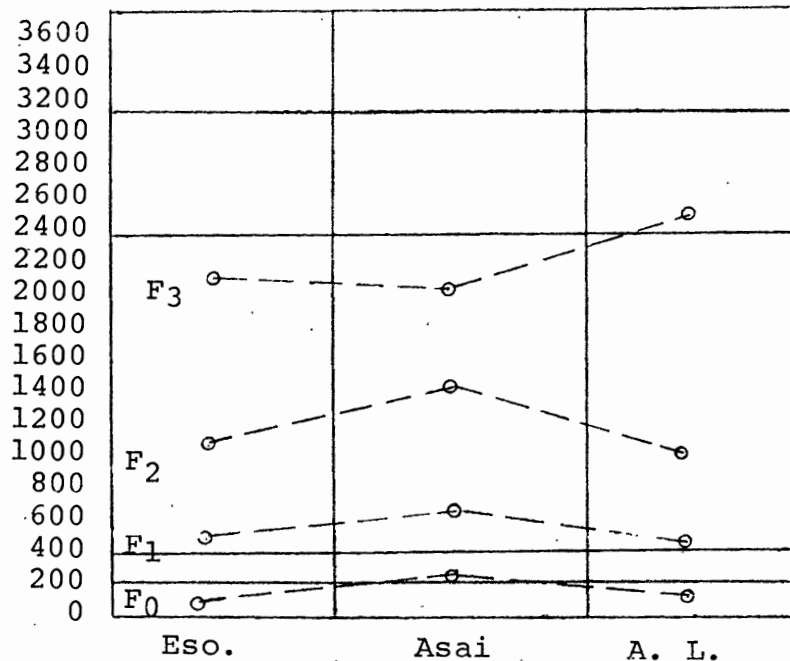
MALE SPEAKERS

F_0, F_1, F_2, F_3 - /i/
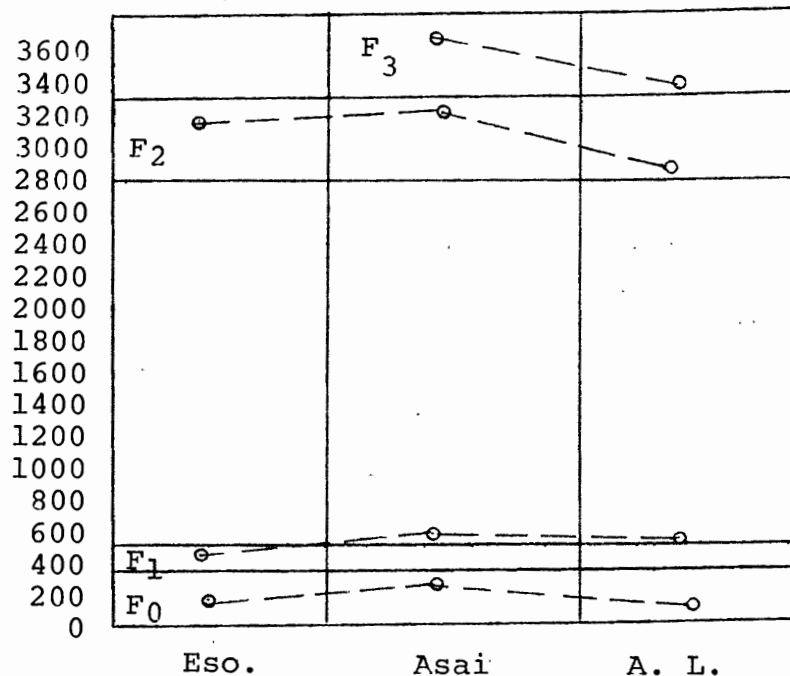
FEMALE SPEAKERS

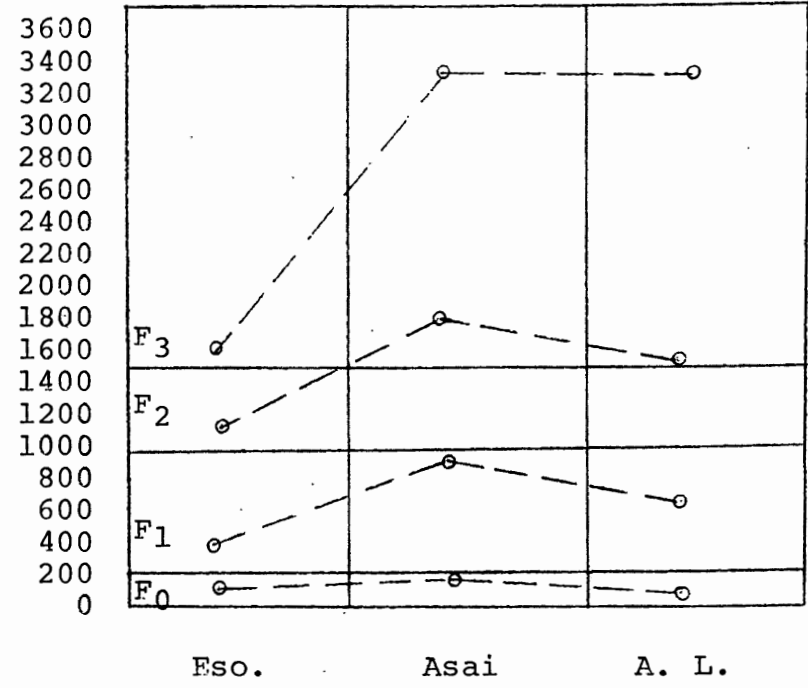F_0, F_1, F_2, F_3 - /i/

Note: Connecting lines denote normal speaker's frequency for F_0, F_1, F_2, F_3

Figure 6.   Comparison of Alaryngeal Speakers to Normal Speaker s on Fundamental and Formant Frequencies, Vowel /i/.

*No information available for F_3 for esophageal female speaker.

TABLE V

VARIABILITY OF FREQUENCIES OF FORMANTS
AND FUNDAMENTAL FREQUENCY
MALE SUBJECTS

VOWEL	F_0	F_1	F_2	F_3
Normal				
/æ/	128	725-750	1800-1900	2700-2750
/a/	120	650-750	1250-1350	2550-2650
/u/	128-136	300	1200-1500	2250
/i/	136-144	250-300	2300-2400	3000-3250
Esophageal				
/æ/	64-80	1000	2000-2100	3000-3400
/a/	72-80	300-750	1000-1200	1450-2150
/u/	72	350-400	1050-1200	2400-2500
/i/	60-72	300-350	1000-1050	1700-2400
Asai				
/æ/	160-176	800	1350-1650	2200-2550
/a/	152-184	850-1000	1550-1600	*
/u/	144-168	650-750	1500-1700	*
/i/	160	475-600	1250-1450	2000-2100
Artificial Larynx				
/æ/	96-104	775-825	1400-1450	2200
/a/	96-104	800-950	1350-1750	2050-2300
/u/	104	475-500	950-1300	2625-2650
/i/	104	300-350	850-950	2400-2500

*denotes no information available

TABLE VI

VARIABILITY OF FREQUENCIES OF FORMANTS
AND FUNDAMENTAL FREQUENCY
FEMALE SUBJECTS

VOWEL	F_0	F_1	F_2	F_3
Normal				
/æ/	176-200	800-850	2100-2250	2750-2850
/a/	192	900-1000	1350-1500	*
/u/	208	375-400	1400-1600	2750-2850
/i/	208-232	300-400	2600-2800	3200-3250
Esophageal				
/æ/	160-168	750-850	2400-2475	3550-3650
/a/	112-136	300-350	1100-1200	1500-1675
/u/	112-144	375-450	1000-1175	1900-2375
/i/	104-120	300-400	3100-3150	*
Asai				
/æ/	160-184	750-800	2250-2450	3000-3450
/a/	160-168	800-900	1750-1850	3250-3400
/u/	192-208	300-500	1150-1200	1950-2000
/i/	184-200	350-475	3150-3250	3500-3600
Artificial Larynx				
/æ/	80	800-900	2300-2400	3200-3250
/a/	80	550-700	1400-1500	3300
/u/	80	500-550	950-1000	3000-3100
/i/	80	400	2600-2950	3300-3500

*denotes no information available

speech can be observed by referring to Table IV. The esophageal male speaker shows a considerably lower fundamental frequency as compared with the normal, i.e., the mean fundamental frequency is approximately 70 Hz for the four vowels produced by the esophageal speaker while approximately 130 Hz was produced by the normal subject. For the vowel /æ/ the esophageal speaker produced formants which are higher in frequency than those of the normal speaker; however, this pattern did not occur for the other vowels. For the vowel /a/, lower formants were found. The vowel /u/ very closely approximates the normal /u/ across the formant range, while /i/ did not follow a particular pattern, but varied greatly.

The Asai male speaker showed consistently higher values for both fundamental frequency and formant 1 for all vowels investigated. A mean fundamental frequency of 163 Hz for the four examined vowels is considerably higher than the normal mean fundamental frequency of 129 Hz. Formant 2 is not consistent with normal, most variant is /i/, in which the frequency is nearly 1000 Hz lower than the normal. Data on formant 3 for the Asai male is sketchy and inconsistent. In many Asai vowel productions, no third formant could be located on the sonographic display.

The artificial larynx male used a fundamental frequency of approximately 102 Hz overall. His use of an electronic device negates any need for discussion of fundamental frequency as this factor is controlled by the mechanical

vibrator. As with the Asai speaker, this subject used higher formant 1 frequencies than the normal speaker, although in no case was there a considerable difference. Formant 2 was found to be inconsistent and showed no trend in variance as compared to normal with only /ɑ/ and /u/ relatively close to normal. Formant 3 values likewise showed a pattern of variance from normal and were very inconsistent.

Examination of the data collected indicates very little relationship among the three types of alaryngeal speech investigated. Fundamental frequencies for the esophageal and Asai speakers vary considerably (artificial larynx not considered for reasons stated above). Relationships among the formant frequency values also varies as illustrated in Figures 3 through 6, although some vowels produced by the same speakers showed normal patterns. Any further comparisons would be unwarranted by the data.

For the female speakers similar comparisons can be made to those seen in the male subjects. The esophageal female speaker used an average fundamental frequency of approximately 132 Hz, which is considerably below the 202 Hz found in the female normal speaker. Across all formants, the differences between normal and esophageal speaker were seen to be varied, with no pattern noted. Likewise, no pattern could be found for the fundamental frequency relationships of the individual vowels to their normal counterparts. It must be concluded the esophageal female voice represented here varies

widely and inconsistently from the normal voice examined.

The female Asai speaker studied was noted to have a fundamental frequency of 179 Hz. Of the female alaryngeal speakers, this most closely approximated the normal female voice. For the vowel /u/ all formants were lower, whereas, for the vowel /i/, all formants were higher. For the remaining two vowels studied, /æ/ and /ɑ/, no pattern of variance could be determined.

In examining the results of vowel production of the artificial larynx female subject, it appears /i/ is the most like normal production of formants 1, 2, and 3, with each formant being only slightly elevated from those of the normal subject. The overall picture presented by the artificial larynx female is far less varied than the other two alaryngeal female speakers. The low fundamental frequency is the result of a chosen frequency by the subject in using the mechanical device and does not reflect any physical characteristic of the subject's voice production. The implications of this chosen low fundamental frequency and its effects on formant values are uncertain.

Examination of the data collected on the female alaryngeal speakers shows similarities among all three speakers on the vowel /æ/ (see Figure 3). The vowel /ɑ/ (see Figure 4) shows fairly consistent performance on formants 1 and 2 across all speakers; however, formant 3 relationships are widely varied and do not follow the pattern seen in lower

formants. The third formant of the esophageal speaker is extremely low in relation to the Asai and artificial larynx speakers, and since the normal speaker's formant 3 is not identifiable no relationship to normal can be noted.

The vowel /u/ shows inconsistent performance across all speakers on fundamental and all formant frequencies (see Figure 5). As with /ɑ/, formant 3 is particularly varied across alaryngeal speakers, and only the artificial larynx female shows a third formant within range of normal.

Vowel /i/ shows quite close relationships to normal at all levels (see Figure 6). Fundamental frequency values follow a pattern established on other vowels with Asai most closely approximating the normal value. Formant 1 values across speakers are all near normal values, while formant 2 values, all somewhat higher than the normal, are more inconsistent with normal. The exception on formant 2 is the artificial larynx speaker, whose performance on all frequencies was seen as very consistent with normal. For formant 3 no values could be determined for the esophageal speaker, and the Asai and artificial larynx speakers show results similar to those seen in the formant 2 relationships to the normal performance.

It can be seen by referring to Figures 3 through 6 that female speakers showed more consistent performances and most closely approximated normal on vowels /æ/ and /i/. This differs from results found for male speakers where /æ/ and

/ɑ/ seem to most closely approximate normal. The notable exception in the male performances was the esophageal speaker, whose values were very inconsistent with normal. Among female speakers also, the esophageal performance was least predictable when compared to normal values.

DISCUSSION

The results of this investigation reveal a number of areas in which inconsistency and unpredictable variability predominate. There are several possible explanations for these findings; these explanations, in addition to the small number of subjects, may provide clues to areas of question for further study.

It is possible that the vowels produced by each subject may not have been those vowels modelled in the explanation given to the subject before recording. Individual differences in vowel production were not compensated for by averaging a larger group of subjects, as was the case in the Peterson and Barney study (1952), and thus an individual difference would alter the results of the study.

Another factor which must be considered is acoustic noise produced in the vocal tracts of the alaryngeal speakers, in the form of breathiness, aperiodicity of vibration at the level of sound production, and possible stoma noise. This noise factor makes identification of formants difficult in sonographic analysis. Here again, a larger sample population

may have been effective in overcoming some of these individual problems through averaging values obtained from a number of subjects.

When the normal subjects for this investigation were compared with the Peterson and Barney (1952) data, some notable differences were found. (See Tables VII and VIII.) Differences are also very prominent when comparisons are drawn between previously reported studies of esophageal fundamental and formant frequency values, and those found in this

TABLE VII

COMPARISON OF NORMAL SUBJECTS WITH
PETERSON AND BARNEY SUBJECTS
MALE SPEAKERS

VOWEL	F_0	F_1	F_2	F_3
/æ/				
P&B Norms	127	660	1720	2410
Normal	128	741	1850	2733
Difference	+1	+81	+130	+323
/a/				
P&B Norms	124	730	1090	2440
Normal	120	716	1300	2600
Difference	-4	-14	+210	+160
/u/				
P&B Norms	141	300	870	2240
Normal	130	300	1300	2250
Difference	-11	0	+430	+10
/i/				
P&B Norms	136	270	2290	3010
Normal	138	175	2358	3100
Difference	+2	-95	+68	+90

TABLE VIII

COMPARISON OF NORMAL SUBJECTS WITH
PETERSON AND BARNEY SUBJECTS
FEMALE SPEAKERS

VOWEL	F_0	F_1	F_2	F_3
/æ/				
P&B Norms	210	860	2050	2850
Normal	221	350	2700	3216
Difference	+11	−510	+650	+366
/ɑ/				
P&B Norms	212	850	1220	2810
Normal	192	933	1416	*
Difference	−20	+83	+196	
/u/				
P&B Norms	231	370	950	2670
Normal	208	383	1466	2783
Difference	−23	+13	+516	+113
/i/				
P&B Norms	235	310	2790	3310
Normal	221	350	2700	3216
Difference	−14	+40	−90	−94

*denotes no information available

investigation. Results of several studies have shown esopha-
geal fundamental frequency values to shift downward from
normal. This trend also was reflected in the performance of
both male and female esophageal speakers in this investiga-
tion. Formant frequency values for the vowel /u/ for the
male esophageal speaker are most similar to those found in
the studies of esophageal speakers by Sisty and Weinberg
(1972). Formant values for the other three vowels were not
as similar, with /ɑ/ being the least comparative. Female
esophageal values in the Sisty and Weinberg (1972) study

show some similarities to the female esophageal speaker in this investigation of F_1 and F_2 of /i/ and /u/. The vowels /æ/ and /a/, however, are not comparable.

The Asai voice has not been as widely studied; however, the 97.08 Hz mean fundamental frequency for Asai males reported by Curry and Snidecor (1973) differs considerably from the 163 Hz mean fundamental frequency of the male Asai speaker in this investigation. The two female Asai speakers seen by Curry and Snidecor (1973) showed such disparity that no average was determined. One Curry and Snidecor (1973) female showed a mean fundamental frequency of 150 Hz. The female Asai speaker in this investigation showed a mean fundamental frequency of 180 Hz for the vowels studied. This suggests that a higher fundamental frequency is available to the female Asai speaker, thus a closer approximation to normal female fundamental frequency.

Phonetic context of the vowels in this investigation may have resulted in variance from previous studies. Peterson and Barney (1952) imbedded the vowels they studied between consonants /h/ and /d/, to form monosyllable words, e.g., /hid/, /hud/, and /hæd/. The /h/ phoneme is not available to laryngectomees as production occurs at the level of the vocal folds. The phoneme /d/ would have created a glide characteristic to the sonographic display due to its voiced production. This would have made sonographic analysis of the target vowels more complicated. Unvoiced consonants

isolate the vowel serving to make it a distinct entity, thus facilitating sonographic analysis. The subjects in the current study produced the four vowels studied within the context of consonants /t/ and /p/. The initiating phoneme /t/ was chosen because it is a plosive and this feature aids the laryngectomee in initiating phonation. The arresting consonant /p/ was chosen primarily for its voiceless component.

The phonetic contextual difference described may have altered the results as compared with Peterson and Barney (1952). This difference, however, would not have an effect upon comparisons of the alaryngeal speakers with normals in the current study.

It is apparent, one must use caution in drawing conclusions from comparisons of single members of populations of alaryngeal speakers with single members of the normal speaking population. Nonetheless, from the results obtained in this investigation, it may be suggested that quantitative data are useful sources of information for consideration when selecting the type of alaryngeal voice to be used by a prospective laryngectomee. Information on fundamental frequency relationships to the normal fundamental frequency of male and female speakers is of importance.

If data on the female alaryngeal speakers are representative of a larger population, the Asai speaker most closely approximated the normal speaker in fundamental frequency. In view of the frequent confusion in the perception

of male and female voices in esophageal speakers (Curry and Snidecor, 1961), this higher fundamental frequency could be influential in selecting the Asai procedure with the purpose of obtaining a fundamental frequency more similar to normal female fundamental frequencies.

CHAPTER V

SUMMARY AND IMPLICATIONS

SUMMARY

Laryngectomy, as a treatment for malignant laryngeal lesions, requires the patient to seek a substitute method of producing speech. Three types of alaryngeal speech were described: esophageal, Asai, and artificial larynx. One consideration in deciding which mode of speech is best for the patient is how closely each type of alaryngeal speech approximates normal.

This investigation was an objective examination of esophageal, Asai and artificial larynx speech as compared with normal in males and females. Comparisons were based on production of four vowels /i/, /æ/, /ɑ/, /u/. Eight subjects were selected, one male and one female in each type of alaryngeal speech. All speakers were recorded producing the four vowels three times. Each production was analyzed sonographically to determine fundamental frequency and formants 1, 2, and 3.

The following information was obtained to answer the questions posed:

1. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by esophageal

speakers and normal speakers?

Fundamental frequencies produced by male and female esophageal speakers were 60 and 70 Hz lower than their normal counterparts, respectively. Formant frequency values for both esophageal speakers showed little relationship to normal, varying both higher and lower in no predictable pattern.

2. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by Asai speakers and normal speakers?

The male Asai speaker produced a fundamental frequency approximately 30 Hz higher than the normal male speaker. Formant 1 was found to be lower than normal; however, this trend did not continue in formants 2 and 3.

Female Asai fundamental frequency values, only 20 Hz lower than normal, suggest that the Asai technique may provide the female laryngectomee with the closest approximation to normal.

3. What is the difference in fundamental frequency and formants 1, 2, and 3 of four vowels produced by speakers using electronic vibratory devices and normal speakers?

Laryngectomees using an artificial larynx select the fundamental frequency they desire in operating the mechanical device. These chosen values are considerably lower than normal. As with the other alaryngeal speakers, formant values were scattered above and below normal formant frequencies.

4. What is the difference between male and female

production of these vowels in each of the four modes of speech?

Male alaryngeal speakers showed much greater consistency in the production of fundamental frequency over the range of the four vowels. Female alaryngeal speakers varied as much as 53 Hz from one vowel to another. Artificial larynx speakers are not considered because of the mechanical production of fundamental frequency.

5. What is the difference between esophageal, Asai and electronic vibratory device speakers for the four vowels examined?

Esophageal speakers produce a low fundamental frequency when compared to Asai speakers. Artificial larynx also produces a low fundamental frequency, but this is at the discretion of the user. Data on formant frequencies were inconsistent and provided minimal information.

Conclusions derived from this investigation must be guarded due to the small sample population. If the data on these subjects is representative of larger populations, it may be inferred that the fundamental frequency of female Asai speakers is considerably closer to normal than the other two types of alaryngeal speech studied. Male alaryngeal speakers appear to have a broader choice of fundamental frequency ranging above and below normal values.

Further study is suggested on the quantitative aspects of alaryngeal speech using a considerably larger population

of subjects. Extreme variations noted in some instances in formant relationships only clarify the need to minimize individual differences through a larger sample.

IMPLICATIONS

Clinical

The lack of conclusive evidence found in this investigation points strongly toward a need for further research before clinical implications may be drawn. Certain other aspects of the various alaryngeal speech methods should be considered by the physician and speech pathologist before the final determination is made as to what form of speech the new laryngectomee will adopt. Considerations such as length of time to learn the new process, motivation for learning a new mode of speech, attitude of the patient and others important in the patient's environment, occupational needs, and financial needs must be part of the decision-making process.

All forms of alaryngeal speech sound different than the so-called normal voice. While Asai speech is somewhat breathier, it does provide the laryngectomee with immediate voice use following the surgical procedures with no learning process before communication can commence. A major consideration for the prospective Asai patient is the need for three surgical procedures before voice can be used. The length of time between operations varies, but an average of three

months from laryngectomee to speech is common. Additionally,
there is the disadvantage of requiring that a finger be
placed over the stoma in order to direct air through the
Asai "tube"; this factor should be considered as it may be
troublesome to some patients.

Esophageal speech, long the preferred method of alaryn-
geal speech by physicians and speech pathologists, is not
always successfully learned and must be placed in the proper
perspective as a learned skill. It often requires several
years to achieve. There is less surgical intervention neces-
sary in esophageal speech, and this may be an advantage for
some patients.

Artificial larynx speech using a handheld electronic
vibratory device is only recently becoming an acceptable sub-
stitute for voice. While some learning must take place to
produce acceptable speech easily, there is no need to learn
air injection as with esophageal speech. This device is used
by some patients as a means of communication until esophageal
speech can be mastered, and by some, as a permanent means of
oral communication.

All factors mentioned must be considered by patient,
physician, and speech pathologist in order to make the wisest
choice for the new laryngectomee. In addition, general physi-
cal and mental health and the severity of the malignancy and
resultant surgical intervention are factors which will affect
selection of speech mode. Before clinical decisions can be

aided by the objective data obtainable by vocal analysis, a much larger and more varied population of laryngectomees of all types will need to be examined.

Research

It is recommended that this investigation be conducted on a larger scale, using alaryngeal speakers with a range of vocal skills from poor to good. In this way, there is less chance the data will be contaminated by individual differences. It would be interesting to note the various formant relationships that would arise in a larger study, and their correlation with normal formant relationships. It is this aspect of vocal analysis that is most intriguing because it offers insight into the actual characteristics of the vowels produced.

It would be illuminating to include, within another quantitative investigation, identification of the vowels examined by subjecting them to listener judgment. The possibility of alteration of results due to differences between those vowels requested and those actually produced by alaryngeal speakers could possibly be lessened by having panels of listeners identify vowels they perceive prior to subjecting the vowels to sonographic analysis.

The obvious need for more research in all areas of alaryngeal speech, especially formant construction is clear. More and more laryngectomees are surviving surgery each year and requiring professional assistance with decisions regarding

a new voice. While objective data of the nature suggested cannot be the only criteria for decision-making, it can be extremely useful if provided by a representative sample of significant size.

BIBLIOGRAPHY

ASAI, R., Laryngoplasty After Total Laryngectomy. _Archives of Otolaryngology_, 95, 114-119 (1972).

BOGERT, Bruce and Peterson, G. E., The Acoustics of Speech. _Handbook of Speech Pathology_, L. E. Travis, ed., New York: Appleton-Century-Crofts (1957).

BOONE, D., _The Voice and Voice Therapy_. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. (1971).

BROAD, D. J. and Peterson, G. E., The Acoustics of Speech. _Handbook of Speech Pathology_, L. E. Travis, ed., New York: Appleton-Century-Crofts (1971).

COLEMAN, R. O., Male and Female Voice Quality and Its Relationship to Vowel Formant Frequencies. _Journal of Speech and Hearing Research_, 14, 565-577 (1971).

CURRY, E. T. and Snidecor, J. C., Physical Measurement and Pitch Perception in Esophageal Speech. _Laryngoscope_, 71, 415-424 (1961).

CURRY, E. T., Snidecor, J. C. and Isshiki, N., Fundamental Frequency Characteristics of Japanese Asai Speakers. _Laryngoscope_, 83, 1759-1763 (1973).

DENES, P. B. and Pinson, E. N., _The Speech Chain_. Bell Telephone Laboratories (1963).

DIEDRICH, W. M. and Youngstrom, K. A., _Alaryngeal Speech_. Springfield, Illinois: Charles C. Thomas, Publisher (1966).

FAIRBANKS, G., _Voice and Articulation Drillbook_, second edition. New York: Harper and Row, Publishers (1960).

FROESCHELS, E., _Twentieth Century Speech and Voice Correction_. New York: Philosophical Library, Inc. (1948).

GRAY, G. W. and Wise, C. M., _The Bases of Speech_. New York: Harper and Brothers, Publishers (1959).

HANLEY, T. D. and Thurman, W. L., _Developing Vocal Skills_. New York: Holt, Rinehart and Winston (1962).

HOLLIEN, H., Vocal Pitch Variation Related to Changes in Vocal Fold Length. Journal of Speech and Hearing Research, 3 (2), 150-156 (1960).

HOLLIEN, H., Vocal Fold Thickness and Fundamental Frequency of Phonation. Journal of Speech and Hearing Research, 5 (3), 237-243 (1962).

HOLLIEN, H. and Curtis, J., A Laminagraphic Study of Vocal Pitch. Journal of Speech and Hearing Research, 3 (4), 361-371 (1960).

KAPLAN, H. M., Anatomy and Physiology of Speech. New York: McGraw-Hill Book Company (1971).

KARLAN, M. S., Two Stage Asai Laryngectomy Utilizing a Modified Tucker Valve. American Journal of Surgery, 116, 597-599 (1968).

MARSHALL, R. C., Conversion from Asai to Esophageal Speech. Journal of Speech and Hearing Disorders, 37, 262-266 (1972).

MILLER, A. H., First Experiences with the Asai Technique for Vocal Rehabilitation After Total Laryngectomy. Annals of Otology, Rhinology and Laryngology, 76, 829-833 (1967).

MILLER, A. H., Four Years Experience with the Asai Technique for Vocal Rehabilitation for the Laryngectomized Patient. Journal of Laryngology and Otology, 85, 567-576 (1971).

MYSAK, E. D., Pitch and Duration Characteristics of Older Males. Journal of Speech and Hearing Research, 2 (1), 46-54 (1959).

PERKINS, W. H., Speech Pathology, An Applied Behavioral Science. St. Louis: The C. V. Mosby Co. (1971).

PETERSON, G. E. and Barney, H. L., Control Methods Used in a Study of the Vowels. Journal of the Acoustical Society of America, 24, 175-184 (1952).

POTTER, R. K., and Peterson, G. E., The Representation of Vowels and Their Movements. Journal of the Acoustical Society of America, 20 (4), 528-535 (1948).

POTTER, R. K. and Steinberg, J. C., Toward the Specification of Speech. Journal of the Acoustical Society of America, 22, 807-820 (1950).

SCHULTHESS, G. v., Pseudoepiglottis and Pseudoglottis from Pedical Skin Flap. _Acta Otolaryngologica_, 72, 225-228 (1971).

SHIPP, T., Frequency, Duration, and Perceptual Measures in Relation to Judgments of Alaryngeal Speech Acceptability. _Journal of Speech and Hearing Research_, 19, 417-422 (1967).

SISTY, H. L. and Weinberg, B., Formant Frequency Characteristics of Esophageal Speech. _Journal of Speech and Hearing Research_, 15, 439-448 (1972).

SNIDECOR, J. C., and Curry, E. T., How Effectively Can the Laryngectomee Expect to Speak? _Laryngoscope_, 70, 62-67 (1960).

SNIDECOR, J. C., _Speech Rehabilitation of the Laryngectomized_. Springfield, Illinois: Charles C. Thomas, Publisher (1968).

SNIDECOR, J. C., Speech Without a Larynx. _Handbook of Speech Pathology_, L. E. Travis, ed., New York: Appleton-Century-Crofts (1971).

WEINBERG, B. and Bennett, S., A Study of Talker Sex Recognition of Esophageal Voices. _Journal of Speech and Hearing Research_, 14 (2), 391-395 (1971).

WEINBERG, B. and Bennett, S., Selected Acoustic Characteristics of Esophageal Speech Produced by Female Laryngectomees. _Journal of Speech and Hearing Research_, 15, 211-216 (1972).

WEINBERG, B. and Bennett, S., A Comparison of the Fundamental Frequency Characteristics of Esophageal Speech Measured on a Wave-By-Wave and Averaging Basis. _Journal of Speech and Hearing Research_, 15, 351-355 (1972).

Zemlin, W. R., _Speech and Hearing Science, Anatomy and Physiology_. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. (1968).

APPENDIX A

EXAMPLES OF SONOGRAPHIC REPRESENTATION OF NORMAL AND ALARYNGEAL SPEAKERS PRODUCTION OF VOWEL /æ/
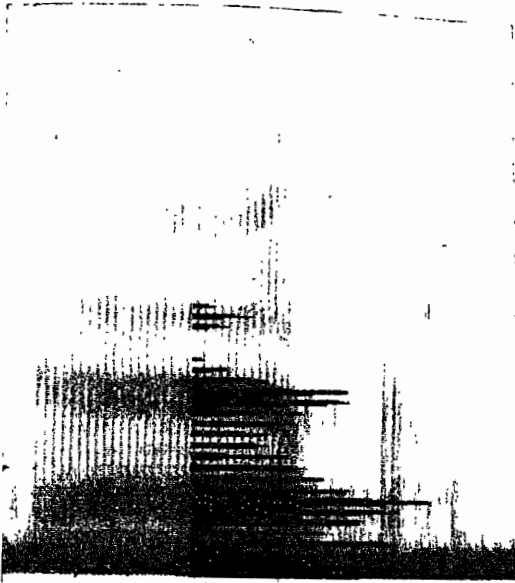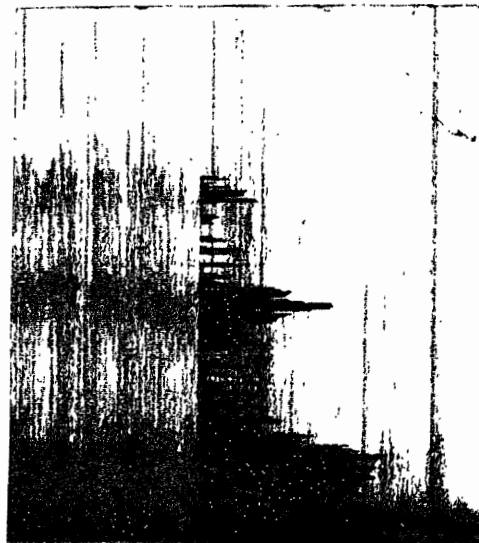

Normal Male Subject /æ/
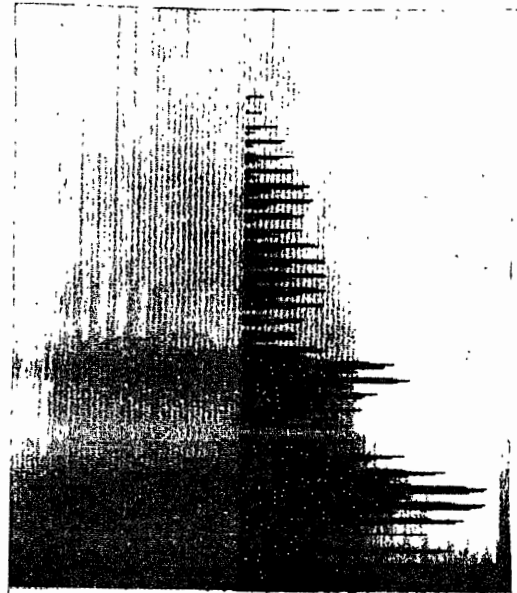
Normal Female Subject /æ/
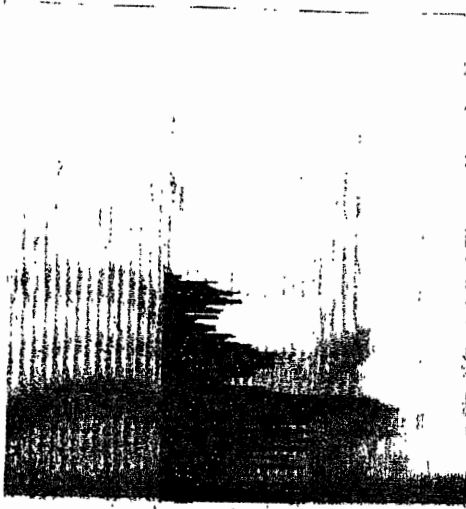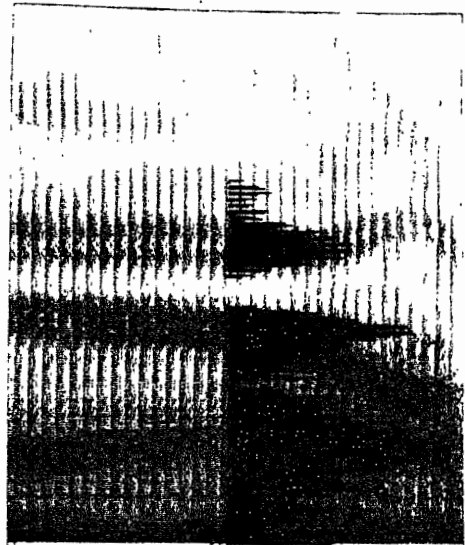
Esophageal Male Subject /æ/

Esophageal Female Subject /æ/

Asai Male Speaker /æ/

Asai Female Speaker /æ/

Artificial Larynx Male
Speaker /æ/

Artificial Larynx Female
Speaker /æ/

APPENDIX B

1. The following are the specifications for Kay Sonograph
    Model 6061B, used in this investigation as described on
    page 22.

SPECIFICATIONS

Frequency range:	85 to 16,000 Hz in two ranges
Displays available:	Frequency-vs-amplitude-vs-time (conventional)
	Frequency-vs-amplitude-vs-time (contour)
	Amplitude-vs-frequency
	Amplitude-vs-time
Analysis time:	1.3 minutes
Effective resolution:	80-8000 Hz 45 and 300 Hz
	160-16000 Hz 90 and 600 Hz
AGC range:	Variable 20 to 40 dB down to 10
Frequency calibration:	Switchable at 50,500, or 1000 Hz intervals
Response:	± 2 dB over entire range
Recording time:	80-8000 Hz 2.4 sec.
	160-16000 Hz 1.2 sec.
Amplifier characteristics:	Flat or 13 dB high-frequency pre-emphasis
Input impedance:	200,600, or 10,000 Ω, Switchable
Recording medium:	Nickle-cobalt plated turntable
Microphone supplied:	Altec-Lansing 681A dynamic

Power supply: 117V, 50/60 CPS, 100 W., Regulated

Dimensions: 25" h x 20" w x 18½" d

Weight: 125 lbs. approx.

2.   The following are specifications for the Uher tape recorder, Model 4000 Report-L, as described on page 21.

Tape Speed: 3 3/4 inches per second

Frequency Response: 40 - 16,000 Hz at 3 3/4 ips

Dynamic Range: 56 dB

Wow and Flutter: ± 0.15% at 7½ ips