Acoustic Analysis of Vowels in Alaryngeal Speech

G. V. M. Hari Prasad
B. Rajashekhar
Assistant Professors
Speech Pathology and Audiology,
K.M.C., Manipal
N.P. Natraja
Professor & Head, Dept. of Speech Sciences,
A.I.I.S.H., Mysore

The success of rehabilitating an alaryngeal speaker mostly depends on the efficacy with which the individual is able to use his voice and convey messages intelligibly. Evaluating the factors affecting the intelligibility of the alaryngeal speakers is important for the rehabilitation of the laryngectomees.

Changes in the speech production mechanism secondary to laryngectomy are reflected in the acoustic characteristics of alaryngeal speech in many ways, (Robbins, Fisher, Blom and Singer, 1984; "Sisty and Weinberg 1972, Weinberg, 1986). Both tracheo esophageal and esophageal speech are characterised by altered fundamental frequency, speaking rate, duration and intensity characteristics. Of these two modes of alaryngeal speech, tracheo esophageal mode is more acceptable, intelligible and nearer to normal laryngeal speakers (Robbins, Fisher, Blom and Singer 1984; Rajashekhar, Nataraja, Rajan, Hazarika, Murthy and Venkatesh, 1990; Rajashekhar, 1991). These altered characteristics which highlight some of the differences between normal and alaryngeal speech, serve to identify parameters of speech important for clinical evaluation and management.

The literature on esophageal speech presents a different picture in terms of the effects of laryngectomy on vocal tract transmission characteristics. Damste (1958) considers that the rest of the vocal tract (the pharyngeal and oral cavities) behaves substantially the same in both normal and esophageal speakers and hence no changes happen in the phonetic events occurring in this region. In contrast, the studies of Rollings (1962) on English speaking laryngectomees and Kytta (1964) on Finnish Speaking laryngectomees revealed that vowel formant frequencies of the esophageal speakers were generally higher than those seen in normal speakers. Sisty and Weinberg (1972) have also observed that the average vowel formant frequency values associated with esophageal speech were elevated and interpreted this to support the view that laryngectomees exhibited a reduced vocal tract length.

No reports on the formant frequency characteristics of T.E. speakers were available to the investigator. Since the T.E. speakers also have an altered vocal tract due to surgical excision of the larynx similar to the esophageai speakers elevated vowel formant frequencies are expected.

The information about the formant frequencies for vowels in the alaryngeal speakers is valuable in understanding the physiology of esophageal and tracheo esophageal speech production and documenting changes in vocal tract function of the alaryngeal speakers.

Methodology

Subjects: Three groups of male speakers namely, tracheo esophageal with Blom Singer's prosthesis, esophageal and normals matched in terms of age,

sex and number participated in the study. All of them were screened for hearing, motor and other sensory abilities and found to be normal in these aspects.

Five subjects who had a tracheo-esophageal puncture (TEP) as a secondary procedure, having undergone laryngectomee earlier and using Blom-Singer's voice prosthesis were selected or the study. The mean age of this group was 57.4 years with a range of 50-69 years. Alaryngeal speakers who used esophageal mode of speech formed the second group. The mean age of this group was 53 years with a range of 37-67 years. Five normal laryngeal speakers matched for age and language with the alaryngeal speakers matched for age and language with the alaryngeal speakers participated in the study. This group had no speech, voice or hearing impairments. The mean age of the group was 50 years ranging from 38-67 years.

Material

A set of six words were segmented from a standard Kannada passage read by the subjects. These words consisted of vowels /a/. */il,* /u/, */ol,* and */e/,* which were free from any contextual influences like following or preceding semi-vowels, glides and nasal sounds.

Method

All the subjects after adequate familiarisation read the passage at comfortable loudness levels and rate into a microphone placed at a distance of 15 cms from the mouth. The reading samples of all the subjects were recorded on hi-bias metal cassettes using professional stereo cassette deck (Akai-CS-M4).

The speech samples of each subjects were digitised at the rate of 8khz using 12 bit VSS data input and output card by feeding the signal from tape deck to the speech interface units through line feedings. The digitised samples were used for analysis of formant frequencies.

Formant frequencies (F1, F2, F3) for each vowel /a/ */\l,* /u/, */ol,* and */e/* were measured directly from the spectrogram display with sectioning on the screen of the computer. Formant frequency estimates were made by measuring the mid point of the visible dark bands of energy appropriate to the first three vowel resonances.

The mean and standard deviation was computed for all the three groups and "t" test was applied to find out the significance of difference between the groups.

Results and Discussion

Tables I, II, and III shows the mean and standard deviation of formant frequencies F1, F2 & F3 for various vowels.

An overview of the tables reveal that the alaryngeal speakers tend to exhibit higher values of vowel formant frequencies F1, F2 & F3 than the normal speakers. Of the two groups of alaryngeal speakers higher formant frequencies were observed in the esophageal speakers than the TE speakers.

Vowel	Oesophageal		TEP		Normal	
	Mean	Range	Mean	Range	Mean	Range
/a/	880 (98)	784-1016	639 (190)	370-832	648 (9)	636-658
/u/	443 (75)	382-533	357 (61)	249-392	405 (29)	381-455
/ol	657 (110)	533-800	610 (106)	492-764	468 (63)	392-517
/il	419 (127)	251-531	363 (60)	257-398	322 (63)	263-398
/e/	628 (107)	520-784	497 (170)	266-658	494 (111)	376-652

Table - 1 : The Mean, S.D. (in parenthesis) and Range of F1 (Hz) of /a/. *lil, lul, lol,* and */e/* in Oeso, TEP and Normal Speakers.

Vowel	Oesophageal		TEP		Normal	
	Mean	Range	Mean	Range	Mean	Range
/a/	1780 (212)	1552-2074	1580 (163)	1302-1689	1520 (170)	1288-1678
/u/	1350 (331)	916-1694	1210 (230)	909-1537	1318 (303)	909-1670
/ol	1288 (103)	1160-1411	1192 (188)	916-1422	952 (69)	891-1035
/i/	2110 (648)	1284-2337	1897 (408)	1559-2588	2124 (305)	1803-2459
/e/	2075 (258)	1819-2467	2000 (253)	1682-2384	1843 {167)	1687-2070

Table - II : The Mean, S.D. (in parenthesis) and Range of F2 (Hz) of /a/, /i/, /u/, *lol,* and */e/* in Oeso, TEP and Normal Speakers.

Vowel	Oesophageal		TEP		Normal	
	Mean	Range	Mean	Range	Mean	Range
/a/	3146 (223)	2856-3470	3074 (345)	2713-3617	2612 (235)	2321-2980
/u/	2777 (648)	1823-3472	2603 (365)	2305-3214	2466 (100)	2313-2586
/o/	2696 (598)	2048-3294	3021 (214)	2833-3347	2671 (207)	2556-2980
/i/	3168 (540)	2586-3752	3019 (394)	2586-3485	2660 (288)	2339-2911
/e/	3094 (484)	2698-3466	3026 (343)	2467-3356	2706 (180)	2472-2933

Table - III: The Mean, S.D. (in parenthesis) and Range of F3 (Hz) of /a/. /i/, /ul, lol, and /e/ in Oeso, TEP and Normal Speakers.

However there was significant difference between the alaryngeal speakers and normals in the formant frequencies F1, F2 and F3 of vowel *lol* only.

The data also reveals a high value of variability among the alaryngeal speakers with the esophageal speaker demonstrating higher variability than the TE speakers. The formant frequencies (F1, F2 F3) of vowels /a/. *Iil, I* u/, *lol,* and /e/being higher in the alaryngeal speakers than the normals speakers is in agreement with the results reported by Sisty and Weinberg (1972); Rajashekhar (1991); Rollins, (1967); Kytta (1964), laryngectomee speakers exhibited higher values of formant frequencies than normal speakers. The formant frequency (i.e., F3-F1-F3-F2, F2-F1) relationship for various vowels among the alaryngeal speakers was different from that of normal speakers. This feature may be one of the factors leading to reduced intelligibility of speech in alaryngeal speaker. Studies by Mitzell et.al. (1985); Gates et.al. (1982); Blom et.al. (1986); Raja Shekhar et. al. (1990) and Rajashekhar (1991) have reported reduced intelligibility of speech in esophageal speakers.

Thus total laryngectomy results in major changes in articulatory aerodynamics and produces alteration in vocal tract morphology (Weinberg 1986). The changes in formant frequencies seem to support the statement and are in agreement with Sisty and Weinberg (1972) that removal of larynx does alter the vocal tract transmission characteristics. The changes in the formant frequency relationship also reflect the altered vocal tract transmission characteristics as well as the altered articulatory patterns adopted by the alaryngeal speakers in the production of speech.

References

1. Blom, E.D., Singer, M.I., Hamaker, R.C. (1986) A prospective study of tracheo esophageal speech, Arch. Oto Laryngol. Head Neck Surg. 112, 440-447.

2. Damste, P.H. (1958) Oescphageal speech after laryngectomy, Groningen, Netherlands; Gebr. Hoitsema.

3. Gates, G. A., Ryan, W., Cooper, J.C., et.al., (1982) Current status of laryngectomee rehabilitation. I. Results of therapy. Am. J.Otolaryngol, 3, 1-7.

4. Kytta, J., (1964) Finnish oesophageal speech after laryngectomy. Sound spectrographic and lineradiographic studies. Acta Otolaryng (Stockholm) Suppl. 195, 7-91.

5. Mitzell, S., Andrews, M.L., & Bowman, S.A., (1985) Acceptability and intelligibility of tracheo esophageal speech. Arch. Otolaryngol, Head Neck Surg. 111, 213-215.

6. Rajashekhar (1991) Acoustic analysis of alaryngeal speech (T.E.P. with B.S. prosthesis and esophageal modes) unpublished thesis submitted to the University of Mysore.

7. Rajashekhar, B., Nataraja, N.P., Rajan, R., Hazarik, P., Murthy, P.S., and Venkatesh, C.s. (1990) Comparison of esophageal and T.E.P. mode of alaryngeal speech in a single laryngectomee. The J. of Indian Speech and Hearing Association 7, 43-46.

8. Rtbbins, J., Fisher, H., Blom, E., and Singer M.I. (98) A comparative acoustic study of normal esophageal and tracheo esophageal speech production. J.S.H.D. 49, 202-210.

9. Rollins, W.J. (1962) A comparative study of vowel formants of esophageal and normal speaking adults, doctoral dissertation, Wayne State University.

10. Sisty, N., & Weinberg, B (1972). Formant frequency characteristics of oesophageal speech J.S.H.R., 15,439-448.

11. Weinberg, B. (1986) Acoustical properties of esophageal and tracheo esophageal speech. In R.Keith and T,Dailey (Eds) Laryngectomee rehabilitation (22nd Ed.) San Diego: College Hill Press.