Air Pressure and Airflow Differences between Esophageal and Tracheoesophageal Speech of Cantonese

Manwa L. Ng

University of Hong Kong
manwa@hku.hk

Abstract

The present study attempted to investigate the aerodynamic differences between esophageal (SE) and tracheoesophageal (TE) speech of Cantonese. Airflow data was obtained from sustained vowels, and pressure below the pharyngoesophageal (PE) segment was estimated from /ip^{h}ip^{h}i/ syllables produced by the alaryngeal speakers. Results indicated that SE speech was associated with a lower rate of airflow and a higher pressure below the PE segment than TE speech. Based on the pressure-flow relationship, it is inferred that the estimated neoglottal resistance is greater in SE speakers than in TE speakers. It is speculated that such difference in neoglottal resistance may be related to the use of different air reservoir mechanism between SE and TE speakers.

1. Introduction

As a surgical procedure for late stage laryngeal cancer, total laryngectomy involves the removal of the entire larynx. Due to the loss of vocal apparatus during the procedure, laryngectomized patients learn to regain phonation by adopting an alternative voicing method. Among the different types of alaryngeal phonation, standard esophageal (SE) and tracheoesophageal (TE) speech do not involve the use of an external device for sound generation. SE and TE speakers phonate by vibrating the pharyngoesophageal (PE) segment. The PE segment is composed of the inferior pharyngeal constrictor muscle, the cricopharyngeus muscle, and the upper portion of the esophageal sphincter [1]. The structural differences between the PE segment and the vocal folds are believed to contribute to the significant perceptual and acoustical differences between SE/TE and normal laryngeal (NL) phonation [cf. 2-13].

Although PE segment is used as the new voicing source (neoglottis) in both SE and TE phonation, significant differences in intelligibility, frequencies, duration and intensity have been reported [cf. 2,9-14]. The differences between SE and TE phonation may be related to the way air is stored. In SE phonation, the upper part of the esophagus is used as a new air reservoir, which can retain up to only 80 c.c. of air [15]. However, during TE phonation, air from the lungs enters the esophagus through the tracheoesophageal fistula created between the trachea and the esophagus just inferior to the PE segment [16,17]. The average lung capacity of an adult male is around 3,000 c.c. [18]. The significant difference in air reservoir may explain the differences between SE and TE phonation.

A large number of studies have reported the perceptual, acoustic, temporal, and vocal intensity characteristics of SE and TE speech [2-15,17,19,20]. Yet, relatively few studies reported the aerodynamic characteristics of SE and TE phonation. In studying six superior male SE speakers, Snidecor and Isshiki [21,22] found a lower mean rate of airflow when sustaining the vowel /i/ and reading passage associated with SE when compared to NL phonation. They reported a mean flow rate of 59.67 c.c./s for SE speech, as compared to 219 c.c./s for NL speakers in reading the first paragraph of the Rainbow passage. They attributed the lower mean flow rate in SE speech to the use of upper esophagus as air reservoir, and the PE segment as the new sound source (the neoglottis). In studying Italian TE speech, Motta, Galli, and Di Rienzo [23] found that greater airflow rate and volume were associated with better speech performance. Schutte and Nieboer [24] compared the airflow and sub-neoglottal (subpseudoglottic) pressure characteristics associated with SE and TE speech, and they found a higher sub-neoglottal pressure and greater airflow through the neoglottis in TE speech. However, in the study, sub-neoglottal pressure was obtained directly by using a pressure sensor inserted in the upper esophagus of alaryngeal speakers, which was invasive and might interfere with air intake especially for SE speakers.

Despite the handful aerodynamic studies of SE and TE phonation, specific and accurate information including neoglottal impedance and sub-neoglottal pressure during phonation is not available. There is a paucity of information regarding the aerodynamic differences between SE and TE phonation. In addition, information concerning the effect of air reservoir on the aerodynamic characteristics of neoglottal vibration is lacking. The present study served as an extension of the studies reported by Snidecor and Isshiki [22], Motta et al. [23], and Schutte and Nieboer [24]. The aerodynamic differences between SE and TE speakers of Cantonese were investigated. Mean rate of airflow during vowel prolongation was measured and the pressure below the PE segment (sub-neoglottal pressure) was estimated. Results were compared between SE and TE speakers. Based on the airflow and air pressure measurements, effects of air reservoir mechanism on neoglottal resistance in SE and TE phonation were discussed.

2. Methods

2.1. Speakers

Ten SE and twelve TE Cantonese male speakers participated in the present study. Fifteen SE and TE speakers were initially recruited for the experiment. Five SE speakers and three TE speakers did not complete the speech tasks and were therefore excluded from the study. The alaryngeal participants were superior speakers carefully selected from the New Voice Club of Hong Kong by practicing speech therapists. The alaryngeal speakers were all physically healthy, with ages between 55-73 years who had no reported history of speech, language, and/or hearing problems, except that associated with laryngectomy. While all TE speakers were Provox-valve users, all SE speakers were reportedly using injection as their main method of air intake.

2.2. Speech Tasks

The speech tasks used included vowel prolongation and /ip^{h}ip^{h}i/ syllable production. The participants were instructed to sustain the vowels /i/, /a/, /ɔ /, and /u/ five times for as long as they could. The order in which the vowels were produced was randomized for each speaker before the recordings took place. To estimate the pressure below the PE segment during speech production, the speakers were asked to produce the syllable /ip^{h}ip^{h}i/ five times. All speech samples were produced at a comfortable level of loudness.

2.3. Recording Procedure

In order to familiarize themselves with the speech tasks and the recording environment, the speakers were instructed to practice the speech tasks several times prior to the actual recording. A brief instruction of the recording procedure was given to each speaker before the recording. During the recording of aerodynamic signals, a facemask was held against the speaker's face to ensure a tight coupling between the face and the mask and to prevent a leak of DC airflow. A catheter was inserted into speaker's oral cavity via the corner of the mouth. It was used to measure the intraoral pressure during /ip^{h}ip^{h}i/ syllable production. The speakers were instructed not to bite the catheter during the experiment. Throughout the recording, the catheter was frequently checked for blockage by speaker's saliva. The airflow and air pressure signals were transduced and stored in computer for analyses.

2.4. Instrumentation and Measurements

Aerodynamic measurements were obtained by using an airflow and air pressure transduction system (MS-110, Glottal Enterprises) via an undivided facemask. The mask was used to cover both the face and the nose of the speaker, but not the tracheostoma. Before each recording, the system was carefully calibrated according to the user's manual. Mean peak rate of airflow in vowel prolongation was measured from the sustained vowels. In order to calculate the mean rate of airflow for each vowel, the maximum point on the airflow signal contour for each vowel was selected and the corresponding flow value was recorded. The recorded values were averaged for the three productions of each vowel.

Similar to estimating subglottal pressure in laryngeal speakers, the pressure below the PE segment for SE and TE speech was estimated from /ip^{h}ip^{h}i/ production by averaging the peak intraoral pressure values measured at the two /p^h/ phonemes. The technique of estimating subglottal pressure by means of intraoral pressure during the closure period of voiceless stops was proposed and discussed by Rothenberg [25], Smitheran and Hixon [26,27], and later validated by Lofqvist, Carlborg, and Kitzing [28]. The same technique was used in the present study to estimate the sub-neoglottal pressure in SE and TE speakers. Intraoral pressure values were measured during the closure period of the first and second stops of the /ip^{h}ip^{h}i/ syllable. The pressure measurement directly indicates the power of the air supply to the phonatory mechanism. The stronger is the air supply, the higher are the pressure values.

3. Results

3.1. Mean rate of airflow in vowel utterance

Table 1 shows the average airflow rate associated with different vowels produced by SE and TE speakers of Cantonese. It is apparent that TE speakers produced the Cantonese vowels with higher airflow rate than SE speakers for all four Cantonese vowels (see Figure 1). To assess the effects of vowel and phonation type on airflow rate, a mixed-design multi-factorial analysis was carried out using *phonation type* as the between-subjects variable and *vowel* as the within-subjects variable. Results of the two-way repeated-measure Analysis of Variance (ANOVA) revealed no significant interaction between *phonation type* and *vowel*, and no significant main effect for *vowel*. However, a significant main effect was found for phonation type; SE speech was associated with significantly lower airflow rate than TE speech [$F(1, 20) = 69.558$, $p = 0.000 < 0.001$].

Table 1. *Mean airflow rate (mL/s) for esophageal (SE) and tracheoesophageal (TE) speakers in vowel prolongation.*

Average peak airflow rate (mL/s)					
	Speakers	Vowels			
		/i/	/a/	/ɔ/	/u/
SE	1	53.4	75.4	84.0	62.2
	2	67.0	44.4	76.4	45.4
	3	72.0	69.0	123.4	112.0
	4	38.6	154.0	52.8	109.6
	5	76.6	66.0	37.2	69.8
	6	50.6	75.8	54.8	47.2
	7	74.2	65.2	92.0	79.4
	8	60.6	67.8	137.4	57.8
	9	56.6	60.8	71.8	64.0
	10	71.4	58.0	110.6	70.0
	Mean	62.1	81.5	84.0	71.7
TE	1	137.6	117.6	138.4	155.0
	2	87.0	150.0	130.0	60.6
	3	113.0	102.8	320.2	150.8
	4	107.0	110.0	123.0	140.0
	5	197.4	189.4	219.6	127.8
	6	156.0	143.4	139.8	122.8
	7	107.4	173.4	138.8	125.6
	8	100.6	319.6	143.8	114.8
	9	132.4	123.8	210.0	131.0
	10	133.4	114.6	140.2	125.2
	11	126.0	151.2	139.0	330.0
	12	112.4	220.2	121.8	134.6
	Mean	127.9	134.7	138.7	135.3

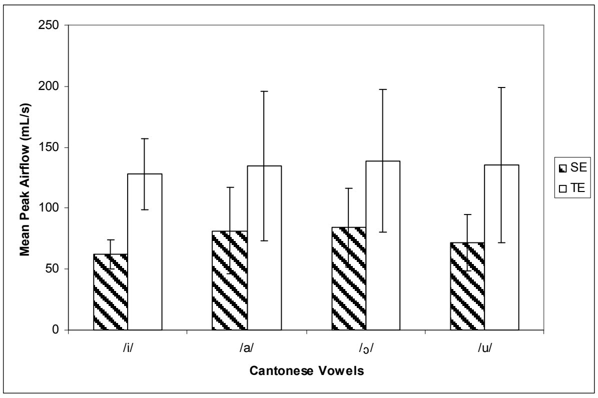

Figure 1. *Mean peak airflow rate associated with Cantonese vowels produced by esophageal (SE) and tracheoesophageal (TE) speakers.*

3.2. Estimated Pressure below the PE segment

Sub-neoglottal pressure was interpolated from the intraoral pressure measurement during the closure period of the first and second stops in the /ip^{h}ip^{h}i/ syllable, and results are shown in Table 2. Results of a Mann-Whitney U test indicated that SE speakers exhibited significantly greater estimated sub-PE segment pressure than did TE speakers ($U = 15.5$, $p = 0.003 < 0.01$). This indicates that, despite the use of upper esophagus as a new air reservoir, SE speakers are still able to maintain a sufficiently high pressure underneath the PE segment for phonation.

Table 2. *Average estimated sub-neoglottal pressure (cmH$_2$O) for esophageal (SE) and tracheoesophageal (TE) speakers.*

	Speakers						
	1	2	3	4	5	6	
SE	10.56	27.90	25.84	25.44	31.00	25.46	
TE	17.26	24.28	23.30	23.18	21.82	22.06	
	7	8	9	10	11	12	mean
SE	24.60	29.40	27.06	24.00	--	--	25.13
TE	24.00	22.82	24.80	22.44	22.20	23.14	22.61

4. Discussion

The present study examined the aerodynamic characteristics associated with SE and TE speech of Cantonese. Airflow rate was measured from sustained vowels, and oral pressure was obtained from /ip^{h}ip^{h}i/ syllable production, based on which neoglottal resistance was inferred. The average airflow rates found for Cantonese SE speakers in vowel prolongation were higher than that reported by Snidecor and Isshiki [22], where English-speaking SE speakers exhibited a mean rate of airflow of 59.67 mL/s with a range of 25 to 97 mL/s in reading the first paragraph of the Rainbow passage. Such discrepancy in airflow rate may be related to the language difference and/or the use of different speech materials. The use of sustained vowels in the present study might help yield a slightly higher average airflow rate than passage reading due to the presence of consonantal sounds.

Despite the use of steady state vowels, both the airflow rates associated with SE and TE speakers were highly variable (as shown by the error bars in Figure 1). This indicates that, even with superior SE and TE speakers, sound production was still unsteady and exhibited high perturbation. This may be related to the lack of good control of the neoglottal vibration system. In addition, although the present data showed that airflow rate across different vowels was not significantly different, the high vowels /i/ and /u/ appeared to be associated with slightly lower average peak airflow than the low vowels /ɔ / and /a/ (see Figure 1). This may be related to the intrinsic pitch of vowels in laryngeal phonation [cf. 29,30], according to which high vowels are associated with higher pitch and greater laryngeal tension possibly due to anterior tongue pull [cf. 31,32]. Based on this, the reduced airflow rate may imply a greater PE segment tension during production of high vowels by SE and TE speakers.

Airflow rate and air pressure data obtained from SE and TE speakers in the present investigation showed interesting differences. Airflow data from vowel prolongation indicated that, although both speaker groups were using the PE segment as the neoglottis, SE speakers produced lower airflow rate than TE speakers, regardless of the vowel being produced (see Table 1 and Figure 1). Yet, SE speakers showed significantly greater pressure below the PE segment than TE speakers (see Table 2). This finding of greater sub-neoglottal pressure in SE speakers contradicts with the results reported by Schutte and Nieboer [24]. This discrepancy in airflow rate may be related to the use of invasive intra-esophageal pressure measurement technique in their study. The placement of a distal sensor as used by Schutte and Nieboer in the esophagus may inevitably affect the pressure measurement. In addition, this may likely affect the way the SE speakers intake air.

In the present study, undivided mask was used to collect the air escaped through the nostril and the mouth opening; thus the airflow measurement directly reflected the airflow through the neoglottis, known as the trans-neoglottal airflow rate. In laryngeal phonation, transglottal airflow is directly proportional to transglottal pressure differential, and the proportionality constant is referred to as the glottal resistance. It has been shown that glottal resistance can be a good indicator of the adductory behavior of the vocal folds [18]. It is directly related to the tension of glottal closure. Similarly, the pressure differential and airflow associated with SE and TE speech can be expressed as:

$$\Delta P_{ng} = Q_{ng} \times Z_{ng} \quad (1)$$

where ΔP_{ng} is trans-neoglottal pressure differential, Q_{ng} is trans-neoglottal airflow rate, and Z_{ng} is the neoglottal resistance. The equation indicates that neoglottal resistance provided by the vibrating PE segment during SE and TE phonation is given by the ratio of the trans-neoglottal pressure differential (or simply the sub-neoglottal pressure) and the air flowing through the neoglottis. Our data show that SE speakers of Cantonese exhibited a lower trans-neoglottal airflow rate and a greater sub-neoglottal pressure than TE speakers. Following the above equation, SE speakers had a higher trans-neoglottal resistance, and thus a tenser neoglottis, than TE speakers.

In TE phonation, pulmonary air is used to set the PE segment into vibration. The expulsion of air from the lungs is activated and controlled by inhalatory and exhalatory muscles, in a way similar to laryngeal phonation [5,18]. According to the anatomy of respiratory system, these muscles are independent of those constituting the neoglottis. Control of air expulsion appears to be totally separated from the control of the neoglottal tension. However, in the case of SE phonation, air is stored in the upper part of the esophagus which also makes up part of the neoglottis. Upon phonation, air is expelled by increasing the pressure inside the upper esophagus. This is done by tensing the upper esophagus, which inevitably tenses the constriction of PE segment. It follows that the neoglottal resistance in SE speakers may be higher than that of TE speakers. The use of upper esophagus as the air reservoir in SE speakers may explain why the trans-neoglottal airflow in SE speakers is lower than that in TE speakers. Apparently, such discussion is conjectural and solely based on the aerodynamic data obtained. More direct information such as imaging and physiological data (e.g. EMG) of the neoglottis is needed in order to better understand the PE segment vibratory behavior in both SE and TE phonation

5. Conclusions

The present study investigated the airflow and air pressure differences between SE and TE speech. Data on mean airflow rate in vowel prolongation and estimated sub-neoglottal pressure indicated that SE speakers exhibited a lower rate of airflow and a higher pressure below the PE segment than did TE speakers. Based on this finding, it can be inferred that the neoglottal resistance is greater in SE speakers than in TE speakers. The greater neoglottal resistance in SE speakers appears to be due to the use of the same muscle group controlling both the air reservoir and the PE segment, while the PE segment and the air reservoir in TE phonation are controlled by two separate groups of musculature.

6. References

[1] Edels, Y., 1983. Pseudo-voice: its theory and practice. In Y. Edels (Ed.), *Laryngectomy: Diagnosis to Rehabilitation* (pp. 107-141). Sydney, Australia: Croom Helm, 107-141.

[2] Gandour, J.; Weinberg, B., 1985. Production of intonation and contrastive stress in esophageal and tracheoesophageal speech. *Journal of Phonetics* 13, 83-95.

[3] Gandour, J.; Weinberg, B.; Petty, S.H.; Dardarananda, R., 1987. Vowel length in Thai alaryngeal speech. *Folia Phoniatrica et Logopaedica* 39, 117-121.

[4] Gandour, J., Weinberg, B., Petty, S. H., & Dardarananda, R. (1988). Tone in Thai alaryngeal speech. *Journal of Speech and Hearing Disorders, 53*, pp. 23-29.

[5] Hartman, D.E., 1979. Perceptual and acoustic characteristics of esophageal speech. In *Laryngectomee Rehabilitation*, R.L. Keith; F. L. Darley (eds.). Cleburne, TX: College-Hill Press, 87-106.

[6] Liu, H.; Wan, M.; Wang, S.; Wang, X.; Lu, C., 2005. Acoustic characteristics of Mandarin esophageal speech. *Journal of the Acoustical Society of America* 118(2), 1016-1025.

[7] Liu, H.; Ng, M.; Wan, M.; Wang, S.; Zhang, Y., 2007. Effects of place of articulation and aspiration on voice onset time in Mandarin esophageal speech. *Folia Phoniatrica et Logopaedica* 59, 147-154.

[8] Liu, H.; Ng, M.; Wan, M.; Wang, S.; Zhang, Y., 2008. The effect of tonal changes on voice onset time in Mandarin esophageal speech. *Journal of Voice* 22(2), 210-218.

[9] Ng, M.; Chu, R., 2009. An acoustical and perceptual study of vowels produced by alaryngeal speakers of Cantonese. *Folia Phoniatrica et Logopaedica* 61, 97-104.

[10] Ng, M.; Gilbert, H.R.; Lerman, J.W., 1998. Perceptions of tonal changes in normal laryngeal, esophageal, and artificial laryngeal male Cantonese speakers. *Folia Phoniatrica et Logopaedica* 50(2), 64-70.

[11] Ng, M.; Gilbert, H.R.; Lerman, J.W., 2001. Fundamental frequency, intensity, and vowel duration characteristics related to perception of Cantonese alaryngeal speech. *Folia Phoniatrica et Logopaedica* 53(1), 36-47.

[12] Ng, M.; Kwok, C.I.; Chow, S.F., 1997. Speech performance of adult Cantonese-speaking laryngectomees using different types of alaryngeal phonation. *Journal of Voice* 11(3), 338-344.

[13] Ng, M.; Wong, J., 2009. Voice onset time characteristics of esophageal, tracheoesophageal, and laryngeal speech of Cantonese. *Journal of Speech, Language and Hearing Research* 52, 780-789.

[14] Robbins, J.; Fisher, H.B.; Blom, E.D.; Singer, M.I., 1984. A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders* 49, 202-210.

[15] van den Berg, J.; Moolenaar-Bijl, A.J., 1959. Criopharyngeal sphincter, pitch, intensity and fluency in esophageal speech. *Practica Oto-Rhino-Laryngologica* 21, 298-315.

[16] Panje, W., 1981. Prosthetic vocal rehabilitation following laryngectomy: the voice button. *Annals of Otology, Rhinology and Laryngology* 90, 116-120.

[17] Singer, M.I.; Blom, E.D., 1980. An endoscopic technique for restoration of voice after laryngectomy. *Annals of Otology, Rhinology and Laryngology* 89, 529-533.

[18] Zemlin, W.R., 1987. *Speech and Hearing Science – Anatomy and Physiology* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

[19] Ching, T.Y.C.; Williams, R.; van Hasselt, A., 1994. Communication of lexical tones in Cantonese alaryngeal speech. *Journal of Speech and Hearing Research* 37, 557-571.

[20] Christensen, J.M.; Weinberg, B., 1976. Vowel duration characteristics of esophageal speech. *Journal of Speech and Hearing Research* 19, 678-689.

[21] Isshiki, N.; Snidecor, J.C., 1965. Air intake and usage in esophageal speech. *Acta Oto-Laryngologica* 59, 559-574.

[22] Snidecor, J.C.; Isshiki, N., 1965. Air volume and air flow relationships of six male esophageal speakers. *Journal of Speech and Hearing Disorders* 30, 205-216.

[23] Motta, S.; Galli, I.; Di Rienzo, L., 2001. Aerodynamic findings in esophageal voice. *Archives of Otolaryngology - Head & Neck Surgery* 127(6), 700-704.

[24] Schutte, H.K.; Nieboer, G.J., 2001. Aerodynamics of esophageal voice production with and without a Groningen voice prosthesis. *Folia Phoniatrica et Logopaedica* 54, 8-18.

[25] Rothenberg, M., 1982. Interpolating subglottal pressure from oral pressure. *Journal of Speech and Hearing Disorders* 47, 219-220.

[26] Smitheran, J.R.; Hixon, T.J., 1981. A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders* 46, 138-146.

[27] Smitheran, J.R.; Hixon, T.J., 1982. A reply to Rothenberg. *Journal of Speech and Hearing Disorders* 47, 220-223.

[28] Löfqvist, A.; Carlborg, B.; Kitzing, P., 1982. Initial validation of an indirect measure of subglottal pressure during vowels. *Journal of the Acoustical Society of America* 72(2), 633-635.

[29] Peterson, G.E.; Barney, H.L., 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175-184.

[30] Lehiste, I.; Peterson, G.E., 1961. Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33, 419-425.

[31] Ladefoged, P., 1964. Some possibilities in speech synthesis. *Language and Speech* 7, 205-214.

[32] Lehiste, I., 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.