# Whispering by Individuals Using Tracheoesophageal Speech

**Jeff Searl,** *Kansas City, Kansas*

**Summary: Objectives.** This study compared whispering attempts by adults using tracheoesophageal (TE) speech with those by adults with a larynx. Comparisons were based on listener judgments, visual-perceptual assessment of spectrograms, and measures of the acoustic signal.

**Study Design.** This was a prospective, cross-sectional study.

**Methods.** Seventeen TE and 10 laryngeal speakers produced sentences in a whisper and in their spoken voice. Listeners judged sentences as whispered or spoken. Judges signal-typed the spectrograms based on presence-absence of a "voicing bar." Speaking rate, articulation rate, percent pause, and dB sound pressure level were measured.

**Results.** Twenty-nine percent of TE speakers were perceived to be whispering on whisper attempts; most others were perceived to be using spoken voice while attempting to whisper. Spectrograms of TE whispering were most often categorized as "mostly voiced." Speaking and articulation rates were slower for TE speakers. There was a significantly greater reduction in speaking rate from spoken to whisper for the TE group. Percent pause did not differ significantly between groups and speaking mode. TE speakers had a significantly smaller difference in dB sound pressure level between spoken and whisper modes.

**Conclusions.** Some individuals using TE speech can whisper based on auditory-perceptual judgment, but most were perceived to be speaking during these attempts. The fact that some TE participants could whisper indicates the behavior is possible and might be considered a therapeutic target if it is of importance to an individual. The percentage of TE speakers who can learn to whisper, and the optimal training approach, are yet to be determined.

**Key Words:** Tracheoesophageal speech–Whisper–Pharyngoesophageal segment–Neoglottis–Acoustic.

## INTRODUCTION

Individuals using tracheoesophageal (TE) speech set pharyngoesophageal (PE) tissue into vibration by diverting air from the lungs through a voice prosthesis. Understanding the extent to which the PE segment can be actively controlled has been of interest to clinicians intent on maximizing TE voice capabilities. Pre-dating the development of TE voice production, investigators reported that the PE voice source during esophageal speech was mediated principally by aerodynamic factors without active control to adjust parameters such as pitch and loudness.[1–3] However, Moon and Weinberg[4] reported that participants using TE speech were able to increase their fundamental frequency (F0) on a consistent basis, whereas trans-source airflow rate was maintained or decreased; four of five participants could also decrease F0 while maintaining or increasing the airflow rate. This control was interpreted as evidence for myoelastic as well as aerodynamic control of the TE voice.

Other studies have supported the conclusion that PE segment activity is not merely an aerodynamically controlled event. Support has come from studying various aspects of TE speech including F0 control and analysis of the voiced-voiceless distinction. The F0 range of the TE voice is restricted relative to the range in laryngeal speakers, but most have the ability to modulate F0 to some extent.[5–8] Cantonese speakers after total laryngectomy have demonstrated even greater range and control of F0 than TE speakers using a nontonal language.[9] F0 can be altered by only adjusting the driving air pressure to the voice source. However, Deschler et al[10] demonstrated that sound pressure level (SPL) of sentences by TE speakers accounted for a small amount of the variability (<10%) in F0. This finding was interpreted as further indication that aerodynamics is not the sole factor mediating behavior of the PE tissues.

Control over production of the voiced-voiceless distinction also provides insight about PE segment control. Perceptual confusions are common along the voicing dimension in TE speech, particularly for plosives and fricatives.[11–13] However, some can produce the voicing contrast accurately as determined by auditory-perceptual judgment.[14–16] Furthermore, acoustic studies confirm that many TE speakers use acoustic features to distinguish cognates including voice onset time, among other parameters, that are also used by laryngeal speakers.[15–17] More directly are high-speed video data indicating that some TE speakers do produce a PE opening gesture when producing voiceless stop consonants.[18] Electromyographic (EMG) evidence has also been reported demonstrating decreased muscle activity during voiceless sounds, suggesting active devoicing.[19]

The accumulating evidence suggests active control of the PE segment is possible, but the limits and determining factors of that control have not been fully explored. Active control over PE properties such as opening and closing, degree of closing force to adjust airflow resistance, and alteration to intrinsic stiffness are surely varied across individuals. However, understanding the range of this control can provide a framework for further understanding the PE voice production process and potentially could lead to the establishment of reasonable therapeutic targets for

Accepted for publication April 17, 2017.

Present affiliation and address as listed for the corresponding author. Bulk of this work occurred at the affiliation listed.

From the Hearing and Speech Department, University of Kansas Medical Center, Kansas City, Kansas.

Address correspondence and reprint requests to Jeff Searl, Department of Communicative Sciences and Disorders, Michigan State University, Room 109, Oyer Speech and Hearing Building, 1026 Red Cedar Road, East Lansing, MI 48824. E-mail: searljef@msu.edu

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997



http://dx.doi.org/10.1016/j.jvoice.2017.04.007

TE speech training that maximize the voice. Whispering can provide unique insight into PE segment control. Whispering by a person with a larynx is accomplished by creating turbulent glottal airflow without vocal fold vibration.[20] Vocal fold positioning can vary and is described in more detail elsewhere.[21] In TE speech, the PE tissues are ideally reconstructed so that there is a zone of tissue approximation at the top of the esophagus that is not overly tonic so as to preclude TE voicing, but sufficiently positioned to allow a pressure differential, allowing vibration to occur. Whispering would require volitional opening of the PE segment to allow generation of turbulent airflow without PE voicing in a manner analogous to a laryngeal speaker who whispers by maintaining some degree of glottal opening to create a turbulent glottal airstream without true vocal fold voicing. Some degree of neuromuscular integrity in the muscles making up the PE source would be needed (unless the PE is already hypotonic without the ability to produce TE voicing). A recent needle electrode EMG study in 34 individuals using TE speech called into question the integrity of the PE neuromuscular status after laryngectomy.[22] Classification of the EMG signal types revealed that all participants had neurogenic injury, with 20 judged to be severe in nature.

In addition to potentially learning about PE control capabilities, there is inherent communication value to whispering, about which nothing has been reported for people using TE speech. Whispering, although a rare mode of communication, is both functional and highly valued in specific situations.[23] In the social-anthropological realm, behaviors that are rarely invoked but widely present across cultures are usually considered behaviors of high relevance.[24] Humans use whispering in social contexts to limit exchanges to those in their immediate area to signal emotional closeness[23]; to initiate playful interaction, maintain secrecy, avoid disturbing others, and induce curiosity in listeners[24]; and to express affective states such as fear or aggression.[25] Whispering is also considered to be the appropriate form of communication in when voiced speech would be disruptive to others, such as in a movie theater.[26] Determining whether it is possible to whisper using TE speech, and subsequently the mechanism by which it occurs and means of training the skill, could have communicative value for the person.

The purpose of this study was to evaluate differences between whispered and spoken sentences produced by individuals using TE speech compared with individuals using laryngeal speech. Whispered sentences relative to spoken sentences were assessed via auditory-perceptual judgment, visual-perceptual judgment of spectrograms, and duration and dB measures of the acoustic waveform.

## MATERIALS AND METHODS

### Participants

Seventeen individuals with a laryngectomy using TE speech as their primary communication method participated. Table 1 includes demographic and medical information on these individuals who were recruited from the community at large. Inclusion criteria were (1) 18 years of age or older, (2) self-report of TE speech as their primary mode of communication, and (3) self-reported functional hearing in day-to-day interactions (aided or unaided). Exclusion criteria were (1) additional diagnoses known or expected to impact verbal communication such as stroke, head injury, or neurological disease, (2) surgical extension that included the oral tongue, oropharynx, velopharynx, or esophagus, and (3) evidence of hyper- or hypotonicity in the PE segment. Evidence for hyper- or hypotonicity was derived from one or more of the following: auditory-perceptual screening by the investigator during initial participant recruitment and continuing during the consent discussion; endoscopy or fluoroscopy data (or reports thereof) if the patient was able to provide that; history of Botox injections to help the TE voice; observation

**TABLE 1.**
**Demographic and Medical Information for Participants Using Tracheoesophageal Speech**

| SS | Age | Sex | Years Post Surgery | Extent of Surgery | RND | Radiation Therapy | Myotomy | TE Puncture Timing | Prosthesis Type |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 57 | M | 7 | Recon | N | Primary | N | Primary | BS Indwelling |
| 2 | 63 | M | 8 | Standard | Y | Primary | N | Primary | BS Indwelling |
| 3 | 54 | M | 1 | Standard | Y | Primary | N | Primary | Provox Vega |
| 4 | 69 | M | 2 | Standard | Y | Postop | Y | Primary | BS Low Pressure |
| 5 | 65 | M | 4 | Standard | N | Postop | Y | Secondary | Provox Vega |
| 6 | 64 | M | 3 | Recon | N | Primary | N | Primary | Provox NID |
| 7 | 58 | M | 6 | Recon | Y | Postop | N | Primary | Provox Vega |
| 8 | 64 | M | 5 | Standard | N | Postop | N | Secondary | Provox NID |
| 9 | 57 | M | 2 | Recon | Y | Primary | N | Secondary | Provox Vega |
| 10 | 65 | M | 8 | Standard | Y | Primary | N | Secondary | BS Low Pressure |
| 11 | 48 | M | 6 | Standard | Y | Postop | Y | Secondary | BS Indwelling |
| 12 | 60 | F | 3 | Standard | Y | Postop | Y | Primary | BS Low Pressure |
| 13 | 56 | F | 5 | Recon | N | Primary | N | Secondary | BS Indwelling |
| 14 | 63 | F | 7 | Standard | N | Postop | N | Secondary | BS Indwelling |
| 15 | 66 | F | 4 | Standard | Y | Postop | N | Primary | BS Low Pressure |
| 16 | 60 | F | 1 | Standard | Y | Primary | N | Secondary | Provox NID |
| 17 | 59 | F | 8 | Recon | N | Primary | Y | Secondary | Provox NID |

BS, Blom-Singer; Recon, reconstruction; RND, radical neck dissection; SS, TE participant.