Articulatory aerodynamics, contact pressures and sense of effort during tracheoesophageal speech. Jeff Searl

Articulatory Aerodynamics, Contact Pressures and Sense of Effort During Tracheoesophageal Speech

Jeff Searl

Hearing and Speech Department, University of Kansas Medical Center, Kansas City, Kansas, United States of America

Abstract

Objectives: The purposes of this study were to: 1) describe sense of effort (SOE) associated with speech, oral air pressure (Po), articulatory contact pressure (ACP), and subneoglottal pressure (Psub) in tracheoesophageal (TE) speech, and 2) evaluate the relationships between these measures. The intent was to inform about articulatory alterations in TE speech with a future goal of guiding therapeutic interventions that serve to balance the physical demands of producing intelligible TE speech with the perceived work of generating the speech.

Patients/Materials and Methods: Thirty four TE speakers provided ratings of SOE. ACP, Po and Psub were measured during production of alveolar consonants.

Results: Po, ACP, Psub, and SOE were higher than what is reported for non-laryngectomized speakers. The primary locations identified by subjects as having increased SOE were the lips and tongue, followed by the brain (i.e., cognitive effort), voice, and lungs. Individuals who reported elevated SOE had higher ACP and Po than those who did not report an increase in SOE. SOE, ACP and Po were all moderately to strongly correlated with each other.

Conclusions: More than half of the TE speakers reported elevated SOE. ACP, and also Po appear to be possible physiologic correlates of internal SOE. These parameters could serve a useful therapeutic function as feedback to individuals with a goal of optimizing physical effort while maintaining the perceptual integrity of the consonants being produced.

Keywords: oral pressure, contact pressure, articulation, effort
1. Introduction

Two fundamental alterations to speech production that result from total laryngectomy are the removal of the voice source (i.e., the vocal folds) and separation of the upper from the lower airway with the trachea routed to a stoma at the base of the neck. These anatomical alterations create two primary issues to address in the alaryngeal speech rehabilitation process. First, a means of introducing a vibration into the vocal tract must be established. Second, a mechanism for creating airflows and pressures within the oral cavity and pharynx is needed to allow production of other important sounds for speech such as bursts and frication which are integral to the production of stops, fricatives and affricates.

In general, the tracheoesophageal (TE) speech production process more closely parallels the non-laryngectomy speech process than does esophageal or electrolaryngeal speech. In TE speech, a connection between the lower and the upper airway is re-established through a surgically created fistula in the common wall between the trachea and esophagus (see van As and Fuller [2006] for a brief historical review). Air flowing through a one-way valve placed in the fistula can set tissue in the upper esophagus and lower pharynx (pharyngo-esophageal [PE] segment) into vibration (Singer & Blom, 1980). As such, TE speech utilizes a voice source comprised of biologic tissue inherent to the speaker, unlike artificial larynx speech which utilizes an external device and a mechanical vibration that is external to the body. Additionally, TE speech utilizes pulmonary airflow, unlike esophageal and electrolaryngeal speech, to drive neoglottal vibration. This pulmonary air source provides a large volume of air that provides certain advantages over esophageal speech such as increased phrase durations, fewer pauses, and potentially louder and more stable voice (Max, Steurs & DeBruyne, 1996; Pauloski, 1998; Robbins, Fisher, Blom, Singer, 1984). Additionally, the larger volume of air provides an opportunity to recreate a more consistent and sustained air stream through the pharynx and oral cavity that can be utilized for creating pressure build-up for bursts and turbulent air flow for frication. Without this connection to the lower airway, artificial larynx and esophageal speakers must generate bursts and frication by other means, most commonly by using the lips and tongue to compress the air within the closed oro-pharyngeal space and then quickly releasing it for a burst or releasing and simultaneously moving an articulator to generate a short period of frication.

Not only does the TE speech process more closely parallel non-laryngectomy speech, the speech product has also been found to be a closer match than esophageal of electrolaryngeal speech. For example, although dependent on an individual speaker, TE speech is generally considered to be closer to laryngeal speech in terms of voice quality (Pindzola & Cain, 1988; Williams & Watson, 1987), the mean fundamental frequency (Robbins et al, 1984), pitch variability (Trudeau, 1994), intensity (Max, Steurs & DeBruyne, 1996), and speaking rate.
(Pauloski, 1998). TE speech can provide a functional communication method for those who have had a laryngectomy and it can do so in a relatively short time frame. Despite success with TE speech, however, it has consistently been found to be less intelligible than speech from non-laryngectomized speakers (e.g., Ainsworth & Singh, 1992; Doyle, Danhauer & Reed 1988). Reasons for the reduced intelligibility could vary widely across speakers. The vast majority of research and clinical efforts regarding TE speech have focused on understanding and characterizing the new voice source and seeking means of improving it (e.g., Robbins et al, 1984; van As et al, 1999; van As et al, 2005). Such efforts have high value and have contributed much to maximizing the TE rehabilitation process. Investigations of non-voice source related aspects of TE speech, such as articulation, have received significantly less attention, although it is increasingly clear that other speech behaviors not strictly related to the voice source also can be altered. For example, altered velopharyngeal activity (Searl & Evitts, 2004), prolonged duration of consonants and vowels (Robbins et al, 1984; Searl & Carpenter, 2001), and increased oral air pressure (Po) on stops and fricatives (Saito, Kinishi & Amatsu, 2000) have been reported.

There is sparse empirical data regarding the articulatory adjustments that occur during TE speech. If one goal of the laryngectomy rehabilitation process is to maximize speech intelligibility, it would be useful to understand the TE articulatory behaviors so that appropriate and targeted therapeutic interventions could be designed if needed. The current study, as well as two prior ones (Searl 2002, 2007) were prompted principally with this overall goal in mind. The research focus on articulatory forces and aerodynamics had its genesis from four information sources. The first were general references in clinical training texts regarding the possibility of needing to increase articulatory precision when using TE speech (Gress & Singer, 2005; van As & Fuller, 2007). Second were a set of five articles in recent years noting increased Po on stops and fricatives in TE speech (Motta, Galli & DiRienzo, 2001; Saito et al, 2000; Searl, 2002, 2004, 2007). None of these articles could elucidate specifically why Po was elevated so significantly (on average approximately 2-4 times higher than in laryngeal speech). Searl (2004) speculated that the elevated Po might be an outcome of an intuited or trained response focused on production of stronger bursts and frication to provide a salient (and perhaps unambiguous) acoustic cue to the listener trying to decode a non-normal speech signal. Increased Po is associated with a stronger burst release in non-laryngectomized speakers. Alternatively, the elevated Po might simply be an obligatory outcome of the altered neoglottal aerodynamics. The PE segment has a higher resistance to airflow than do the true vocal folds (Weinberg et al, 1982). This creates a situation where the voice source driving pressure for the laryngectomy is notably higher than that for the non-laryngectomized speaker. The higher subneoglottal pressure could conceivably translate into a strong Po impulse once the PE segment opens (either for voicing on an adjacent voiced phoneme or if the PE segment opens in a devoicing gesture for a voiceless stop consonant). The current study incorporates measurement of subneoglottal air pressure simultaneously with Po.
and articulatory contact pressure (ACP) to better understand the possible relationship among these parameters. Regardless of whether the elevated Po is a trained/intuited response or a direct outcome of neoglottal aerodynamics, it does beg the question, what oral articulatory behaviors are enlisted or required to contain the elevated Po?

The third piece of information prompting the current study was a recent report from this lab indicating that lip ACP during bilabial stop consonants in TE speech were 2-3 times greater than in laryngeal speakers (Searl, 2007). From a clinical training perspective, it is of interest to determine whether the ACP needs to be this high; that is, could adequate Po be generated to result in perceptually sufficient bursts and frications without expending more physical resources than necessary. The issue of ‘excess effort’ involved in articulation has not been explored systematically relative to TE speech although the Po data along with the historical approach of training exaggerated articulation in alaryngeal speakers make this a potentially relevant area of interest. In the TE literature, elevated ‘effort’ has been presented in relation to respiratory activity and neoglottal voice generation (e.g., Monahan, 2005; van As & Fuller, 2007) in TE speakers who have strained voice or aphony. Anecdotally, there are reports within the clinic that speech is at times fatiguing and requires more ‘work’ based on the subjective descriptions of alaryngeal speakers in general. Interestingly, when comparing the magnitude of the Po to the ACP during TE speech in the Searl (2007) study, it is clear that ACP is substantially greater than the Po that is being contained; this held true for both TE speakers and nonlaryngectomized individuals. This opens the possibility that individuals might be trained to reduce ACP (and perhaps reduce the physical work involved in TE articulation), but still have the ability to contain sufficiently elevated Po to generate a strong burst or frication.

One final, related line of thinking contributed to the development of the current series of studies evaluating TE articulation. The Hypo-Hyper Theory of Speech (or Theory of Adaptive Variability) posits that an individual modifies speech behaviors along a continuum depending on the speech situation (Lindblom, 1990; Perkell, Zandipour, Matthies & Lane, 2002). Hyper-articulated or ‘clear’ speech may be utilized in situations where particularly clear speech is desired (e.g., lecturing) or when intelligibility or audibility is at risk (e.g., noisy environment, impaired hearing of the listener). Increased physical effort during hyper-articulation is hypothesized (Lindblom, 1990), although physiological evidence of this has not been sought. The lip ACP data from Searl (2007) might be evidence of hyper-articulated speech although this currently is the sole report on the matter. Other studies documenting elevated Po and less coarticulation in TE speech (as inferred from velopharyngeal aerodynamics [Searl, 2004]) also could be consistent with the interpretation that TE users have modified speech behaviors along a continuum more toward hyper-articulated speech, perhaps to maintain intelligibility. In this case, it may be clinically appropriate to determine the extent to which additional articulatory ‘effort’ is required to maintain an acceptable level of intelligibility. Economy of effort is an under-riding theme in the
hypo-hyper theory of speech; the question would be whether TE speakers are being economical in their articulatory adjustments or are they over- or under-compensating.

The overall goal of this study was to better understand consonant production in TE speech in terms of the articulatory effort and air pressures that are generated. The concept of ‘effort’ is addressed by quantifying the actual ACP generated between articulators during speech and also by global ratings of speech effort from the TE speaker. Sense of effort (SOE) ratings relative to speech production have been used with both normal and disordered speakers for various purposes (Solomon, 2000; Solomon, Robin & Luschei, 2000). SOE ratings have not been reported previously for TE speakers although Searl (2007) offers anecdotal comments in his discussion that some study participants reported varying levels of fatigue or increased ‘work’ to talk. Subneoglottal air pressure measurements are included to allow for more careful interpretation of the tongue-palate contact and Po data.

This study had the following purposes. 1) Describe ACP in TE speech for a set of lingu-alveolar consonants. At present, only data for the lips have been reported. 2) Describe Po and subneoglottic pressure (Psub) during TE speech for a set of lingua-alveolar consonants. Data for both measures have been reported; as such, the current data serve to broaden the data sets available in the literature. 3) Provide the first systematic description of the SOE related to TE speech production. This is to include global ratings of speech effort as well as an indication of the locus (i) of the effort throughout the thorax, neck and head. It is not clear at the moment whether elevated SOE is a phenomenon that rises to a level of significance necessitating attention by researchers or clinicians. 4) Evaluate the strength of the relationships among Po, Psub, ACP, and patient self-ratings of SOE. A strong relationship between Po and ACP has been found previously for the lips, but no data are available for the tongue. Inclusion of Psub will allow more informed interpretation of the Po data and the role that neoglottic alterations compared to oral articulatory changes might be playing in the Po that is measured. Inclusion of the SOE data within the set of planned correlations may allow identification of speech production parameters that vary in relation to internal speech effort state and which could serve as a focus in therapy if the physical expenditure for articulation is inordinately high. Overall, better understanding of the relationships between ACP, Po, Psub and SOE could shed light on the need (or lack thereof) to train higher/lower contact pressures in an effort to optimize both costs (sense of effort) and benefits (intelligible speech).

2. Patients/Materials and Methods

2.1. Subjects

Thirty four individuals who had a laryngectomy and used TE speech participated in this study. The group ranged in age from 46-71 years of age (mean: 61.4 yrs)
and was comprised of 22 males (65%) and 12 females (35%). All speakers met the following criterion: 1) TE speech was the primary mode of communication (patient report); 2) >4 months post laryngectomy (range: 4-38 months); 3) functional hearing ability (hearing screening, best ear, aided allowed); 4) normal speech prior to laryngectomy (patient report); and 5) negative history for stroke, head injury or other neurological disease that might impact speech. Two certified SLPs estimated TE speech intelligibility using a single words and sentence protocol (Yorkston & Beukelman, 1984). Intelligibility scores ranged from 80-97%. The SLPs also rated overall speech proficiency on a 5-point scale with 1=poor and 5=excellent. Ratings for each speaker averaged 4 or greater for all subjects suggesting a fairly highly proficient group of TE speakers.

Only those individuals with a standard laryngectomy were included (i.e., those with surgeries that included pharyngectomy, esophagectomy or glossectomy were excluded). Primary TE puncture was done for 25 of the subjects (74%). Radical neck dissection was completed unilaterally in 11 speakers and bilaterally in 7 others. Radiation therapy was completed by 65% of the participants (9 subjects pre- and 12 subjects post-laryngectomy). Cricopharyngeal myotomy at the time of total laryngectomy is standard in our facility and so 100% of the subjects were known to have some degree of surgical myotomy.

Blom-Singer voice prostheses were utilized by 22 speakers (65%); all of these were indwelling, 20Fr prostheses. Those using smaller diameter prostheses were excluded because the smaller flange size on the 16Fr devices made the prostheses modifications for subneoglottic pressure measurement difficult (see 2.3.3). The remaining 12 speakers used a Provox®2 prosthesis; again, those who may otherwise have qualified for the study, but who used a 17Fr Provox NID were excluded because of prosthesis modification issues. Twenty four speakers (~70%) used digital occlusion to produce voice; of these, 20 occluded using an HME and 4 used a digit directly over the stoma. The remaining 10 speakers used some type of handsfree valve during speech.

All had undergone speech therapy following laryngectomy per patient report, but varied widely in terms of the number of sessions (range: 2-16 visits) and focus of therapy. In general, however, the therapy focused on prosthesis management, voice generation, and use of HME and handsfree devices. Seven of the 34 used a handsfree valve.

2.2. Speech Stimuli

Consonant vowel (CV) syllable strings were constructed using /t/, /d/, /s/, /z/ and /n/ and the vowels /i/, /u/, and /a/ for a total of 15 CV constructions. Each syllable was produced as a 5-syllable string (e.g., tatatatata/) on one breath at comfortable rate and loudness. Three repetitions of each syllable series were produced with stimuli fully randomized for each speaker.
2.3. Instrumentation and Measures

2.3.1. Articulatory Contact Pressure (ACP)

ACP was measured using an Entran EPI-BO flatline transducer (Entran, Fairfield, NJ, USA) mounted on a thin palatal appliance as described in Searl (2003). Briefly, the transducer is 0.5mm thick with a 2.0mm diameter sensing surface at one end of a 2.3mmx6.5mm housing unit and it has response characteristics that have been found to be acceptable for measuring contact pressure during speech (see Searl, 2003; Searl, 2007; Hinton & Luschei, 1992; Thompson et al, 1997). A stone dental casting is made for each speaker. This casting is used to create a 0.5mm thick palatal appliance made from flexible thermo-forming coping material. The Entran transducer is mounted on the palatal mold on the alveolar ridge in midline at a location that yields the highest and most consistent pressure recordings during the 60 minute accommodation period that was imposed to allow speakers to adjust to the presence of the palatal appliance in the mouth (see Searl et al, 2007 for details regarding speaker adaptation to this device). Figure 1 shows the transducer, palatal appliance and also the oral pressure sensing tube (describe below). The transducer signal output was amplified and routed to one channel of a PowerLab 8SP digital recording device (ADInstruments, Colorado Springs, CO, USA). The PowerLab hardware and software system was used for signal conditioning (50Hz low pass filtered; 20kHz digitization, 16 bit precision). See Figure 1 which shows the transducer mounted on the thin palatal appliance.

2.3.2. Oral Air Pressure (Po)

A polyethylene tube (i.d.=2.1mm, o.d.=3mm) was custom molded to run in the corner of the mouth, through the buccal-gingival sulcus to the last maxillary molar where the tube was bent to project to the midline of the hard palate. The tube was attached to the palatal appliance at the bend around the molar and along the course of the projection to the midline of the hard palate. This arrangement was used in order to avoid interfering with anterior tongue movements involved in production of the alveolar consonants. In addition, it was important to insure that the opening of the tube tip was parallel to the oral airstream that would be present during /s/ and /z/ in order to avoid spurious pressure reading of maximum Po. Figure 1 depicts the Po tube attached to the palatal appliance. The tube was connected to a Setra #239 differential pressure transducer (Foxborough, MA, USA) and the output was routed to a second channel of the PowerLab set-up. The Po signal was also low pass filtered at 50kHz to eliminate the voicing component. Both the Po and the contact pressure transducers were calibrated prior to speech recordings from each subject as previously described (Searl, 2003).
2.3.3. Subneoglottic Air Pressure (Psub)

To measure air pressure in the upper esophagus, each speaker's voice prosthesis was modified to accommodate the presence of a polyethylene tube as shown in Figure 2a-b. A small incision was made through the tracheal and esophageal flanges at the 180° mark (i.e., 6 o'clock) so that the sensing tube was run through the flanges on the undersurface of the barrel of the prosthesis. A small amount of adhesive was placed around the tube at each flange to hold it in place and to block air from passing through the incisions. In this way, the sensing tube (1.6mm i.d., 2.2mm o.d.) did not occupy any portion of the voice prosthesis through which air was flowing from the trachea to the esophagus and also allowed the one-way valve to remain fully functional. The tip of this tube projected into the esophagus approximately 2-3mm; the tip was cut at a 90° bevel. An HME base plate was attached to the neck and an In-Health manometer attachment was placed in the collar of the base-plate. The polyethylene tubing was routed out of the stoma through the side port of the manometry collar attachment as depicted in Figure 2b. Again, a small amount of adhesive was applied at the site where the tube passed thru the manometry attachment. The polyethylene tube was then attached to a second differential pressure transducer (Setra #239) and complete stoma coverage could be obtained either with a digit (over an HME or directly over the housing unit itself) or with a hands-free valve in place. Each speaker used whichever method of stoma occlusion that they typically used during TE speech production. This left 4 subjects who typically used direct digital occlusion over the stoma (i.e., non-HME/valve users) who were forced to use digit occlusion over the housing unit opening. They were allowed as much time as needed to experiment with digit coverage using this set-up prior to data recording.
After initial testing, it was determined that a three-way stop-cock valve needed to be positioned in the tubing line after it exited the housing unit collar in order to allow for suctioning through the tube in the event that saliva obstructed the tube tip sitting in the esophagus (as indicated by a baseline shift in pressure being recorded). The investigator was careful to monitor this baseline throughout recording and apply small suction as needed to clear the line. It was also noted that a saliva swallow frequently was successful at clearing the line when it appeared to be plugged as indicated by a return to baseline.

Insertion of the prosthesis was complicated slightly by the presence of the tube, but generally was accomplished without difficulty. The investigator inserted all of the modified prostheses with the polyethylene tube already attached. For the Provox valves, it was possible to use the inserter tool as described by the manufacturer, although passing the modified prosthesis through the inserter did require additional force given the added width of the prosthesis barrel. For the Blom-Singer valves, the gel cap insertion procedure was used as described by the manufacturer. There was sufficient room in the gel cap to accommodate the presence of the polyethylene tube, although the cap could not extend down over the barrel of the prosthesis as far as it normally would if the tube was not present. For the first 12 subjects, flexible fiberoptic nasopharyngoscopy was completed with the scope passed into the upper esophagus. This was done to assess the placement of the modified prosthesis and to determine whether the polyethylene modification survived the insertion process as planned. In no instance was any difficulty noted and the tube appeared to be firmly in place. For the remaining subjects, endoscopy was utilized if there was some question about the insertion process or outcome, but typically this was not necessary. Prior to insertion a fine-tip permanent marker was used to place a mark on the polyethylene tube at the point that it passed out of the tracheal flange of the prosthesis. By looking for the mark on the tube, the investigator could tell whether the tube had dislodged forward or backward relative to the original placement and adjustments could be made as needed. With the modified prosthesis in situ, the neck was prepped for placement of a base unit. The polyethylene tubing was slid through the slit in the collar of the base plate and the base plate was attached to the neck. As with the
other transducers, the Psub transducer was calibrated prior to each subject's recording session.

2.3.4. Sense of Effort (SOE)

Sense of effort was reported in two ways. First, each subject was asked to rate the overall sense of effort that they perceived during typical speech production using a 5-point scale. Each scale point was labeled as follows: 0 = “no extra effort,” 1 = “minimally effortful”, 2 = “mildly effortful”, 3 = “moderately effortful”, and 4 = “severely effortful”. They were asked to consider this as an overall rating of effort and to not think about when their speech is at its best or its worst. They were also asked to consider this as a rating of “physical effort” to produce speech rather than “mental” or “psychological” effort. The rating was used for the correlational analyses in this study. For descriptive purposes, a second SOE report was gathered that forced subjects to identify and rank the locations of the SOE in the body. They were shown a lateral schematic of a torso, neck and head that depicted internal structures such as the lungs, top of the esophagus, throat, palate, tongue and lips. These were labeled on the schematic. A brain was also drawn in the schematic within the skull. Subjects were asked to place marks on the schematic where they perceived effort when speaking. They were told that they could place as many or few marks as they wished. If they made more than one mark on the schematic, they were instructed to place a number on each mark to indicate where they perceive the greatest effort (labeled #1), next greatest effort (#2), and so forth.

2.4. Procedures

Participation required two visits. On the first, subjects provided written consent to be in the study and then completed a history form and had their hearing screened. An audio recording was made as they completed a word and sentence intelligibility test for later scoring by two certified SLPs. A dental casting of the upper arch was then made following standard dental casting procedures; the cast was used to create the palatal appliance described above. Finally, they rated their overall SOE related to speech and marked the SOE figure to identify the loci of the effort. Subjects returned for a second visit 2-4 weeks after the first. Their first task on the return visit was to repeat the SOE ratings (overall rating and loci of effort markings) so that the stability of such ratings could be evaluated. They then wore the palatal appliance for one hour continuously. The contact pressure transducer and oral pressure tube were attached to the appliance throughout the accommodation period. During that hour, they engaged in casual conversation with a research assistant to help them accommodate to the appliance. Following the accommodation period, subjects sat in front of a computer screen that was used to display the stimuli in random order. On each screen was shown a reminder for them to say the stimulus on one breath using their typical pitch and loudness. The rate of presentation of the stimuli on the computer screen was under the control of the investigator to allow sufficient pause time between
productions or to ask for a repeat if a misreading or other anomaly occurred (e.g., too fast, soft/loud, etc.). The oral and subneoglottic air pressure baselines were monitored continuously by the investigator to look for drifts that might suggest obstruction of the tubing and small amounts of suction were applied as needed to clear the saliva.

For the analysis, peak pressures (contact, Po, and Psub) were taken from the central syllable of each syllable train. For each subject, the mean pressure (contact, Po, Psub, respectively) was calculated from the three repetitions produced by the speaker and this average pressure was utilized in the analyses below. Ten percent of the samples were measured a second time by the investigator and also by a trained laboratory assistant to evaluate measurement reliability. Correlations between the investigators first and second measurements were .96, .94, and .99 for ACP, Po and Psub, respectively. Paired t-tests for each of the three pressure measures were not statistically significant (p>.05 for each test) indicating good intra-measurement agreement. Likewise, for inter-measurement agreement, correlations between the investigator’s measurements and the trained assistants were high (.93, .93, and .97 for ACP, Po and Psub) and the paired t-tests were not statistically significant at the .05 alpha level.

3. Results

3.1. Data Collapse

Preliminary analysis was completed to determine whether vowel context influenced any of the three pressure measurements. A Consonant (5) x Vowel (3) ANOVA was calculated for each pressure measurement. None of the three had a statistically significant main effect of Vowel (ACP: F=2.11, p=0.74; Po: F=1.62, p=0.66; Psub: F=2.08, p=0.71). Subsequent analysis was completed after collapsing data across vowel contexts when evaluating differences in the pressure measures as a function of the consonant and when completing the correlation analyses.

3.2. Pressure Measurements

3.2.1. Articulatory Contact Pressure

Descriptive statistics for ACP values are in Table 1. A one-way analysis of variance for repeated measures was calculated to evaluate differences in contact pressures across the set of consonants. The computed F value of 1.94 was not statistically significant using an alpha level of .05 (p=0.62) indicating no difference in tongue-palate contact pressures across the set of five consonants.
Table 1. Means and standard deviations for contact pressure, oral air pressure, and sub-neoglottic pressure for each consonant. All pressures are reported in kPa.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>/t/</th>
<th>/d/</th>
<th>/s/</th>
<th>/z/</th>
<th>/n/</th>
<th>Grand mean (sd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (sd)</td>
<td>1.46 (0.45)</td>
<td>1.34 (0.33)</td>
<td>0.90 (0.31)</td>
<td>0.87 (0.27)</td>
<td>0.46 (0.21)</td>
<td>1.01 (0.31)</td>
</tr>
<tr>
<td>Contact Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (sd)</td>
<td>9.4 (3.8)</td>
<td>8.1 (3.9)</td>
<td>7.6 (2.9)</td>
<td>7.3 (3.3)</td>
<td>8.7 (4.1)</td>
<td>8.2 (3.7)</td>
</tr>
<tr>
<td>Sub-neoglottic Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (sd)</td>
<td>3.47 (1.18)</td>
<td>3.91 (1.48)</td>
<td>4.17 (1.49)</td>
<td>3.18 (0.93)</td>
<td>3.26 (1.57)</td>
<td>3.60 (1.33)</td>
</tr>
</tbody>
</table>

3.2.2. Oral Air Pressure

Table 1 includes means and standard deviations for Po across the five consonants. A one-way ANOVA for repeated measures resulted in an F of 61.09 and a p value of 0.008. This indicated that there were differences in Po across the consonant set. Post hoc comparisons (Tukey’s honestly significant difference test) were completed and indicated that Po for /n/ was significantly lower than that of the other 4 consonants. Values for /s/ and /z/ did not differ from one another but were both significantly lower than the Po for /t/ and /d/. Po for /t/ and /d/ did not differ significantly. Overall, the results suggest lowest pressure on the nasal phoneme, highest pressure on the stop consonants and intermediate pressure on the fricatives.

3.2.3. Subneoglottic Air Pressure

Subneoglottic pressure means and standard deviations are in Table 1. A one-way ANOVA for repeated measures was computed. The F value of 1.08 was not statistically significant (p=0.833) indicating that pressures did not differ significantly across stimuli.

3.3. Sense of Effort

3.1.1. Consistency of Ratings

A paired t-test was computed using each subjects ratings of overall SOE at the first and second visits to evaluate the consistency of such ratings. The resulting t-value of 1.33 was not statistically significant (p=0.72). This suggests that the overall SOE rating was consistent within the time frame of the study.
3.3.2. Magnitude of Overall SOE

Fifteen subjects (44% of the group) rated SOE as 0, indicating no perception of increased effort associated with speech production. The remaining 19 subjects had a mean rating of 2.2, standard deviation of 0.98 and a range from 1-4 on the 5-point scale. The mean rating suggests that, as a group, these 19 individuals perceived themselves as having slightly more than a mild increase in effort associated with speech production. Overall SOE was rated as minimal by 5 subjects, mild by 9, moderate by 4, and severe by 1.

3.3.3. Locus

Fifteen of 34 (44%) did not mark any location as having increased SOE on the lateral schematic. These were the same 15 subjects who rated overall effort as 0. The remaining 19 subjects marked at least one location on the schematic. The number of subjects identifying increased SOE at each site, and the distribution of rankings is reported in Table 2. From the tabled data it can be inferred that 4 subjects identified only one location of effort, 9 identified two sites, 3 identified three sites, and 3 identified 4 sites. The lips and the tongue were the most commonly identified sites of increased SOE followed by the brain, voice and lungs. The brain ratings were interpreted as indication of “mental effort.” The pharynx was never marked as having increased SOE.

Table 2. Descriptive information regarding locus of SOE for the 19 subjects who reported increased SOE. Values beneath each anatomical structure are the number of subjects marking a given site at the respective rank positions.

<table>
<thead>
<tr>
<th>Ranking Position</th>
<th>Lips</th>
<th>Tongue</th>
<th>Pharynx</th>
<th>Voice</th>
<th>Lungs</th>
<th>Brain</th>
<th>Number of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>7</td>
<td>--</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>4</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Number of subjects identifying the site (%)</td>
<td>13 (68.4)</td>
<td>12 (63.2)</td>
<td>0 (0.0)</td>
<td>5 (26.3)</td>
<td>4 (21.1)</td>
<td>9 (47.4)</td>
<td></td>
</tr>
</tbody>
</table>

3.3.4. Differences in Pressure Measures as a Function of Effort Rating

Three t-tests for independent measures were calculated to evaluate differences in Po, ACP, and Psub for those subjects without elevated SOE (N=15) compared to those with some degree of elevated effort (N=19). Table 3 reports the mean pressure values for the subjects divided into two groups considered “no extra effort” (rating of 0 on the 5-point scale) and elevated speech effort (ratings >1). Oral pressure and tongue-palate contact pressure were significantly higher for the subjects who reported increased speech effort compared to those who
reported no excess speech effort (Po: t=2.95, p=0.007; ACP: t=2.41, p=0.015). Psub did not differ significantly between the two groups (t=0.76, p=0.691).

Table 3. Means and standard deviations for the pressure measurements (kPa) for subjects who rated themselves as having no exceptional effort vs. those with some degree of elevated speech effort (>1 on the 5-point scale).

<table>
<thead>
<tr>
<th>Measure</th>
<th>No Speech Effort</th>
<th>Some Speech Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N=15)</td>
<td>(N=19)</td>
</tr>
<tr>
<td>Oral Pressure (Po)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.85 (0.36)</td>
<td>1.58 (0.56)</td>
</tr>
<tr>
<td>Articulatory Contact Pressure mean (ACP)</td>
<td>6.5 (2.7)</td>
<td>7.7 (3.0)</td>
</tr>
<tr>
<td>Subneoglottic Pressure (Psub)</td>
<td>3.4 (1.1)</td>
<td>3.7 (2.6)</td>
</tr>
</tbody>
</table>

3.4. Correlational Analysis

Multiple Pearson product moment correlations among SOE, ACP, Po and Psub were calculated. In order to limit the number of statistics computed, consideration was given to how data could be collapsed across consonants. Correlations with data collapsed across all five consonants are presented in Table 4 to look for more general trends in the TE speech data. Table 3 also presents correlations for subgroupings of the consonants depending on Po status. As noted above, there were significant differences between consonants in terms of Po; it was possible that these differences might obscure or weaken correlations of Po to the other variables in the total data collapse statistics.

Overall, SOE, Po and ACP appeared to be moderately to strongly correlated with one another. SOE was strongly correlated to ACP regardless of the consonant subgrouping (r values ranged from .76 to .86) and was moderately to strongly correlated with Po (r values ranged from .42 to .72). The correlation between ACP and Po was strong for the individual consonant subgroupings (.60 to .79). The subneoglottal pressure measure had notably weaker relationships to the other three variables even though some significant correlations were still present. There was not a significant correlation between Psub and Po for any of the subgroupings or for the combined data set. Excluding the correlations when using combined /t/ and /d/ data, Psub was significantly correlated with SOE and ACP although the r values ranging from .21 to .39 suggested only mildly strong relationships.
Table 4. Correlation Matrices for the four variables of interest. Statistically significant correlations at the .05 level are marked with an asterisk. ACP=Articulatory contact pressure, Po=Oral Pressure, Psub=Subneoglottal pressure, SOE=Sense of effort.

<table>
<thead>
<tr>
<th>All Consonants</th>
<th>Po</th>
<th>ACP</th>
<th>Psub</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOE</td>
<td>.42*</td>
<td>.76*</td>
<td>.24*</td>
</tr>
<tr>
<td>Po</td>
<td>---</td>
<td>.44*</td>
<td>.18</td>
</tr>
<tr>
<td>ACP</td>
<td>---</td>
<td>---</td>
<td>.21*</td>
</tr>
<tr>
<td>/t/ + /d/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOE</td>
<td>.58*</td>
<td>.86*</td>
<td>.20</td>
</tr>
<tr>
<td>Po</td>
<td>---</td>
<td>.79*</td>
<td>.19</td>
</tr>
<tr>
<td>ACP</td>
<td>---</td>
<td>---</td>
<td>.15</td>
</tr>
<tr>
<td>/s/ + /z/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOE</td>
<td>.72*</td>
<td>.80*</td>
<td>.39*</td>
</tr>
<tr>
<td>Po</td>
<td>---</td>
<td>.68*</td>
<td>.13</td>
</tr>
<tr>
<td>ACP</td>
<td>---</td>
<td>---</td>
<td>.29*</td>
</tr>
<tr>
<td>/n/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOE</td>
<td>.57*</td>
<td>.78*</td>
<td>.22*</td>
</tr>
<tr>
<td>Po</td>
<td>---</td>
<td>.60*</td>
<td>.19</td>
</tr>
<tr>
<td>ACP</td>
<td>---</td>
<td>---</td>
<td>.26*</td>
</tr>
</tbody>
</table>

4. Discussion and Conclusions

Maximizing speech intelligibility has been a long-standing goal of both researchers and clinicians working with individuals who utilize TE speech. In addition to maximizing the new voice in terms of its audibility, consistency, quality, and so forth, most clinicians seem willing to recognize the need to address articulatory behaviors, at least for some TE speakers. Unfortunately, there is limited empirical evidence to guide decisions about what aspects of articulation might be the most important to evaluate and address therapeutically. The current study adds to small, but growing knowledge base regarding articulatory behaviors during TE speech with an ultimate goal of guiding clinical decisions about the need (or lack thereof) for articulatory intervention, and specific approaches to that intervention. Consonant production has been an ongoing focus in this line of research, in part because of the important role that consonants play in intelligibility, but also because the anatomical changes from laryngectomy create a substantial alteration in speech aerodynamics (i.e., diversion of the trachea to the base of the neck; re-coupling the lower and upper airways via a fistula; altered neoglottal aerodynamics that might impact air flow upstream of the PE segment). The only significant amount of work on consonant production in TE speech that has received any degree of repeated investigation has been the ability to produce the voiced-voiceless distinction (Jongmans, Hilgers, Pols, van As, 2006; Saito et al, 2000; Searl and Carpenter, 2002). The
current study attempts to broaden the understanding of physiologic characteristics of consonant production in TE speech and to relate these instrumental measures to speakers’ ratings of effort.

4.1. Pressure Measurements

4.1.1. Articulatory Contact Pressure

Tongue-to-palate contact pressures during alveolar consonants produced by this group of TE speakers were substantially higher than what has been reported for a similar set of consonants in non-laryngectomized adults (Searl, Evitts and Davis, 2007). Both studies used the same instrumentation, although the speech stimuli varied somewhat. The slight differences in stimuli, however, are not likely to explain the approximate doubling of contact pressures demonstrated by the TE speakers. Lip ACP data from TE speakers also found that TE speakers produced bilabial phonemes with ~2 times higher pressure (Searl 2007). It is interesting to note that in both the current study and Searl (2007) that the nasal phonemes also were produced with this elevated contact pressure. Given the low oral pressure requirement for nasal phoneme production, one might hypothesize that only limited articulatory contact pressures would be invoked to contain the limited air pressure. However, it seems the case that the speech system does not attempt to modulate contact pressures during speech depending on a particular consonant (no differences as a function of consonant have been found for TE or laryngeal speakers in the studies from our lab). Rather, it seems that the speech system merely targets a contact pressure level that will be sufficient to contain air pressures across the range of consonants to be produced. For TE speakers, it appears that the speech system simply re-adjusts the ACP to a notably higher level. From an economy-of-effort perspective, it seems reasonable for the speech system to not have to specifically adjust contact pressures dependent on the consonant. However, the TE data do suggest that the overall level of physical effort expended during consonant articulation is greater than in non-laryngectomized speakers.

4.1.2. Oral Air Pressure
The Po data for this set of alveolar consonants are consistent with what has been reported previously (Motta et al, 2001; Saito et al, 2002; Searl, 2002, 2004, 2007). The pattern of differences in Po across the stimulus set were predictable with lowest Po on the nasal /n/; highest pressure on the stop consonants /p/ and /b/; and intermediate values for the fricatives /s/ and /z/. Overall, the Po values are notably higher than what is reported for non-laryngectomized adults (again, 2-3 times higher). The lower pressure on the nasal phonemes in TE speech is presumably for the same reason that low pressures occur on nasal in non-laryngectomy speakers. That is, the velopharyngeal port is open, and all else being equal in terms of respiratory drive and oral articulatory behaviors, a Po drop will occur because air can escape through the nose.
4.1.3. Subneoglottal Pressure

Although pressure within the esophagus of TE and esophageal speakers has been reported previously, the data are sparse. Subglottic pressure in non-laryngectomized speakers is expected to be within a range of approximately 5-10cmH2O (~0.49-0.98kPA) (Wilson & Leeper, 1992). The overall mean Psub for the TE speakers in this study was 3.5 times greater than the range expected from normal speakers. This is not surprising given earlier studies that have described the elevated resistance of the neoglottis (e.g., Weinberg et al, 1982).

A few comments regarding the procedure for obtaining the Psub measure in this study are warranted given the novelty of the approach. The intent was to avoid placing a tube within the lumen of the prosthesis which would diminish the cross-sectional area through which pulmonary air would flow and which would have held open the one-way valve at the esophageal end of the prosthesis. The current approach was generally successful although the prosthesis insertion process was somewhat uncomfortable for a small number of subjects, presumably because of the increased overall diameter of the prosthesis passing through the fistula. Once in place, subjects did not complain of pain or other sensations of discomfort, although when directly asked they generally reported that they could feel something different. No leakage of saliva around the prosthesis was noted for any of the subjects and their TE voices were not perceptibly different when the modified prosthesis was in place (voice recordings with the usual and the modified prosthesis are available for judging, but that has not been done to date). The problem of saliva blocking the esophageal tip of the sensing tube was generally easy to manage using the suction arrangement. Since that time, we have found it simpler to pass a length of acrylic fishing line through the sensing tube all the way to the esophageal tip to disperse saliva bubbles. The fishing line is flexible, but has enough stiffness to not coil up in the tube. By placing marks along the length of the sensing tube all the way out to the distal end that is attached to the transducer, a comparable length of fishing line can be passed through the tube to reach the esophageal tip.

4.2. Sense of Effort and Relationship among Measurements

As far as the author is aware, this is the first attempt to systematically describe the SOE associated with speech in general from a group of individuals using TE speech. Increased effort has been discussed for TE speakers with a focus mainly on issues of voice production (e.g., excess respiratory effort, neck tension, strained voice or aphonia, etc.). A sizeable portion of the group (~45%) did not report any sense of increased effort. However, 55% of the group reported at least some amount of increased effort, most in the mild range of the rating scale. It may be tempting to diminish reports of effort in the minimal or mild range. However, speech should not be effortful. In fact, non-laryngectomized individuals are usually not aware of any sense of effort associated with speech production unless they have been fatigued or they are required to significantly increase
loudness (see Solomon, 2000). For that reason, speakers with any degree of elevated SOE were grouped together for portions of the analysis.

Inspection of the data regarding the locus of the SOE suggests that the lips and the tongue were the most commonly identified sites of effort with over 60% of subjects marking either or both articulators. ACP also differed significantly when comparing those who reported no increased effort to those who reported some elevated effort. It seems reasonable to conclude that subjects SOE ratings and identification of the lip and tongue as sites of effort are a reflection of the physical effort involved in articulation as documented in the ACP measure. The correlational analysis also can be interpreted as support for that position. That is, ACP may be a measurable, physiologic correlate of the subjective rating of SOE. If true, it may be possible to use the ACP signal as feedback to an individual in an effort to manipulate SOE. Care would have to be taken to track the aerodynamic (Po) and perceptual outcome (consonant intelligibility or precision) when such manipulations are attempted. The intent would be to limit physical effort but still produce a strong consonant with a burst or frication that is sufficient for the listener to accurately decode what is being said. The fact that Po and ACP were also strongly correlated means that caution must be taken to not have the unintended effect of weakening the stop consonant or fricative. Of course, such future work is contingent on first determining whether the findings of increased SOE by TE speakers is replicated in other studies and determining whether the increased effort rises to a level that warrants consideration for manipulation.

The number of subjects identifying the voice as a source of effort was somewhat lower than anticipated. Knowing that the rehabilitation process generally focuses on the new voice and that neoglottal resistance and Psub are expected to be elevated (which it was in this study), one might have expected subjects to focus more on the neoglottis when identifying locus of effort. However, this was a highly proficient group of speakers who were not having trouble with the voicing aspects of TE production. It seems reasonable to expect that in other TE speakers, the locus of effort very well could be on the voice, with or without concurrent sense of effort in the articulators. Also, it may be that subjects had simply accommodated to the elevated driving pressures below the neoglottis so that it was no longer a salient perception to the speaker.

The fact that almost half of the group identified the brain as a site of effort is also telling. This was interpreted to reflect a sense of elevated cognitive load. Further exploration of this concept of cognitive effort might be important, particularly if means of reducing this effort can be identified. It may be that factors such as time-post laryngectomy or TE puncture are critical, or it may be that manipulations of other behaviors that reduce physical effort (e.g., reducing ACP) might also translate into lower cognitive loads.

The relationship of Psub to the other variables is less easy to interpret. Psub was not significantly correlated to Po. This suggests that the elevated Po that has
been consistently documented in TE speakers is not likely a direct outcome of elevated pressures below the neoglottis. It may be that the oral articulatory behaviors that result in increased ACP on these consonants is primarily responsible for increasing Po. For example, when a TE speaker is producing an alveolar stop, they may have been trained (or intuited the need) to generate a strong burst. To do so, they may have the tongue contact the palate more forcefully, and perhaps with greater surface contact (though this has not been documented to date). The net result might be further reduction of the oropharyngeal cavity volume, and subsequently greater compression of the air within the vocal tract (i.e., increased Po), relative to a production that is done less forcefully and with less complete contact between tongue and palate. Future work that incorporates methods of surface area contact between articulators, or even better, volume of the vocal tract, could help elucidate this issue.

Psub was correlated with SOE, but the strength of the correlations was weaker than that between ACP and SOE. The correlation between Psub and ACP also was fairly weak (although statistically significant for some of the consonant set). The ACP-Psub correlation suggests that individuals who demonstrate increased ACP also have higher Psub. TE speakers might be considered to vary along a continuum of physical effort (reflected in contact and air pressure measures) and that the level of effort may not be specific to one aspect of the process, but rather pervades most or all of the production parameters. Future studies that allow for simultaneous recordings of measures that might reflect effort across the speech system (respiratory, neoglottal, oral articulatory, etc.) are needed to explore this situation further. Results of this study suggest that contact pressures as well as some air pressure measures could serve as a physiologic correlate to perceptions of speech effort.

4.3. Future Directions

Studies are currently underway in which TE speakers are provided with ACP and/or Po biofeedback as they attempt to decrease ACP. The goal is to see if an individual can reduce the physical work of articulation (ACP), but maintain Po. Acoustic recordings of such efforts will be used to help determine whether the perceptual product is degraded as these manipulations are made. Studies of effort (both subjective and physiologic) in electrolaryngeal and esophageal speakers are of interest. It will be particularly interesting to compare measures of effort within speakers who are able to utilize more than one method of alaryngeal communication to see if one is more likely than another to induce greater articulatory demands or perceptions of greater articulatory effort.

5. Acknowledgements

Thanks are given to all of the individuals with a laryngectomy who willingly gave of their time. A portion of this work was supported by a Technology Innovation Enhancement Grant, Bowling Green State University.
6. References


