Voice quality and surgical detail in post-laryngectomy tracheoesophageal speakers

Jacobi, I.; Timmermans, A.J.; Hilgers, F.J.M.; van den Brekel, M.W.M.

DOI
10.1007/s00405-015-3777-4

Publication date
2016

Document Version
Final published version

Published in
European Archives of Oto-Rhino-Laryngology

Eur Arch Otorhinolaryngol
DOI 10.1007/s00405-015-3777-4

HEAD AND NECK

Voice quality and surgical detail in post-laryngectomy tracheoesophageal speakers

I. Jacobi[1] · A. J. Timmermans[1] · F. J. M. Hilgers[1,2,3] · M. W. M. van den Brekel[1,2,3]

Received: 30 June 2015 / Accepted: 2 September 2015

Abstract The objective of this study is to assess surgical parameters correlating with voice quality after total laryngectomy (TL) by relating voice and speech outcomes of TL speakers to surgical details. Seventy-six tracheoesophageal patients' voice recordings of running speech and sustained vowel were assessed in terms of voice characteristics. Measurements were related to data retrieved from surgical reports and patient records. In standard TL (sTL), harmonics-to-noise ratio was more favorable after primary TL + postoperative RT than after salvage TL. Pause/breathing time increased when RT preceded TL, after extensive base of tongue resection, and after neck dissections. Fundamental frequency (f0) measures were better after neurectomy. Females showed higher minimum f0 and higher second formants. While voice quality differed widely after sTL, gastric pull-ups and non-circumferential pharyngeal reconstructions using (myo-)cutaneous flaps scored worst in voice and speech measures and the two tubed free flaps best. Formant/resonance measures in/a/ indicated differences in pharyngeal lumen properties and cranio-caudal place of the neoglottic bar between pharyngeal reconstructions, and indicate that narrower pharynges and/or more superiorly located neoglottic bars bring with them favorable voice quality. Ranges in functional outcome after TL in the present data, and the effects of treatment and surgical variables such as radiotherapy, neurectomy, neck dissection, and differences between partial or circumferential reconstructions on different aspects of voice and speech underline the importance of these variables for voice quality. Using running speech, next to sustained/a/, renders more reliable results. More balanced data, and better detail in surgical reporting will improve our knowledge on voice quality after TL.

Keywords Head and neck cancer · Total laryngectomy · Voice · Speech

Introduction

Tracheoesophageal (TE) speech presently is the preferred method of restoring oral communication after total laryngectomy (TL), because in many aspects it considerably outperforms esophageal (E) and electrolarynx (EL) speech [1]. A major advantage over E speech is the speed of acquisition and the close to normal phonation time and speaking rate of TE speech, whereas its wider intonation/pitch variability and speech intelligibility also outperform EL speech [1]. Nevertheless, TL still has a major impact on speech, swallowing, and psychosocial well-being [2–4]. For TE speech, significant correlations were found between voice quality and quality of life measures, fatigue, sentence duration, anxiety to speak, and the frequency of making telephone calls. Female patients exhibit a greater voice handicap and significantly lower quality of life scores than males [4–6]. In studies on the relationship between acoustic measures and patient-experienced voice quality, self-assessed voice-related quality of life correlated with

I. Jacobi and A. J. Timmermans contributed equally.

✉ M. W. M. van den Brekel
m.vd.brekel@nki.nl

1 Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

2 Institute of Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands

3 Department of Oral and Maxillofacial Surgery, Academic Medical Center, Amsterdam, The Netherlands

Published online: 22 September 2015

acoustic measures of intensity and temporal aspects of speaking (pauses, duration) [7, 8]. Studies that matched professional listeners' judgments of voice quality of sustained/a/with acoustics or visual signal typing of the spectrogram pointed out the importance of harmonics and formants, also in higher spectral regions [9, 10].

Voice quality and speaking effort differ widely within the TE population [5, 8, 11]. The tonicity of the pharyngoesophageal (PE) segment (also called neoglottis), and therewith voice quality, is based on the adaptation and vibration dynamics of the pharyngeal mucosa [12]. Dependent on the individual anatomy, the surgical procedures performed and possibly radiotherapy, variation occurs in muscular control, position and length of the vibrating segment, and mass and stiffness of the PE segment. Each of these characteristics can affect voice (and swallowing) function.

In comparison to the quasi-symmetric vocal folds, the vibrating neoglottis consists of amorphic vibrating elements in the wall of the PE segment. The whole vibrating segment is in general larger (more mass) and neurologically less controllable than the vocal folds are. Furthermore, in view of the fact that air pressure control is needed to initiate and extend vibration, it seems a 'drawback' that the PE segment below and at the neoglottic region is expandable, while the (sub)glottic larynx and trachea are stabilized through their cartilage framework. After TL, the laryngeal differences between the sexes are lost and the limited neurological control, the myo-elastic properties, mass, size, and diameter of the neoglottis and its surrounding tissues bring about a lower frequency and more irregular voice, decreased dynamic range, and less aerodynamic voice and f0 control [13–16].

Although post-TL voice quality and control are known to differ substantially between patients, studies discussing the morpho-physiology and surgical characteristics and their (interacting) effects on post-TL functioning are still sparse. In the literature, various variables were found to affect functional outcomes. Among these, besides the extent of the resection, are the surgical method of pharynx closure and reconstruction (muscle closing techniques, donor site tissue properties), the conservation of the posterior pharyngeal wall, the degree and level of neoglottic closure during phonation (presence and place of the neoglottic bar and distance and intensity of contact between posterior and anterior wall), the pressure built up below the neoglottic bar during phonation (intraluminal pressure), the diameter of the pharynx (pharyngeal and esophageal volume and extension), previous or post-operative (chemo-)radiotherapy, and (the extent of) neck dissections [14, 16–28]. Although the extent of the surgical resection is primarily dictated by tumor extent, surgical techniques, such as neurectomy and upper esophageal myotomy, and the technique of pharynx (muscle) closure and type of reconstruction thus seem to be important phonosurgical aspects of TL.

In this retrospective study, we aim to identify patient and surgical parameters that correlate with voice quality after TL. Therefore, voice and speech outcomes of TL speakers were assessed and related to their patient and surgical details.

Methods

Patient data

Over time, voice and speech recordings of 86 TE speakers of sufficient quality and extent were collected. For 76 of these, sufficient surgical detail could be retrieved for meaningful analyses. The vast majority of the 76 patients participated between 2007 and 2014 in voice prostheses studies at the Netherlands Cancer Institute (NKI). The included patients underwent TL between 1983 and 2013. Almost all patients underwent TL at the NKI. In 13 cases, with the patient's informed consent, surgical reports were retrieved from other Dutch hospitals. At the NKI, common procedures during TL include the creation of a separate tracheostoma in the inferior skin flap, sectioning of the sternal head of the SCM muscles to obtain a flatter stoma area, a short myotomy of the upper esophageal sphincter, T-shaped closure of the pharyngeal defect, and a primary TE puncture (TEP) with direct fit of the voice prosthesis (VP) [29]. Between 1990 and 2002, unilateral neurectomy of the pharyngeal plexus was performed on a regular basis.

Medical records were assessed for demographic and surgical information, e.g., sex, age, site and TNM classification of the tumor, indication for TL, and prior and postoperative treatments in the head and neck area, staff surgeon performing the TL, surgical extent (indicated by surgical details such as the level of the trachea resection, base of tongue resection and tumor location), remaining pharyngeal mucosa and reconstruction procedure, pharyngeal mucosa and muscle closure technique, neck dissection, myotomy of the upper esophagus, and plexus pharyngeus neurectomy. Since mass properties play a large role in voicing, as a general indicator of tissue properties, body mass index (BMI) was assessed. The later VP lengths were used as a possible surrogate marker for the tracheoesophageal wall thickness and its potential effect on voicing.

Outcome measure to assess voice and speech

Voice and speech recordings included a read aloud standard Dutch text (length: 151 words) and sustained/a/. The read aloud text allows analysis of speech (including f0, %voiced and %unvoiced speech) and no-speech (%pause and/or breathing time). The chosen outcome measures are based on the available literature on voice quality, and have importance for laryngectomized patients in view of perceptual and experienced voice quality.

1. Speaking fundamental frequency (f0): we assessed f0 in terms of its minimum, maximum, mean/median, and range across the read aloud text. The mean indicates the height of he normal speaking frequency; the range of f0 indicates the possibility to produce variation in intonation.
2. %Voiced and %unvoiced speech: TE speech is generally less voiced than laryngeal speech; higher amounts of %voiced indicate better TE speech.
3. %Pause/breathing time (calculated from the ratio of pause and breathing time to the overall reading time [7]: %pause/breathing time was assessed to measure the ease of reading a whole piece of text. Comparable to laryngeal speakers, TE speakers tend to pause/breathe at phrase ends indicated by the text (e.g., at comma's, points). Therefore, the total percentage of pause was assessed and not, e.g., the number.
4. Band energy difference (BED; between 0–500 and 4000–5000 Hz, according to van As-Brooks et al. [10]; dB) in sustained/a/as a measure for spectral tilt (further referred to as "spectral tilt"). The tilt of the spectrum is related to the pressure build-up and the force of (neo-)glottic closure during vibration; it indicates effort. Spectral tilt has been shown to correlate with perceived voice quality in laryngectomized patients [30].
5. Harmonics-to-noise ratio (in dB; HNR) in sustained/a/, indicating the degree of acoustic periodicity. HNR has been found to correlate with perceived voice quality, pleasantness, and intelligibility [9, 10].
6. Presence of formants and harmonics in the area of the third and fourth formant (HNR F3/F4); in sustained/a/ [9, 10], which were found to correlate with intelligibility and voice quality.
7. The position of and the distance between the first two formant*s* (spectral peaks) were assessed in sustained/a/ as indicator for differences in the vocal tract resonance cavities (lumen). Higher formants indicate shorter vocal tracts (less distance from voice source to lips). The formant distance was found to correlate with intelligibility and pleasantness [9].

All analyses were performed with Praat [31].

Statistics

Descriptive statistics were performed. For pairwise comparisons between groups on the continuous acoustic variables we used Mann–Whitney *U* tests, for three samples the Kruskal–Wallis (alpha = 0.05/3) followed by Mann–Whitney *U* tests. Pearson's correlation was applied for relationships between voice measures. Spearman's correlation was applied for relationships between prosthesis size and acoustic measures or age. The Fisher's exact test was applied between categorical variables. Analyses were performed with IBM SPSS Statistics 22.0.

Results

Patient and surgical characteristics

The 76 included patients underwent a TL for the following indications: 24 as primary cancer treatment, and 52 after RT ($N = 47$) or chemoradiation (CCRT; $N = 5$). Of these latter 52, 38 concerned salvage treatment, 5 a second primary, and 9 a dysfunctional larynx after prior (chemo-) radiotherapy. Of the 24 primary TL patients, 16 underwent postoperative RT and 3 postoperative CCRT. Five patients were not irradiated at all. The patients' demographic, medical and surgical characteristics are shown in Table 1.

The majority of patients had a glottic or supraglottic carcinoma with tumor-negative lymph nodes. Mean age at TL was 59 years. Simultaneously with the TL, 39 patients (52 %) underwent a neck dissection (ND), including at least levels 2–4. The vast majority underwent a primary TEP (90 %) with insertion of a Provox2 VP. The most common VP length at primary TEP was 8 mm, and during follow-up 6 or 8 mm (36 and 30 %, respectively).

Forty-seven patients were treated with standard TL (sTL) without reconstruction (62 %), and 35 (75 %) underwent short myotomy of the upper esophageal musculature. Fifteen of the sTL patients (32 %) had a neurectomy of the pharyngeal plexus. Only in 8 patients (10 %), a significant part of the base of the tongue was resected. In 30 patients, the trachea was transected above the third ring (64 %).

Besides the 47 patients with a sTL, 10 underwent a sTL with PM-muscle flap reinforcement of the pharyngeal suture line or closure of a pharyngocutaneous fistula. Pharyngeal closure techniques, such as layers and to what extent the constrictor pharyngeal muscle remnants were sutured together, were not consistently reported in detail. Nineteen patients underwent an extended pharyngeal resection requiring reconstruction of the lumen. There were two circumferential reconstructions with a free flap, four with a tubed gastric pull-up, and one with a full gastric pull-up. There were twelve partial reconstructions of the pharyngeal wall; one with a free radial forearm flap and eleven with a PM myocutaneous flap. Sixteen different

Table 1 Patient and surgical characteristics; * in "Extent of TL" refers to details in "Lumen reconstruction"

	N (%)
Total	76 (100)
Male	66 (90)
Mean age at TL (range)	59 (29–85)
Mean age at recording (range)	66 (42–88)
Primary site	
Supraglottic	20 (26)
Glottic	37 (49)
Subglottic	4 (5)
Oropharynx	4 (5)
Hypopharynx	9 (12)
Proximal esophagus	2 (3)
T classification (initial)	
Tis-T1	20 (26)
T2	17 (22)
T3	11 (15)
T4	28 (37)
N classification (initial)	
N0	54 (71)
N+	22 (29)
Indication TL	
Primary	24 (32)
Salvage	38 (50)
Second primary	5 (7)
Dysfunctional larynx	9 (12)
BMI at TLE	
<18	5 (7)
18–25	33 (43)
25–30	28 (37)
>30	5 (7)
Unknown (TL elsewhere)	5 (7)
TL timing in relation to RT	
TL for RT failure	44 (58)
TL for CCRT failure	5 (7)
TL for RT failure + postoperative RT	3 (4)
TL + postoperative RT	16 (21)
TL + postoperative CCRT	3 (4)
No RT	5 (7)
Extent of TL	
Standard TL	47 (62)
Standard TL with PM flap reinforcement	10 (13)
Extended TL with non-circumferential resection*	12 (16)
Extended TL with circumferential pharyngeal resection*	7 (9)
Lumen reconstruction*	
PM myocutaneous flap	11 (58)
Free flap	3 (16)
Gastric pull-up	5 (26)
Neck dissection ≥level 2–4	
No	37 (49)
Unilateral	18 (24)

Table 1 continued

	N (%)
Bilateral	21 (28)
Extended resection base of tongue	
No	68 (90)
Yes	8 (10)
Neurectomy (standard TL, $N = 47$)	
Yes	15 (32)
No	32 (68)
Short myotomy (standard TL, $N = 47$)	
Yes	35 (75)
No	12 (25)
Neurectomy (standard TL, including PM flap for reinforcement, $N = 57$)	
Yes	16 (28)
No	39 (68)
Unknown	2 (4)
Short myotomy (standard TL, including PM flap for reinforcement, $N = 57$)	
Yes	42 (74)
No	13 (23)
Unknown	2 (4)
Transection trachea (standard TL, $N = 47$)	
<3rd ring	17 (36)
>3rd ring	30 (64)
Tracheoesophageal puncture (TEP)	
Primary TEP	69 (91)
Secondary TEP	7 (9)

surgeons/surgical teams were involved in the surgeries in this patient cohort.

Voice and speech outcomes

Standard TL

There were several interactions between various treatment parameters: statistical analyses confirmed differences in the frequency of ND and pre- or post- (C)RT treatment (Fisher's exact test, $p = 0.005$) as well as the frequency of neurectomy ($p = 0.027$). Patients with a bilateral ND were predominantly treated by postoperative (C)RT (60 %) and all but one underwent neurectomy. Patients with a unilateral ND were predominantly treated by (C)RT before TL (62 %), and had no neurectomy (67 %). In patients without ND, most had (C)RT prior to TL (83 %) and only half of them underwent neurectomy. Next to differences in subcategory frequencies, these correlations between treatment groups limited an analysis of relationships with voice.

Within the sTL speakers ($N = 47$), speaking f0 and the percentage of voicedness in the read aloud text were significantly correlated ($r = 0.666$, $p < 0.001$): the higher f0, the more voicing during speech. The voicedness in the read aloud text correlated significantly with the harmonics-to-noise ratio (HNR) measured in sustained/a/: the more voicing in the text, the better the harmonics-to-noise ratio, ($r = 0.492$, $p < 0.001$,). Spectral tilt (BED) correlated marginally with the f0-range in the text ($r = 0.297$, $p = 0.045$): the higher the range of f0 in the text, the more tilt in the spectrum of sustained/a/. Speaking fundamental frequency correlated significantly with other f0-measures: the higher the speaking mean, the higher the minimum, maximum, and range were ($r = 0.665$, maximum $r = 0.874$, range $r = 0.705$, $p < 0.001$). There was no significant correlation with pause/breathing time.

Speakers' sex The median speaking f0 across the read aloud text was lower for males (median 108 Hz, ranging from 49 to 238 Hz) than for females (140 Hz, range 33–277 Hz). The minimal f0 in running speech was higher for females (71 Hz vs. 41 Hz, $p = 0.009$, $U = 40.0$). All other differences in voice outcome between male ($N = 40$) and female ($N = 7$) speakers were insignificant.

Age, BMI and thickness of the tracheoesophageal wall There was a significant decline in prostheses lengths in older speakers ($p = 0.018$, $r = -0.343$, $N = 47$), but there were no associations between voice measures and age, prosthesis length (tracheoesophageal party wall thickness), and/or BMI.

Pharyngeal closure The variety of surgical teams ($N = 16$) and their muscle closure techniques as well as underspecified procedure descriptions precluded evaluations of an effect of pharyngeal closure technique or the extent of the residual pharyngeal wall.

Short myotomy and neurectomy Most patients in the sTL group underwent a short myotomy (35/47) and had no neurectomy (32/47). There were no significant main effects of myotomy on voice measures, only a trend towards increased pause/breathing time during speech in TE speakers with myotomy compared to those without myotomy (24 vs. 20 %, $p = 0.057$, $U = 132.0$). TL speakers with neurectomy showed significantly more voicing during speech (55 vs. 33 %, $p = 0.035$, $U = 129.5$), and higher speaking f0 (122 vs. 103 Hz, $p = 0.058$, $U = 157.0$).

(Chemo-)radiotherapy In view of the subsample sizes, for effects of (C)RT [salvage surgery after (C)RT failure] vs. primary TL with postoperative (C)RT, we excluded the smallest subgroups: three patients with no history of (C)RT, and one patient with pre- and post (C)RT, leaving a subset of 43 patients. While the harmonics-to-noise ratio was significantly better after primary TL with postoperative (C)RT when compared to those patients receiving preoperative (C)RT ($p = 0.042$, $U = 118.0$), pause/breathing time during the reading task was longer (21 vs. 24 %, $p = 0.048$, $U = 120.5$).

Neck dissection of at least level 2 to 4 Neck dissection (no/unilateral/bilateral) had a significant effect on the percentage of pause/breathing time in running speech, and this effect increased from no- to uni- to bilateral ND [20 vs. 24 vs. 29 %, $\chi^2(2) = 8216$, $p = 0.016$]. Post-hoc tests showed significantly higher percentage of pause/breathing time after bilateral ND than without ND ($p = 0.004$, $U = 77.5$).

Extensive base of tongue resection Significantly lower first formants were found after extensive tongue resection (589 vs. 656 Hz, $p = 0.033$, $U = 69.5$) and higher percentage of pause/breathing time during the reading task (25 vs. 21 %, $p = 0.039$, $U = 71.5$). Extensive base of tongue resections were equally distributed in terms of ND and timing of (C)RT, the factors found to affect pause/breathing time as well.

Standard TL with PM flap for reinforcement

In sTL group with PM flap reconstructions for reinforcement (10 patients), speech and voice measures were comparable to the sTL group without such additional flap (Table 2; Figs. 1, 2, 3). In the whole sTL group including the 10 reinforcement flaps ($N = 57$) analysis of the influence of myotomy, neurectomy, neck dissection, RT, and base of tongue resection rendered similar results as in the sTL group without PM flap reinforcement. Several effects were even more evident in this larger group: again, females (9/57) had a higher minimum f0 and significantly higher second formants (F2 1452 vs. 1398 Hz, $p = 0.030$, $U = 117.0$). In this larger group, next to an effect of age on tracheoesophageal party wall thickness, prosthesis length tended to be shorter when (C)RT preceded TL [χ^2 (1) = 5610, $p = 0.018$]. After neurectomy, next to more voicing during the running speech task, the median (131 vs. 101 Hz, $U = 201.0$, $p = 0.040$), maximum (204 vs. 172 Hz, $U = 191.5$, $p = 0.033$) and range (162 vs. 122 Hz, $U = 197.0$, $p = 0.043$) of the speaking f0 were significantly higher.

Extended TL with reconstruction

Several voice measures differed significantly between sTL speakers ($N = 57$) and the group with TL plus (near-)total pharyngectomy with lumen reconstruction ($N = 19$). Three speakers showed less than 5 % of voicing during the running speech task, including two of the five total pharyngectomies with a gastric pull-up (one full and one tubed). These three speakers were excluded from f0 analyses.

In running speech, compared to the sTL group, for this reconstruction group, we found significantly lower values for: the median f0 (67 vs. 117 Hz, $p < 0.001$, $U = 195.5$), the lowest and highest f0 (22 vs. 39 Hz, $p = 0.011$, $U = 281.0$ and 114 vs. 179 Hz, $p < 0.001$, $U = 183.5$), the f0 range (81 vs. 136 Hz, $p < 0.001$, $U = 205.0$), and voicedness (18 vs. 37 %, $p < 0.001$, $U = 245.0$) (see Fig. 1 (including laryngeal f0-means according to Traunmüller and Eriksson [32]). Moreover, in this group, the 2nd formant of/a/was slightly higher (1426 vs. 1400 Hz, $p = 0.048$, $U = 376.5$). There were no significant differences in measures of harmonics-to-noise ratios and spectral tilt.

Sample sizes in the pharyngeal reconstruction subgroups were small. However, there was a clear trend in measures across running speech as well as in sustained/a/: overall, there was a trend for two tubed free flap reconstructions to show more favorable results of f0 (median, maximum, minimum and range), as well as voicedness across the read aloud text (Table 2; Fig. 1), pause/breathing time and spectral tilt (Figs. 1, 3), harmonics-to-noise ratio (Fig. 3),

Table 2 Mean and standard deviation of acoustic voice measures per subgroup

	Standard TL		Standard TL + PM for reinforcement		Partial reconstruction (PM/RFF)		Tubed free flap		Gastric pull-up ± PM		Full gastric pull-up + PM	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Text												
f0 median	119	50	134	73	61	20	135	8	67	39	–	–
f0 min	48	32	59	36	26	11	38	0	43	19	–	–
f0 max	197	72	206	94	93	35	225	23	149	32	–	–
f0 range	148	59	147	71	67	23	187	23	106	19	–	–
% voiced	40	21	38	18	23	12	41	18	8	6	4	–
% pause	23	9	25	11	24	6	17	4	28	6	20	–
% unvoiced	36	21	37	20	53	18	43	14	65	11	79	–
/a/												
HNR	1.8	3.9	0	2.4	−0.1	1.5	4.4	2.5	−1.2	1.4	−3.7	–
HNRF3/4	78	16	82	14	81	12	101	0	78	26	52	–
Spectral tilt	−16	8	−17	8	−12	8	−31	6	11	10	0.4	–
F1 Hz	644	39	636	37	657	31	592	88	668	44	672	–
F2 Hz	1393	68	1427	76	1440	86	1498	27	1433	102	1367	–
F3 Hz	2853	178	2896	205	2908	252	2481	222	2845	87	2924	–

Two of five of the gastric pull-ups had to be excluded from the f0 analyses due to the absence of voicing during the reading task

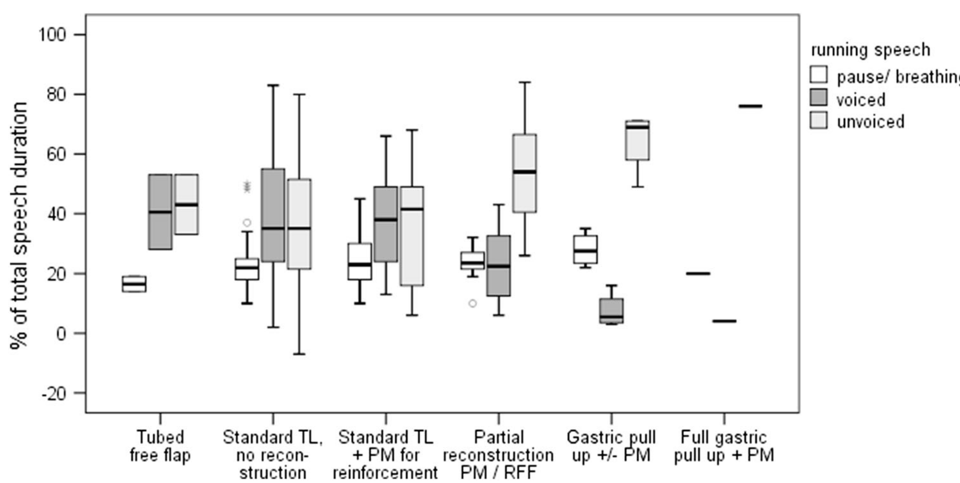

Fig. 1 Amount of voiced and unvoiced speech, and pause/breathing time as a percentage of the duration of the running speech task. Percentage voiced speech as *box-plots* (representing the median, 95 % CI and interquartile range). Most favorable groups/outcomes to the *left*, least normal to the *right*

as well as the harmonics-to-noise ratio in the region of the 3rd and 4th formant (not plotted). Speakers with PM myocutaneous flap reconstruction and those with gastric pull-up reconstruction showed the least favorable outcome (depicted in Figs. 1, 2, 3). The two tubed free flap reconstruction were comparable to or even better than the best sTL voices. In comparison to the other TL speakers, vocal tract resonance cavity measures (formants) showed rather high 2nd formant and low 1st formant values (large F2–F1 distance) for these two patients (Table 2).

Discussion

The aim of the present retrospective study is to identify surgical parameters that correlate with voice quality after TL by relating voice and speech outcomes of TL speakers to details of their surgical procedure.

In the sTL speakers, there are several interesting correlations: higher speaking f0 correlates with more voicedness, and voicedness correlates with a better harmonics-to-noise ratio. Moreover, in running speech, there is a trend

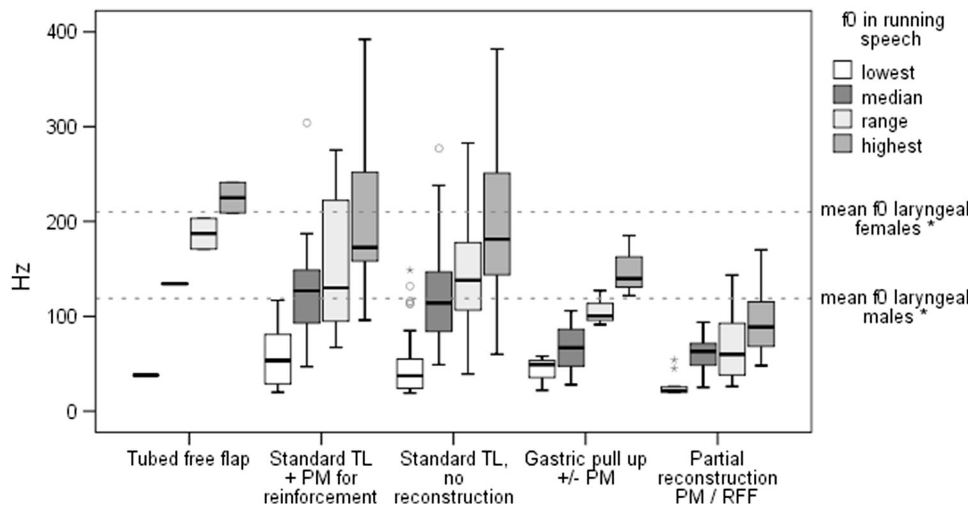

Fig. 2 *Boxplots* of lowest, median, and highest fundamental frequency (f0) during the running speech task. Three patients had to be excluded from the f0 analyses due to less than 5 % of voicing during the running speech task. The broken horizontal lines depict mean f0 values for groups of laryngeal female and male speakers according to Traunmüller and Eriksson [32]. From *left to right:* most to least favorable speaking frequency and range

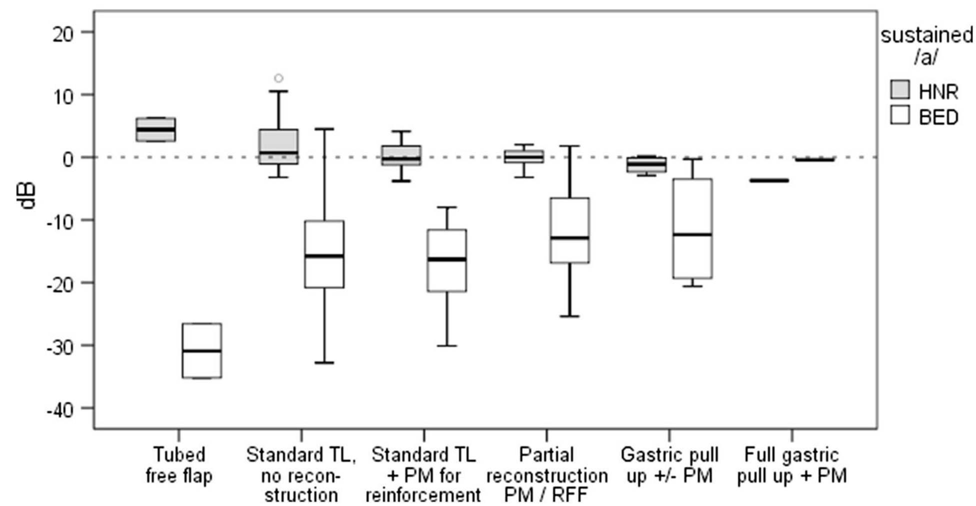

Fig. 3 *Boxplots* of harmonics-to-noise ratio (HNR) and spectral tilt, both from sustained/a/. To the *left*: most favorable (signal type I; [10]), to the *right*: least favorable (signal type IV; [10])

towards a higher f0 in females, with significant higher minimum f0 in running speech. Females also differed in formants (vocal tract cavities) with higher second formants. These f0- and formant differences might be due to gender-dependent behavior, with females trying to produce a higher pitch (by using more tension and changing the height of the neoglottis). Yet, differences between the sexes in the average vocal tract/esophageal lumen size can not be ruled out.

Other studies did not find f0 differences between the sexes, when assessing f0 in sustained/a/[6]. In sustained/a/, however, f0 is not necessarily representative for normal speaking f0.

One of the effects of surgical detail present in this cohort of sTL speakers is that pause/breathing time increases from no- to uni- to bilateral neck dissections, and is more pronounced after extensive tongue resection, and when postoperative (C)RT is used. Another interesting finding is that patients with primary TL and postoperative (C)RT showed better harmonics-to-noise ratios than patient who had a salvage laryngectomy after prior (C)RT, suggesting that the ‘condition’ of the PE segment is more favorable for voicing after primary TL than after salvage surgery.

Next to longer pause and breathing time, TL speakers with extensive base of tongue resection presented with lower first formants in/a/, possibly the result of ratio changes in the front vs. back cavity and compensatory strategies to still reach the perceptual impression of/a/. The base of tongue plays a major role in speech, as is the case for swallowing. After TL, swallowing deficiencies, especially with solid food, are reported regularly [4, 23, 33, 34]. This dysphagia usually presents in pharyngeal clearance problems and prolonged (oro)pharyngeal transit times, which are the result of both the decreased control of the base of tongue and the pharyngeal wall [35, 36].

To prevent spasm of the neoglottis, most of our sTL speakers underwent a short myotomy of the upper esophageal segment [37]. The need for this short myotomy to prevent spasm is controversial in literature, and several (early) studies supported pharyngeal neurectomy [19, 28, 38, 39], with good PE pressure and higher voices, presumably by the maintenance of some residual pressure in the neoglottis as a function of the contralateral plexus. This was reason why between 1990 and 2002, neurectomy was favored. However, in view of the favorable effects of a short myotomy of the upper esophageal sphincter in one of our studies [37] neurectomy was abandoned. Interestingly, we now find that, in line with literature, speaking pitch is higher and f0 measures are more favorable after neurectomy.

Due to the various details in surgical reporting, the limited number of patients per subgroup, selection biases, and the diverse surgical teams over the 30 years the surgical data were collected, an analysis ofpharyngeal closure technique on voice and speech outcome was not possible retrospectively. Literature on speech failure, intraluminal pressure, fistulae, and swallowing, however, clearly favors muscle closure over non-muscles closure [19, 27, 40, 41].

Our results confirm previous studies in showing worse functional scores for speakers with (partially) reconstructed pharynges [42], although the voice outcome was more favorable for the two tubed free radial forearm flaps (RFF). This could be a selection bias. The worse voice quality after circumferential reconstruction coincides with the literature [25]. Yet, other than in our study, after partial pharyngeal reconstruction, voice (and swallowing) was reported to be comparable to standard TL [25]. Including electro-larynx speakers, a different subset of donor sites, and other voice quality scoring in that study, however, a direct comparison of voice outcome is difficult.

In pharyngeal reconstructions, (full) gastric pull-ups and non-circumferential reconstructions scored worst across running speech as well as in sustained/a/measures. In three of the five (tubed) gastric pull-up patients a measurable f0 was present (however, minimally voicing), which suggests an occasionally sufficient diameter and closure of its lumen. The other two had no voicing (f0).

For sTL, voice measures showed no significant effect of age, BMI or TEP length. With the TEP being well below the PE segment, it is a too imprecise indicator of pharyngeal wall thickness. Nonetheless, in line with previous studies on pharyngeal wall or tissue thickness, TEP length showed a significant effect of age. The fact that age, in contrast to earlier findings [43], is not correlated with any of the acoustic voice parameters studied, might be a selection bias, since the better/fitter patients are probably overrepresented in this cohort of voice recordings. More precise measures to assess tissue properties might lead to a better interpretation of the role of (changes) in tissue properties in TE voice and speech quality.

Next to differences in voice quality, formants (resonance cavity measures) in/a/indicate differences in pharyngeal lumen properties and cranio-caudal place of the neoglottic bar between pharyngeal reconstruction procedures. While the vocal folds in normal adult laryngeal speakers are at the height of C5–C6, imaging of TL speakers during phonation suggests that the neoglottic bar is located higher than the vocal folds (middle of C3–C5) [6, 42]. Roughly speaking, formant frequencies are inversely proportional to the vocal tract length (the ratio of pharynx length to mouth length) with small formant dispersion (F2–F1 distance) indicating larger body size and shape [44]. According to the effect of shortening of the vocal tract (from lips to the neoglottic bar) after laryngectomy, overall, higher formant values were found in TL speakers than in laryngeal speakers [45]. Whereas in sTL speakers the neoglottic bar is usually around the level of C4, this level was found to differ widely in TL speakers with pharyngeal reconstructions (C3–C7) [42]. In our dataset, the second formant and the formant dispersion were highest in tubed RFF, and lowest in full gastric pull-up, explainable by different locations of the neoglottic bar caused by different lumen diameters and tissue characteristics of the reconstruction. Overall, the voice outcome and formants in our data suggest that smaller diameter pharynges and/or more superiorly located neoglottic bars are associated with favorable voice quality and more effortless speech. To compensate for a wide pharynx, external pressure (e.g., by PM flap) might be useful to compensate for a low tonicity. These findings are confirmed in previous studies using videofluoroscopy in which it has been shown that smaller pharyngeal diameters and optimization of the intraluminal pressures favor voicing [14, 16, 24, 37, 42, 46].

As shown on videofluoroscopy, for vibration after TL, pulmonary airstream is sent through the PE segment, and the walls are pushed up until the walls form a neoglottic bar (pharyngeal closure) leading to a Bernoulli-effect, and the walls starts to vibrate. Positioning and muscular control of the vibratory segment play a significant role in f0 alterations, and in some patients there is a striking pharyngeal control, leading to a good control over loudness and dynamic range [47, 48]. Dynamic range of the TL voice has been reported to correlate with the contraction amplitude in the neoglottis [14], and air pressure with esophageal expansion [16]. Although ideally we would like to create the narrowest point at the optimal level with maximal muscular control, we do not have enough knowledge and data to determine the optimal creation of the pharynx and neoglottic bar. A very wide 'pharynx', as in full gastric pull-up has the worst outcome, but a too narrow pharynx, although not necessarily bad for voicing, can interfere with swallowing, and a too muscular pharynx can cause hypertonicity-related voicing problems.

Strength and limitations

Studies discussing the direct impact of surgical detail in TL procedures on functioning are sparse. Much of the literature regarding voice functioning focuses on sustained vowels, which ignores the patient perspective as Kazi et al. [6] pointed out. In contrast, running speech covers different aspects of voice and speech and consequently, next to sustained/a/, predominantly running speech was used in the present study. Although we were able to compare patients with various surgical procedures, the present retrospective study had limited power due to the unbalanced dataset, the limited numbers of patients in various subgroups and the impossibility to carry out meaningful multivariable analyses. In this study, patients of whom voice recordings were available, were mostly patients that participated in voice prosthesis studies and were thus possibly the “better/fitter” TL speakers. This might have caused an unbalanced dataset.

Also, the presumed effect of pharyngeal closure technique could not be assessed because surgical reports were not always clear. Prospective data collection and structured reporting of surgical detail is needed to draw more definitive conclusions.

For the future, imaging data during voicing and data on muscle activity during phonation would help to disentangle relationships between pharyngeal properties, vibrating mass, and surgical procedures, including muscle closure techniques, myotomy and neurectomy, and the role of speaker behavior in voice outcome after TL.

Conclusion

The ranges in voice outcome after TL are related to variables like radiotherapy, neurectomy, neck dissection, and reconstruction procedures. In this patient cohort gender/speaker behavior appears to have an influence on the f0 in running speech. Overall, our results suggest that narrower pharynges and/or more superiorly located neoglottic bars are associated with more favorable voice quality. Patients with pharyngeal lumen reconstructions (i.e., by PM myocutaneous flaps and (tubed) gastric pull-ups) have the poorest voices. In sTL, neurectomy may be favorable.

Acknowledgments The Netherlands Cancer Institute receives a research grant from Atos Medical Sweden, which contributes to the existing infrastructure for health-related quality of life research of the department of Head and Neck Oncology and Surgery.

Compliance with ethical standards

Conflict of interest The Netherlands Cancer Institute receives a research grant from Atos Medical Sweden, which contributes to the existing infrastructure for health-related quality of life research of the department of Head and Neck Oncology and Surgery.

References

1. Searl JP (2014) Reeves IS Chapter 9. Nonsurgical voice restoration following total laryngectomy. In: Ward EC, Van As-Brooks CJ (eds) Head and neck cancer: treatment, rehabilitation, and outcomes. 2nd edn. Plural Publishing, USA, pp 263–300
2. Oozeer NB, Owen S, Perez BZ, Jones G, Welch AR, Paleri V (2010) Functional status after total laryngectomy: cross-sectional survey of 79 laryngectomees using the Performance Status Scale for Head and Neck Cancer. J Laryngol Otol 124(4):412–416
3. Queija Ddos S, Portas JG, Dedivitis RA, Lehn CN, Barros AP (2009) Swallowing and quality of life after total laryngectomy and pharyngolaryngectomy. Braz J Otorhinolaryngol 75(4):556–564
4. Robertson SM, Yeo JC, Dunnet C, Young D, Mackenzie K (2012) Voice, swallowing, and quality of life after total laryngectomy: results of the west of Scotland laryngectomy audit. Head Neck 34(1):59–65
5. Ackerstaff AH, Hilgers FJ, Aaronson NK, Balm AJ (1994) Communication, functional disorders and lifestyle changes after total laryngectomy. Clin Otolaryngol Allied Sci 19(4):295–300
6. Kazi R, Kiverniti E, Prasad V, Venkitaraman R, Nutting CM, Clarke P et al (2006) Multidimensional assessment of female tracheoesophageal prosthetic speech. Clin Otolaryngol 31(6):511–517
7. Lundstrom E, Hammarberg B (2011) Speech and voice after laryngectomy: perceptual and acoustical analyses of tracheoesophageal speech related to voice handicap index. Folia Phoniatrica Logopaedica 63(2):98–108
8. Lundstrom E, Hammarberg B, Munck-Wikland E, Edsborg N (2008) The pharyngoesophageal segment in laryngectomees–videoradiographic, acoustic, and voice quality perceptual data. Logoped Phoniatr Vocol 33(3):115–125
9. D’Alatri L, Bussu F, Scarano E, Paludetti G, Marchese MR (2012) Objective and subjective assessment of tracheoesophageal prosthesis voice outcome. J Voice 26(5):607–613
10. van As-Brooks CJ, Koopmans-van Beinum FJ, Pols LC, Hilgers FJ (2006) Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech. J Voice 20(3):355–368
11. Schuster M, Toy H, Lohscheller J, Eysholdt U, Rosanowski F (2005) Quality of life and voice handicap of laryngectomees using tracheoesophageal substitute voice. Laryngorhinootologie 84(2):101–107
12. Muller-Miny H, Diederich S, Bongartz G, Peters PE (1991) Radiologic findings following supraglottic and total laryngectomy. Der Radiol 31(7):324–331
13. Fitch WT, Giedd J (1999) Morphology and development of the human vocal tract: a study using magnetic resonance imaging. J Acous Soc Am 106(3 Pt 1):1511–1522
14. Reis N, Aguiar-Ricz L, Dantas RO, Ricz HM (2013) Correlation of intraluminal esophageal pressure with the dynamic extension of tracheoesophageal voice in total laryngectomees. Acta cirurgica brasileira/Sociedade Brasileira para Desenvolvimento Pesquisa em Cirurgia 28(5):391–396

15. Deschler DG, Doherty ET, Reed CG, Singer MI (1999) Effects of sound pressure levels on fundamental frequency in tracheoesophageal speakers. Otolaryngol Head Neck Surg 121(1):23–26
16. Takeshita-Monaretti TK, Dantas RO, Ricz H, Aguiar-Ricz LN (2014) Correlation of maximum phonation time and vocal intensity with intraluminal esophageal and pharyngoesophageal pressure in total laryngectomees. Ann Otol Rhinol Laryngol 123(11):811–816
17. Aguiar-Ricz L, Ricz H, de Mello-Filho FV, Perdona GC, Dantas RO (2010) Intraluminal esophageal pressures in speaking laryngectomees. Ann Otol Rhinol Laryngol 119(11):729–735
18. Iwai H, Tsuji H, Tachikawa T, Inoue T, Izumikawa M, Yamamichi K et al (2002) Neoglottic formation from posterior pharyngeal wall conserved in surgery for hypopharyngeal cancer. Auris Nasus Larynx 29(2):153–157
19. Brok HA, Stroeve RJ, Copper MP, Schouwenburg PF (1998) The treatment of hypertonicity of the pharyngo-oesophageal segment after laryngectomy. Clin Otolaryngol Allied Sci 23(4):302–307
20. Albirmawy OA, Elsheikh MN, Silver CE, Rinaldo A, Ferlito A (2012) Contemporary review: impact of primary neopharyngoplasty on acoustic characteristics of alaryngeal tracheoesophageal voice. Laryngoscope 122(2):299–306
21. McIvor J, Evans PF, Perry A, Cheesman AD (1990) Radiological assessment of post laryngectomy speech. Clin Radiol 41(5):312–316
22. Blom ED, Singer MI, Hamaker RC (1986) A prospective study of tracheoesophageal speech. Arch Otolaryngol Head Neck Surg 112(4):440–447
23. de Casso C, Slevin NJ, Homer JJ (2008) The impact of radiotherapy on swallowing and speech in patients who undergo total laryngectomy. Otolaryngol Head Neck Surg 139(6):792–797
24. Fouquet ML, Goncalves AJ, Behlau M (2009) Relation between videofluoroscopy of the esophagus and the quality of esophageal speech. Folia Phoniatrica Logopaedica 61(1):29–36
25. Gadepalli C, de Casso C, Silva S, Loughran S, Homer JJ (2012) Functional results of pharyngo-laryngectomy and total laryngectomy: a comparison. J Laryngol Otol 126(1):52–57
26. Kazi R, Singh A, Mullan GP, Venkitaraman R, Nutting CM, Clarke P et al (2006) Can objective parameters derived from video fluoroscopic assessment of post-laryngectomy valved speech replace current subjective measures? An e-tool-based analysis. Clin Otolaryngol 31(6):518–524
27. Maclean J, Szczesniak M, Cotton S, Cook I, Perry A (2011) Impact of a laryngectomy and surgical closure technique on swallow biomechanics and dysphagia severity. Otolaryngol Head Neck Surg 144(1):21–28
28. Singer MI, Blom ED, Hamaker RC (1986) Pharyngeal plexus neurectomy for alaryngeal speech rehabilitation. Laryngoscope. 96(1):50–54
29. Balm AJ, van den Brekel MW, Tan IB, Hilgers FJ (2011) The indwelling voice prosthesis for speech rehabilitation after total laryngectomy: a safe approach. Otolaryngologia polska/Polish Otolaryngol 65(6):402–409
30. Ng ML, Liu H, Zhao Q, Lam PK (2009) Long-term average spectral characteristics of Cantonese alaryngeal speech. Auris Nasus Larynx 36(5):571–577
31. Boersma P, Weenink D. Praat: doing phonetics by computer (Version 5.3.53) [Computer program]. http://www.praat.org. 5.3.53 ed2013
32. Traunmuller H, Eriksson A (1995) The perceptual evaluation of F0 excursions in speech as evidenced in liveliness estimations. J Acous Soc Am 97(3):1905–1915
33. Chone CT, Spina AL, Barcellos IH, Servin HH, Crespo AN (2011) A prospective study of long-term dysphagia following total laryngectomy. B-Ent 7(2):103–109
34. Maclean J, Cotton S, Perry A (2009) Dysphagia following a total laryngectomy: the effect on quality of life, functioning, and psychological well-being. Dysphagia 24(3):314–321
35. Landera MA, Lundy DS, Sullivan PA (2010) Dysphagia after total laryngectomy. Persp Swallow Swallow Disord (ASHA) 19:39–44
36. McConnel FM, Cerenko D, Mendelsohn MS (1988) Dysphagia after total laryngectomy. Otolaryngol Clin North Am 21(4):721–726
37. Op de Coul BM, van den Hoogen FJ, van As CJ, Marres HA, Joosten FB, Manni JJ et al (2003) Evaluation of the effects of primary myotomy in total laryngectomy on the neoglottis with the use of quantitative videofluoroscopy. Arch Otolaryngol Head Neck Surg 129(9):1000–1005
38. Blom ED, Pauloski BR, Hamaker RC (1995) Functional outcome after surgery for prevention of pharyngospasms in tracheoesophageal speakers. Part I: Speech characteristics. Laryngoscope 105(10):1093–1103
39. Chone CT, Seixas VO, Andreollo NA, Quagliato E, Barcelos IH, Spina AL et al (2009) Computerized manometry use to evaluate spasm in pharyngoesophageal segment in patients with poor tracheoesophageal speech before and after treatment with botulinum toxin. Braz J Otorhinolaryngol. 75(2):182–187
40. Deschler DG, Doherty ET, Reed CG, Hayden RE, Singer MI (2000) Prevention of pharyngoesophageal spasm after laryngectomy with a half-muscle closure technique. Ann Otol Rhinol Laryngol 109(5):514–518
41. Olson NR, Callaway E (1990) Nonclosure of pharyngeal muscle after laryngectomy. Ann Otol Rhinol Laryngol 99(7 Pt 1): 507–508
42. van As CJ, Op de Coul BM, van den Hoogen FJ, Koopmans-van Beinum FJ, Hilgers FJ (2001) Quantitative videofluoroscopy: a new evaluation tool for tracheoesophageal voice production. Arch Otolaryngol Head Neck Surg 127(2):161–169
43. Op de Coul BM, Hilgers FJ, Balm AJ, Tan IB, van den Hoogen FJ, van Tinteren H (2000) A decade of postlaryngectomy vocal rehabilitation in 318 patients: a single Institution's experience with consistent application of provox indwelling voice prostheses. Arch Otolaryngol Head Neck Surg 126(11):1320–1328
44. Evans S, Neave N, Wakelin D (2006) Relationships between vocal characteristics and body size and shape in human males: an evolutionary explanation for a deep male voice. Biol Psychol 72(2):160–163
45. Sisty NL, Weinberg B (1972) Formant frequency characteristics of esophageal speech. J Speech Hear Res 15(2):439–448
46. Schuster M, Rosanowski F, Schwarz R, Eysholdt U, Lohscheller J (2005) Quantitative detection of substitute voice generator during phonation in patients undergoing laryngectomy. Arch Otolaryngol Head Neck Surg 131(11):945952
47. Moon JB, Weinberg B (1987) Aerodynamic and myoelastic contributions to tracheoesophageal voice production. J Speech Hear Res 30(3):387–395
48. Hilgers FJ, Dirven R, Jacobi I, van den Brekel M (2015) Physiology and prospects of bimanual tracheoesophageal brass instrument play. Acta Otorhinolaryngol Italica 35(3):202–207