IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 16, Issue 5 Ver. V (May. 2017), PP 101-111
www.iosrjournals.org

Formant Structure of Vowels Produced By Hindi Esophageal Speakers: A Comparative Study

Kumar Manish, [1] Uday Singh, [2]
[1]*asantara Hearing Care Centre, Manpur, Gaya, Bihar,*
[2] *C. U. Shah Institute Of Audiology & Speech Language Pathology, C. U. Shah Medical College Campus, Surendranagar, Gujarat.*

Abstract:
Aim: *To investigate the characteristics of vocal tract resonance in Hindi Esophageal Speakers.*
Methodology: *Five normal Hindi speakers and five Hindi esophageal speakers participated in this study. They were asked to produce the three corner vowels /a/, /i/, and /u/ in the three different conditions like non-sustained, sustained and word level. The average first three formant frequencies F1, F2, and F3 from these vowels were obtained and submitted for statistical comparisons. For each formant, the formant frequencies measured from the medial 80% of the vowel partition were averaged and data obtained from esophageal speakers were compared with that obtained from laryngeal speakers.*
Result & Discussion: *The results indicated that esophageal speakers were associated with significantly higher formant frequencies (F1,F2 and F3) when compared with laryngeal speakers .This finding suggests that formant frequencies changes across vowels were systematic and similar for esophageal and laryngeal speakers are consistent with those previously reported in English. It can be concluded that vocal tract transmission has been changed in esophageal speakers of Hindi after laryngectomy. It also suggests that pharyngoesophageal (PE), which is used as neoglottis by esophageal speakers, is located higher than vocal folds in normal speakers.*
Summary and Conclusion: *Thus, by comparing this result of Hindi speakers with the previous studies, it is noted that the difference in formant characteristics of vowels between esophageal speech and normal speech to be similar across languages. Therefore, the factor of language difference seems to have little effect to the change in vocal tract transmission after laryngectomy.*
Keywords: *Esophageal Speaker, Formant Structure, Vowel*.

I. Introduction

Human beings communicate with one another principally through speech i.e., we express our intentions in a sequence of articulated sounds to which particular meaning gets attached. Speech, being a convenient form of communication, loss of this power of communication makes our daily life more complicated, less efficient and more frustrating. The most widely known method of alaryngeal speech is esophageal speech particularly when the client is not eligible for surgical voice restoration method i.e. tracheoesophageal speech. Other options such as an alternative form of communication system are one that permits those using it to exchange information with a high degree of ease, flexibility, speed, and accuracy in a wide variety of circumstances. Spoken language appears to meet these conditions.

Esophageal Speech

Esophageal speech, one of the preferred alaryngeal modes in which laryngectomized persons learns to speak using the upper esophageal sphincter as a vibrating site for alaryngeal phonation. Esophageal phonation is produced by injecting air, which is present in the hypo laryngeal space above the esophagus, into the upper esophagus below its constrictive opening, then air is forced out of the esophagus under pressure. This causes the tissue of the upper esophagus to vibrate to produce alaryngeal phonation. The sound is then resonated and articulated into human speech in the vocal tract by the articulators.

Formant Structure

A formant is specified by its center frequency commonly called formant frequency and bandwidth. Some of the formants always have fixed frequency positions irrespective of the vowel phonated, whereas others vary in frequency depending upon the vowel concerned (Cervera, Miralles, & Alvarez, 2001). The former, or fixed formants, are generated in the cavities which remain almost unchanged in volume, for example the resonance cavities of the **t**horax, larynx and nose. The latter, or variable formants, are generated in the cavities of adjustable volume, such as the pharyngeal, oral and labial cavities (Sovijarvi, 1938). According to Liu et al (2009), formant frequency is determined by the configuration of the vocal tract which can be affected by a numbers of factors: (1) the intrinsic size of the vocal tract, (2) the size of the tongue, (3) the size and

configuration of the oral cavity, (4) the size and configuration of the pharyngeal cavity, and (5) the tongue configuration.

Studies particularly in Hindi esophageal speakers have not been reported in literature till date. Hence, the present study gains an importance by highlighting the acoustic characteristics in Hindi speaking esophageal speakers. Further, the present study provides an insight to the speech clinicians in understanding the probable variations in formant characteristics of Hindi esophageal speakers as against the other esophageal speakers.

Aim of Study:

The present study aims at estimating the acoustic characteristics, particularly formant values of three cardinal vowels, viz, [a], [i] and [u] of Hindi speaking esophageal speakers and compare it with formant values of the same vowels obtained from normal individuals.

II. Method

Subjects: Five male esophageal and five male laryngeal speakers participated in this study. Five adult males ranging in age from 47 to 77 yrs (mean age 62; 2 yrs) who had undergone total laryngectomy were taken as the subjects for this study. At the time of this study, all the subjects were ambulatory, had no structural or neurological deficits, had fairly good speech intelligibility and were able to freely phonate. The subjects had no complaints of impaired hearing.

Inclusion Criteria: 1. All subjects should be native speakers of Hindi language with no history of speech and/or hearing problem except that associated with laryngectomy in the case of esophageal speakers. 2. All the subjects should have undergone total laryngectomy with no further dilation. 3. All alaryngeal speakers will be using esophageal speech as their primary mode of communication

The period of use of alaryngeal speech in the subjects varied from 1; 5 yrs to 12; 5 yrs. The control group consisted of 5 adult males ranging in age from 60 to 64 years (mean age 62; 4 yrs) who were normal speakers and had no history of voice or speech disorder. The demographic information regarding the distribution of 5 esophageal subjects is stated in the following table1.

Table 1: Demographic Information Of Esophageal And Laryngeal Speakers

Subject	Age (yrs)	Period since operation	Type of surgery	Normal subjects Age (Yrs)
1	47	1yr.5 months	Total laryngectomy	62
2	65	10 yrs 9 months	Total laryngectomy	63
3	58	12 yrs 6 months	Total laryngectomy	60
4	64	4 yrs 1 month	Total laryngectomy	62
5	77	1 yr 6 months	Total laryngectomy	63

III. Materials

The three cardinal vowels of the vowel triangle namely (1) the low back vowel /a/, (2) the front high vowel /i/, and (3) the back high vowel /u/ articulated by both esophageal and normal speakers under three conditions were used as materials for the study. These conditions include:-

(a) Three vowels (/a/, /i/, and/ u/) in isolation (non sustained articulation)
(b) Three vowels (/a/, /i/, and/ u/) in isolation (sustained articulation)
(c) Three vowels occurring at the meaningful word level such as /maka/, /mika/, and /muka/.

Recording Procedure:

The subjects, both the esophageal as well normal subjects, were asked to sit comfortably in a noise free room. After sufficient rehearsals, they were asked to produce the target vowels in the first two conditions and read the three words. The subjects were instructed to phonate and read at comfortable loudness on to a dynamic microphone, kept at a distance of 10 cm to 15cm from the mouth. The procedure consisted of recording the production of three vowels in phonated and sustained as well as they occurred in the words using CSL (Kay Elemetrics Corp; model 4150) which were later analyzed using the software PRAAT (version 4.4.33). Each subject produced the three cardinal vowels namely /a/, /i/, and /u/ in all the above mentioned conditions at his

typical modal frequency at a comfortable loudness. Each subject articulated the target speech signal 4 times in a random order. The speech signals were recorded using a dynamic microphone to the system and to minimize the recording of the stoma noise for esophageal speakers, the microphone was placed at an angle of about 30 degree above horizontal. The signals were amplified by using a multichannel conditioning amplifier. The signals were then digitized with a sampling frequency of 20 KHz and a quantization rate of 16 bits/sample.

Acoustic analysis:

The speech sample was analyzed by using the software PRAAT. The 1st three formants values (F1, F2, and F3) were obtained by using LPC (linear predicting coding) algorithm of PRAAT. For all vowels produced by laryngeal and esophageal speakers, only 80 % (mid portion) of the vowels were used for the formant measurement to avoid initiation and termination effects. All the analysis was done using the PRAAT software. Broadband spectrogram (300 Hz) of each vowel in all three conditions produced by each speaker were obtained and used to facilitate the identification of the vowel. Once this segment was identified, LPC spectra were computed for the vowel segment using consecutive samples centered in the middle of the vowel segment. The signal was then multiplied by 45ms hamming window and LPC coefficient were computed using an autocorrelation method (Atal& Hanauer, 1971). The LPC order was 23 according to the criterion of Rabiner & Shaffer (1978). LPC was preferred over spectrogram since this technique offers reliable detection of weak nasal formants and as well formants undergoing rapid transitions in frequency areas where parametric spectral analysis typically performs poorly. Much of the recognition of speech using a feature-based approach relies heavily on the detection of formant time-frequency trajectory patterns, which gives the identification not only for the voiced speech sound currently under analysis, but also can provide important cues to pre- and postvocalic speech. The enhanced formant detection properties offered by pole focusing therefore can considerably improve the reliability of formant pattern recognition (Jack, 1988). LPC (or, more correctly, analysis by linear prediction) has become increasingly popular for analyzing formant frequencies, since it gets round typical difficulties of measuring them directly in the sound wave, especially the less well defined formants often found for women and children. A further advantage of using the linear prediction model is that it estimates the resonance modes of the speaker's vocal tract, which is fine since one common reason for studying the formants is the relation between articulation and vocal tract resonance.

Formant frequencies were measured at the midpoint of the vowel. Detection of a formant was based on the visual examination of LPC spectra. The first three peaks from LPC spectrum were selected as F1, F2 & F3 values. All the formants (F1, F2 &F3) were measurable 100% of time in both normal & esophageal group. The yielded list of F1, F2, and F3 values were later averaged and submitted for statistical comparisons. The statistical analysis included T-test & independent sample test for comparing the formant values of normal & esophageal speakers. For the comparison between three conditions among the same group i.e. esophageal or normal group, one way ANOVA and Post hoc test were employed.

IV. Result

The first three formant frequency F1, F2, and F3 of the cardinal vowels /a/, /i/, and /u/ in esophageal speakers and laryngeal speakers were determined from the LPC spectra. The first three peaks from LPC spectra were selected for this purpose. Mean formant values with SD, were determined for the target vowels under three conditions. Independent sample test was carried out to establish the statistical significance in formant values between esophageal & normal speakers. Post Hoc test (Tukey HSD) was employed to compare the formant values amongst the vowels produced under the three conditions in esophageal & normal speakers. The results are discussed below.

Condition-1 (non-sustained production of the vowels)

The mean and the standard deviation of F1, F2, and F3 values of the three vowels produced by esophageal and laryngeal Hindi speakers in isolated condition is given in the table 2.

Table 2: Mean (M), Standard Deviation (SD) of formant values F1, F2, and F3 for vowels /a/, /i/, and /u/ and P value for normal and esophageal speakers. (*P≤ 0.05-significant)

Vowel	Formants	Normal		Esophageal		P
		Mean(M	S.D	Mean(M)	S.D	
/a/	F1	723.44	28.17	835.11	38.73	0.001* 01*
	F2	1211.69	66.65	1315.08	53.24	0.027*
	F3	2694.84	24.83	2870.64	49.85	0.000*
/i/	F1	283.44	24.23	341.81	41.88	0.027*

	F2	2374.09	69.22	2676.43	120.76	0.001*
	F3	3019.75	170.67	3239.48	236.03	0.130
/u/	F1	329.95	17.13	425.68	29.84	0.000*
	F2	737.14	58.23	907.12	133.39	0.031*
	F3	2373.69	89.15	1839.91	192.33	0.000*

It is obvious from the table that all the formant values under considerations are higher in esophageal speakers than in normal speakers except for vowel /a/ where F3 value is lower in esophageal speakers compared to normal speakers. Independent samples test was carried out on the data to determine the significant difference between the normal and esophageal group for the cardinal vowel /a/, /i/, and /u/ in non-sustained condition. The results indicate that for vowel /a/, higher F1, F2, and F3 values in esophageal speakers are statistically significant when compared to the values of normal speakers. In case of the vowel /i/, the values of F1, F2 and F3 are higher in case of esophageal speakers compared to laryngeal speakers. But statistical significance exits only for F1 and F2. Though F3 value is higher in esophageal compared to laryngeal speakers, no statistical significance is observed. For the vowel /u/, F1 and F2 values are higher in esophageal speakers compared to normal speakers and this is statistically significant. But in case of F3, the trend is observed to be reverse, that is, F3 value is statistically higher in the case of normal speakers compared to esophageal speakers. For the purpose of visual comparison, the mean F1, F2, and F3 values across vowels and speakers groups is represented in figure 1.

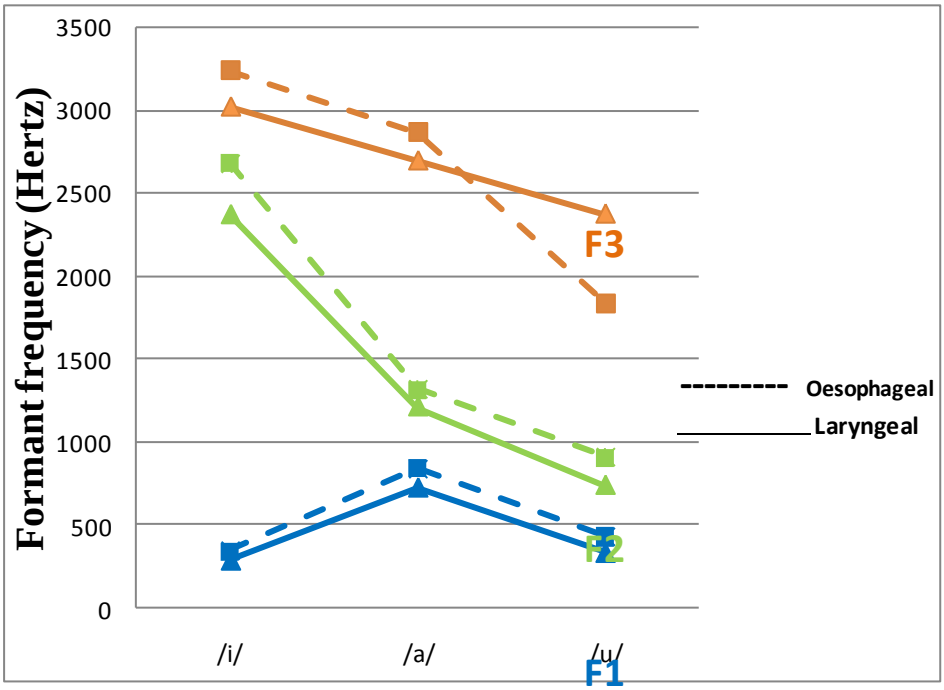

Figure1. Mean F1, F2, and F3 value of the three corner vowels produced by normal and esophageal speakers of Hindi.

From the above graph it can be seen that formant values (F1, F2, & F3) of esophageal speakers are higher compared to laryngeal speakers in all three cardinal vowels /a/, /i/, and /u/ for non-sustained condition except for F3 value of the vowel /u/. In addition, the general trends of F1, F2, and F3 of vowels produced by esophageal speakers closely follow that of laryngeal speakers

Condition-2 (sustained articulation)

The mean and the standard deviation of F1, F2, and F3 values of the three vowels produced by esophageal and normal Hindi speakers in sustained condition is given in the table 2.

Table 3: Mean (M), Standard Deviation (SD) of formant values F1, F2, and F3 for vowels /a/, /i/, and /u/ and P value for normal and esophageal speakers. (*P≤ 0.05-significant)

Vowels	Formants	Normal		Esophageal		P value
		Mean (M)	S.D	Mean(M)	S.D	
/a/	F1	599.57	44.55	782.88	104.35	0.007*
	F2	1052.07	85.61	1321.46	129.42	0.005*
	F3	2232.72	102.47	2820.33	130.76	0.000*
/i/	F1	278.09	35.10	368.22	40.65	0.006*
	F2	2292.54	84.93	2747.73	119.04	0.000*
	F3	3125.33	96.86	3292.03	172.39	0.096
/u/	F1	386.20	42.37	432.10	42.65	0.126
	F2	701.43	52.37	898.70	105.83	0.006*
	F3	2095.75	122.14	1797.95	122.14	0.002*

Independent samples test was carried out on the data to determine the significant increase in the formant values of the cardinal vowels /a/, /i/, and /u/ in sustained production condition. The results show that for vowel /a/, formant values F1, F2, and F3 are significant between esophageal and laryngeal speakers. That is esophageal speakers have higher F1, F2, and F3 values compared to laryngeal speakers. For vowel the /i/, formant values F1 and F2 are significant ie formant values F1 and F2 are higher in esophageal compared to laryngeal speakers. But the higher F3 value is found not to reflect statistical significance, though esophageal speakers have higher F3 value compared to laryngeal speakers. For vowel /u/, formant values F2 and F3 are significant between normal and esophageal speakers but F2 value is higher in esophageal speakers compared to laryngeal speakers and F3 value is higher in laryngeal speakers compared to esophageal speakers. For the purpose of visual comparison, the mean F1, F2, and F3 values across vowels and speakers groups for the vowels produced in sustained condition is represented in figure 2.

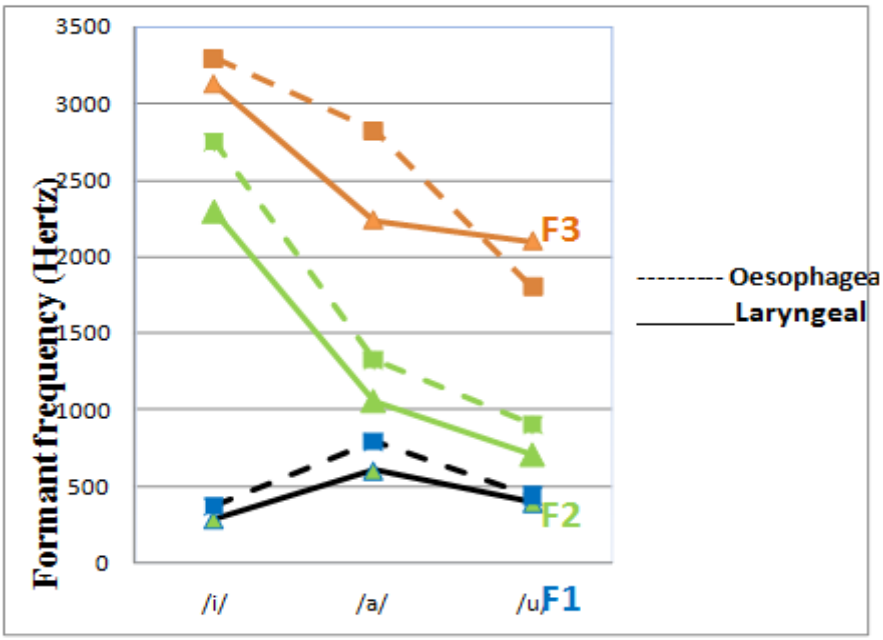

Figure2. Mean F1, F2, and F3 value of the three corner vowels produced by normal and esophageal speakers of Hindi.

From the above graph it can be seen that formant values (F1, F2, & F3) of esophageal speakers are higher compared to laryngeal speakers in all three corner vowels /a/, /i/, and /u/ for sustained condition except for F3 value of the vowel /u/. In addition, the general trends of F1, F2, and F3 for the vowels produced by esophageal speakers closely follow that of laryngeal speakers.

Condition-3 (word level)

The mean and the standard deviation of F1, F2, and F3 values of the three vowels produced by esophageal and laryngeal Hindi speakers in word level is given in the table 4..

Table 4: Mean (M), Standard Deviation (SD) of formant values F1, F2, and F3 for vowels /a/, /i/, and /U/ and P value for normal and esophageal speakers. (*P≤ 0.05-significant)

Vowels	Formanttt	Normal		Esophageal		P
		Mean(M)	S.D	Mean(M)	S.D	
/a/	F1	747.56	78.35	908.04	47.4	0.004
	F2	1366.14	37.55	1527.82	43.6	0.000
	F3	2970.23	353.9	3260.60	88.9	0.113
/i/	F1	339.22	52.38	431.92	62.8	0.035
	F2	2162.25	146.5	2701.33	237.	0.003
	F3	2492.46	333.5	3137.39	164.	0.275
/u/	F1	438.81	47.22	488.91	58.6	0.175
	F2	727.67	80.94	870.64	86.2	0.027
	F3	2368.56	204.5	1922.62	318.	0.030

Independent samples test was carried out on the data to determine the significant difference between the normal and esophageal group of the cardinal vowels /a/, /i/, and /u/ produced at the word level. The results show that for vowel /a/, formant values F1 and F2 are significant between esophageal and laryngeal speakers and F3 is not significant though F3 value is higher in esophageal speakers compared to laryngeal speakers. That is formant values F1 and F2 are higher in esophageal speakers compared to laryngeal speakers. For the vowel /i/, formant values F1 and F2 are significant and F3 is not significant though F3 value is higher in esophageal speakers compared to laryngeal speakers. The formant values F1 and F2 are higher in esophageal speakers compared to laryngeal speakers. For vowel /u/, formant values F2 and F3 are significant that is, F2 value is higher and F3 value is lower in esophageal speakers compared to laryngeal speakers. The formant value F1 is not significant though F1 value is higher in esophageal speakers compared to laryngeal speakers. For the purpose of visual comparison, the mean F1, F2, and F3 values across vowels and speakers group in word level condition is graphically represented in figure 3.

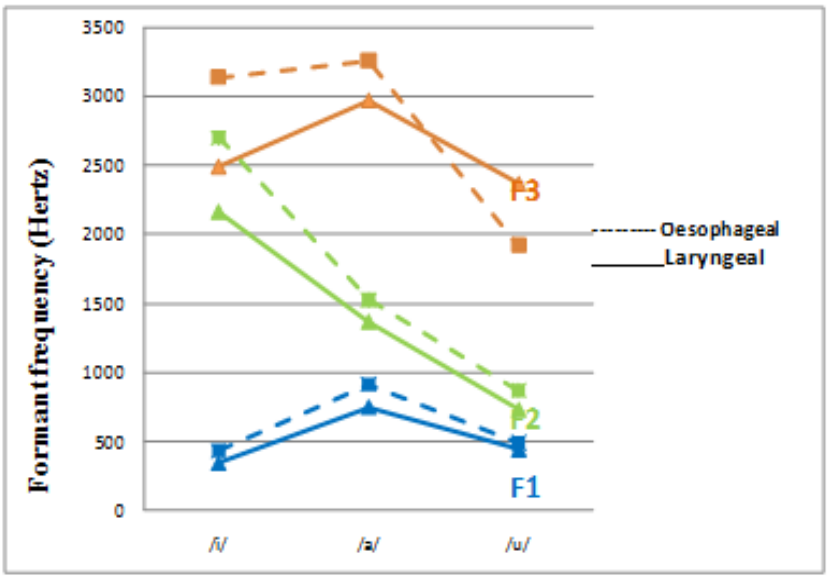

Figure3. Mean F1, F2, and F3 value of the three corner vowels produced by normal and esophageal speakers of Hindi.

.

From the above graph it can be seen that formant values (F1, F2, & F3) of esophageal speakers are higher compared to laryngeal speakers for all the three cardinal vowels /a/, /i/, and /u/ at word level condition except for F3 value of the vowel /u/. In addition, the general trends of F1, F2, and F3 of vowels produced by esophageal speakers closely follow that of laryngeal speakers.

Comparison across vowels

To compare the formant values across the three conditions in case of esophageal speakers as well in normal speakers, Post Hoc test was applied and the result is discussed below.

Table 5: P value of formant values F1, F2, and F3 for vowels /a/, /i/, and /u/ between three conditions (isolated, sustained, and word level) for esophageal group.

Vowels	Formant	Non-sustained Vs sustained	Non-sustained	Word level Vs Sustained
/a/	F1	0.485	0.263	0.37
	F2	0.992	0.005*	0.006*
	F3	0.692	0.000*	0.000*
/i/	F1	0.684	0.034*	0.146
	F2	0.785	0.970	0.902
	F3	0.904	0.690	0.441
/u/	F1	0.973	0.110	0.158
	F2	0.992	0.861	0.915
	F3	0.951	0.825	0.652

By applying Post Hoc test using multiple comparison through Tukey HSD method, it is noted that there is no statistically significant difference between the vowels articulated in non-sustained & sustained conditions. Statistical significance is noted for F2 and F3 values for /a/ both in non-sustained Vs word level production and word level vs. sustained production. For the vowel /i/ significance is noted only for the F1 under non-sustained Vs word level condition. But no significance is noted for the vowel /u/ under any three conditions.

Table 6: P value of formant values F1, F2, and F3 for vowels /a/, /i/, and /u/ between three conditions (isolated, sustained, and word level) for laryngeal group.

Vowels	**Formant**	**Non-sustained Vs**	**Sustained Vs Word level**	**Word level Vs Sustained**
/a/	F1	0.010*	0.768	0.003*
	F2	0.007*	0.008*	0.000*
	F3	0.013*	0.145	0..000*
/i/	F1	0.975	0.101	0.070
	F2	0.464	0.020*	0.167
	F3	0.741	0.850	0.425
/u/	F1	0.088	0.002*	0.113
	F2	0.670	0.971	0.802
	F3	0.028*	0.998	0.031*

By applying Post Hoc test using multiple comparison through Tukey HSD method, it is noted that there is statistically significant difference for F2 & F3 value and no significance for F1 value for vowel /a/ under sustained Vs non-sustained conditions. Statistical significant is also noted for F2 value under non-sustained Vs word level production and F1, F2, & F3 values under sustained Vs word level conditions. For the vowel/i/, significance is noted only for F2 value under non-sustained Vs word level production. For the vowel /u/, significance is noted for F3 value under non-sustained Vs sustained, F1 under non-sustained Vs word level and F3 value under word level Vs sustained condition.

V. Discussion

The present study attempted to examine the vocal tract transmission property of Hindi speaking esophageal speakers by analysing the formant values of three cardinal vowels. The source filter theory of vowel production assumes the independence between the source function and the vocal tract transfer function. (Fant,1960). That is, the formants are independent of fundamental frequency (pitch of the voice). Changes in one system should not alter the other system according to Fant's acoustic theory of speech production. Accordingly, after the removal of the larynx (source function), esophageal speakers should still be able to produce vowels as before laryngectomy if the formant structure (vocal tract transfer function) remain unchanged (Liu et al 2005). But our data indicate that average F1, F2, and F3 values of all the three cardinal vowels produced by esophageal speakers are significantly higher than those produced by normal speakers. This could be due to change in both the source as well the transmission function. In the case of the former the laryngeal source is being replaced with pseudo glottal vibrations and the resonant cavity volume is considerably reduced thereby increasing the formant frequencies. The formant characteristics of vowels produced by Hindi esophageal and normal speakers in the three different conditions are discussed based on T-test, independent sample test, ANOVA and Post Hoc test (Tukey HSD).

The reasons for the effective change of formants:

The significant difference in the formant values of vowels between esophageal and laryngeal speech indicates that the vocal tract transfer function has been changed due laryngectomy. According to the source – filter theory (Fant, 1960), the length of the vocal tract can significantly affect the formant characteristics of vowels. The shorter the length of the vocal tract, the higher is the formant frequencies. The higher formant frequencies in esophageal speech are therefore believed to be due to a shortened effective vocal tract length compared to laryngeal speakers. As Sisty & Weinberg (1972) and Cervera et al (2001) noted, increased average vowel formant frequency in esophageal speech could be attributed to the overall vocal tract length caused by position of neoglottis in esophageal speakers.

Another reason which supports the higher formant frequency in esophageal speaker is made available from the physiological studies. During laryngectomy, laryngeal structure has to be removed and so esophageal speakers have to learn to adopt the PE segment as a new sound source. Edels (1983) noted that the PE segment is formed by the superior pharyngeal constrictor, cricopharyngeus, and the upper part of esophagus. Cineradiographic images obtained from esophageal speakers of Finnish and English revealed that the PE segment of esophageal speakers is located higher than the vocal folds in laryngeal speakers (Kytta, 2001; Diedrich et al, 1966; Damste et al, 1969). The PE segment was generally found at the level of C4-C6, whereas vocal folds are located at the level of C5-C6 in laryngeal speakers (Zemlin, 1998). Bentzen et al (1976) investigated the esophageal phonation using X-ray videotapes. They found that there were high-seated PE segment at the level of the C6-C7. They observed that better esophageal voice quality in some superior esophageal speakers was associated with high-seated PE segment. Accoding to Diedrich and Youngstrom (1966), the PE segment might vibrate at a level as high as in the hypopharynx at C2-C3 upon phonation in some good esophageal speakers. An upward displacement of the PE segment to the level of C3-C5 during phonation was also found in most of the superior tracheoesophageal laryngectomees (Van et al 2001). This altered location of the new sound source results in a shortened effective vocal tract, which explain why formant frequencies associated with vowels produced by esophageal speakers are significantly higher than those produced by laryngeal speakers.

Comparison of formant values across the two groups:

Condition 1- (non-sustained)

In the present study it was noted that the increased formant values in esophageal speakers was significantly higher than normal speakers. That is the difference in formant values of normal & esophageal speakers were highly significant. This holds good not only for the vowels articulated in a non-sustained way but also in other conditions.

The first formant

The formant value F1 for vowel, / a/ was higher by 111.64 Hz in esophageal speakers compared to laryngeal speakers. The result is probably due in part to the unrounded nature of /a/ and to the fact that /a/ is generated anatomically more posteriorly articulated than the other back vowels. In the view of Fant (1960), though F 1 is generally dependent more on the back cavity volume than on the volume of other cavities, the vowel /a/ is an exception. Its F1 is affected equally on a percentage basis by a change in the front cavity volume and by a change of the back cavity volume (Kytta 1964). It is evident that the latter cavity volume plays an exceedingly prominent part in the F 1 of /a/, even though /a/, as distinguished from the other back vowels, would be equally dependent upon the front cavity volume. The mean difference of F1 for /u/ was 95.72 Hz, which was less compared to back vowel /a/.The F 1 value for the front vowel /i/ was 58.37 Hz higher than the mean value of laryngeal speakers. It may be due to small cavity volume and the ridge of the tongue being in closest proximity to the palate. (Kytta 1964)

The second formant:

The average increase of F2 value for vowel /u/ and /a/ in the esophageal speakers as compared to the normal were169 Hz &103.39 Hz. It may because the vowel /u/ presupposes the higher tongue- ridge position and thus smaller distance to the palate compared to vowel /a/.The difference in F2 value for the vowel /i/ was 302.34 Hz higher in esophageal speaker as compared to normal speakers. Fant (1960) stated that F 2 of /**i**/ is clearly a half-wavelength resonance of the back cavity and esophageal speakers have less back cavity as compared to the normal. This causes higher resonance in the case of esophageal speakers and hence higher F2 value in them.

The third formant:

The third formant is of secondary importance in the back vowels since it plays no part in the intelligibility of speech (Kytta 1964). Fant (1960) said that F3 of back vowels is chiefly dependent on the parts in front of tongue constriction. The mean difference of F3 value for /a/ was 175.80 Hz higher in esophageal

speaker as compared to laryngeal speaker. But the average F3 for vowel /u/ was lower in esophageal speakers (1839 Hz) compared to the normal speaker (2373 Hz). This may be related to the difference in tongue configuration between esophageal & normal speaker. This decreased value in F3 was consistent with previous finding reported by Liu et al (2005).

Condition 2-sustained:

The first & second formant:

The difference in F1 values for the vowels / a/, /i/ and /u/ were higher by 183.31 Hz, 90.13 Hz and 45.90 Hz respectively in esophageal speaker as compared to the normal speaker. The difference in F2 for the vowels /a/, /i/, and /u/ were higher by 269.38 Hz, 455 Hz and 197.27 Hz respectively in the esophageal speaker as compared to the laryngeal speaker. The increased values in F1 & F2 are consistent with previous finding reported by Sisty & Weinberg (1972) for English speaking esophageal speakers. Cervera et al (2001) reported similar findings in the study of acoustic analysis of Spanish vowels produced by esophageal speakers. This is also compatible with the higher formant values observed in Dutch-speaking tracheo esophageal speakers who were also using the PE segment as the neoglottis (Van et al, 1997).

The third formant:

The mean difference of F3 values for /a/, and /i/ were 587.60 Hz and 166.70 higher in esophageal speaker as compared to normal speaker. But the average F3 for vowel /u/ was lower by 297.79 Hz in esophageal speakers as compared to the normal speakers. This may be related to the difference in tongue configuration between esophageal & normal speaker. This decreased value in F3 was consistent with previous finding reported by Liu et al (2005).

Condition 3-word level:

The first & second formant:

Number of studies has been carried out estimating the formant characteristics of vowels produced by esophageal speaker in word context in different languages in western countries. Cervera et al (2001) studied the formant characteristics of vowels in the context of nasal, liquid, fricative & stops and they found higher formant values in esophageal speaker as compared to the normal speakers. In the present study, vowel formant frequencies are studied in the context of a single consonant [K] and it was found that the difference in F1 for the vowels / a/, /i/ and /u/ was 160.48 Hz, 92.70 Hz and 50.10 Hz higher in esophageal speakers as compared to the normal speakers. Similarly the difference was found to be 161.67 Hz, 539.07 Hz and 142.97 Hz, in the case of F2 for the same vowels in the esophageal speakers as compared to the laryngeal speakers. Our results are consistent with the study reported for Spanish language by Cervera et al (2001). The increased values in F1 & F2 are also consistent with previous findings reported by Sisty & Weinberg (1972). The reason for this finding can be attributed to coarticulation effects (Stevens & House, 1972). Another reason cited for this finding is that esophageal speakers articulate vowels with front and higher tongue position relative to the tongue position in the normal speakers (Delattre, Liberman,& Gerstman, 1952; Peterson & Barney,1952).

The third formant:

The mean difference in F3 value for /a/, and /i/ were higher by 290.36 Hz and 194.92 Hz respectively in esophageal speaker as compared to the normal speaker. But the mean difference of F3 for vowel /u/ was lower by 445.94 Hz in esophageal speakers as compared to normal speakers. This may be related to the difference in tongue configuration between esophageal & normal speaker. This decreased value in F3 is consistent with previous finding reported by Liu et al (2005).

Comparison among same group in between three conditions:

The formant values across the three conditions in case of esophageal as well in normal speakers are discussed by applying Post Hoc test through Tukey method.

Vowel /a/:

For multiple comparisons, Post Hoc test was applied through Tukey HSD for vowel /a/ and it is noted that among normal group there is statistically significant difference for F2 & F3 values and no significance for F1 value for vowel /a/ under sustained Vs non-sustained conditions. Statistical significance is also noted for F2 value under non-sustained Vs word level production and F1, F2, & F3 values under sustained Vs word level conditions. This finding can be attributed to coarticulation effects (Stevens &House, 1972) On the other hand among esophageal group, it is noted that there is no statistically significant difference between the vowels articulated in non-sustained & sustained conditions. But statistical significance is noted for F2 and F3 values for

/a/ both in non-sustained Vs word level production and word level vs sustained production. This can be also can be attributed to coarticulation effects (Stevens &House, 1972)

Vowel /i/:

For multiple comparisons, Post Hoc test was applied through Tukey HSD for vowel /i/and it is noted among normal group for the vowel /i/, significance is noted only for F2 value under non-sustained Vs word level production. There is no statistical significance observed under sustained Vs non-sustained and sustained Vs word level conditions. On the other hand among esophageal group, for the vowel /i/ significance is noted only for the F1 under non-sustained Vs word level condition. There is no statistical significance observed under sustained Vs non-sustained and word level Vs sustained conditions in all the three formant values.

Vowel /u/:

For multiple comparisons, Post Hoc test was applied through Tukey HSD for vowel /u/ and it is noted among normal group that for the vowel /u/, significance is noted for F3 value under non-sustained Vs sustained, F1 under non-sustained Vs word level and F3 value under word level Vs sustained condition. There is no statistical significance observed in the formant values F1 & F2 under sustained Vs non-sustained and sustained Vs word level conditions. On the other hand among esophageal group, no significance is noted for the vowel /u/ under any three conditions.

VI. Summary & conclusion:

The primary goal of rehabilitation after laryngectomy is to return the patient as nearly as possible to his preoperative physiological, social and economic status. Achieving this goal depends significantly on the patient's ability to communicate effectively. Esophageal speech, though traditionally considered the method of choice is acquired as an effective mode of communication by meager percentage of patients involving considerable therapy time and variant speech proficiency.

Thus, the results indicate the need for analyzing voice or speech of laryngectomees to plan and monitor the therapy programme, that is, there is need for different treatment goals and priority of these goals in esophageal speakers. Further, the measurement procedures used in this study would be useful in evaluating the therapy techniques and progress made by the laryngectomee. Further, the present study provides an insight to the speech clinicians in understanding the probable variations in formant characteristics of Hindi esophageal speakers as against the other esophageal speakers.

By comparing this result of hindi speakers with the previous studies, it is noted that the difference in formant characteristics of vowels between esophageal speech and normal speech to be similar across languages. Therefore, the factor of language difference seems to have little effect to the change in vocal tract transmission after laryngectomy.

It is noteworthy that vowel formant frequency is in theory determined by the configuration of the vocal tract which can be affected by a numbers of factors: (1) the intrinsic size of the vocal tract, (2) the size of the tongue, (3) the size and configuration of the oral cavity, (4) the size and configuration of the pharyngeal cavity, and (5) the tongue configuration.

Reference

[1]. Aithal, V. (1997). Acoustic and temporal analysis of Tacheo-oesophageal speech.Master's thesis, university of Mysore, Karnataka, Ind

[2]. Bellandese, M. H., Lerman, J. W., & Gilbert, H. R. (2001). An acoustic analysis of excellent female esophageal, tracheoesophageal, and laryngeal speakers. Journal of Speech, Language and Hearing Research, 44, 1315-1320.

[3]. Bennet, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial laryngeal speech. Journal of Speech and Hearing Research, 16, 608-615.

[4]. Bentzen, N., Guld, A., & Rasmussen, H. (1976). X-ray video-tape studies of laryngectomized patients. Journal of Laryngology and Otology, 90, 655-666.

[5]. Case, L. J. (1984). Clinical management of voice disorders. Michigan: Aspen publisher group.

[6]. Cervera, T., Miralles, J. L., & Gonzalez-Alvarez, J. (2001). Acoustical analysis of Spanis vowels produced by laryngectomized subjects. Journal of speech, Language and Hearing Research, 44, 988-996.

[7]. Damste, P. H., Lerman, J. W. (1969). Configuration of the neoglottis: An X-ray study. Folia Phoniatr. 21, 347-358.

[8]. Diedrich, W. M., Youngstorm, K. A. (1966). Alaryngeal Speech. Springfield,IL: Charles C. Thomas.

[9]. Dunn, H. K. (1950). The calculation of vowel resonances, and an electrical vocal tract. Journal of Acoustic Society of America,

[10]. Hariprasad, G.V.M., Rajashekhar, B., & Nataraja, N.P. (1993). Acoustic analysis of vowels in alaryngeal speech. Journal of All India Institute of Speech and Hearing, 24, 1

[11]. Kent, R. D. (1993). Vocal tract acoustics. Journal of Voice, 7, 97-117.

[12]. Kent, R. D., & Read, W. C. (1992). The acoustics analysis of speech. San Diego: Singular Publication group.

[13]. Kytta, J. (1964). Finnish oesophageal speech after laryngectomy: sound spectrographic and cineradiographic studies. Acta Otolaryngol, 195, 1-94.

[14]. Liu, H., Wan, M., Wang, X., & Lu, C. (2005). Acoustic characteristics of Mandarin esophageal speech. Journal of Acoustic Society of America, 118, 1016-1025.

[15]. Liu, H., & Ng, M. (2009). Formant characteristics of vowels produced by Mandarin Esophageal speakers. Journal of Voice, 23, 255-260.
[16]. Manwa, C., & Rhoda, C., (2009). An acoustical and perceptual study of vowels produced by alaryngeal speakers of Cantonese. Folia Phoniatrica et Logopaedica, 61, 97-104
[17]. Nord, L., & Hammerberg, B.(1989) "Analysis of laryngectomee speech-A progress report." In Eurospeech, 2493-2496.
[18]. Peterson, G., Barney, H. (1952). Control methods used in a study of the vowels. Journal of Acoustic Society of America, 24, 175-184.
[19]. Premlatha, B.S., Shenoy, A.M., & Anantha, N (1994). Speech evaluation after NLT and TL a comparative acoustical analysis. Indian Journal of Cancer, 31, 244-
[20]. Rajashekhar,B. (1991). Acoustic analysis of alaryngeal speech. Doctoral thesis, University of Mysore, Karnataka, India.
[21]. Robbins, J., Fisher, H. B., Blom, E. C., & Singer, M. I. (1984). A comparative acoustic study of normal, esophageal and tracheoesophageal speech production. Journal of Speech and Hearing Disorders, 49, 202-210.
[22]. Robbins, J. (1984). Acoustical differentiation of laryngeal, oesophageal, and tracheoesophageal speech. Journal of Speech and Hearing Research 1984, 27, 577-584
[23]. Rollin,W.J. (1962). A comparative study of vowel formants of esophageal and normal speaking adults. PhD Thesis, Wayne state University.
[24]. Rollins. (1962. Cited in Hariprasad, G.V.M., Rajshekar, B., Nataraja, N. (1994). Acoustical analysis of vowels in alaryngeal speech. Journal of All India Institute of Speech and Hearing, 24, 13-15.
[25]. Sisty, N.L., & Weinberg, B. (1972). Formant frequency characteristics of esophageal speech. Jounal of Speech & Hearing Research, 15, 439-448
[26]. Sunitha, M. (1994). Acoustic and Temporal parameter in Malayalam speakers using different types of TEP prosthesis. Master dissertation, university of Mysore, Karnataka, India.
[27]. Van, C. J., Van A. M. A., Beinum, F. J., Hilgers, F. J. M., & Pols, L. C. W. (1997). Formant frequencies of Dutch vowels in tracheoesophageal speech. Institute of Phonetic Sciences,University of Amsterdam, Proceedings, 21, 143-15
[28]. Weinberg. (1986). Cited in Trudeau and Qi (1990). Acoustic characteristics of female Te speech. Journal of Speech and Hearing Disorders, 55, 244-250.
[29]. Weinberg, B., & Bennet, S. (1972). Selected acoustic characteristics of esophageal speech produced by female laryngectomees. Journal of Speech and Heaing Research, 15,
[30]. Ylppo, A., & Sovijurvi, A. (1962**).** Sonagraphic and palatographic studies of full denture, half denture and edentulous cases. Acta odontal scand, 20**,** 257.
[31]. Zemlin, W. R. (1998). Speech and Hearing Science-Anatomy and Physiology. 4th ed.