Alternative Voice After Laryngectomy Using a Sound-Producing Voice Prosthesis

M. van der Torn, MD; M. P. de Vries, PhD; J. M. Festen, PhD; I. M. Verdonck-de Leeuw, PhD; H. F. Mahieu, PhD

**Objective:** To improve the voice quality of female laryngectomees and/or laryngectomees with a hypotonic pharyngoesophageal (PE) segment by means of a pneumatic artificial source of voice incorporated in a regular tracheoesophageal (TE) shunt valve. **Methods:** The new sound source consists of a single silicone lip, which performs an oscillatory movement driven by expired pulmonary air flowing along the outward-striking lip through the TE shunt valve. A prototype of this pneumatic sound source is evaluated in vitro and in six laryngectomees. In vivo evaluation includes speech rate, maximal phonation time, perceptual voice evaluation of read-aloud prose by an expert listener, speech intelligibility measurements with 12 listeners, and self-assessment by the patients. Moreover, extensive acoustical and aerodynamic in vivo registrations are performed using a newly developed data acquisition system. **Results:** The current prototype seems beneficial in female laryngectomees with a hypotonic PE segment only. For them the sound-producing voice prosthesis improves voice quality and increases the average pitch of voice, without decreasing intelligibility or necessitating other pressure and airflow rates than regular TE shunt speech. Pitch regulation of this prosthetic voice is possible, yet limited. **Conclusions:** The mechanism is feasible and does not result in unacceptable airflow resistance. For this new mechanism of alaryngeal voice to become an established technique for postlaryngectomy voice restoration, a voice suitably pitched for male laryngectomees has to be generated and a large part of the melodic and dynamic range of the sound source has to be attainable within physiological airflow rates. **Key Words:** Laryngectomy, artificial larynx, voice prosthesis, alaryngeal speech, voice.

**INTRODUCTION**

Total laryngectomy, as surgical treatment for locally advanced laryngeal tumors, interferes with all functions of the larynx, i.e., phonation, respiration, and occasionally deglutition. Most problems of laryngectomized patients are related to the loss of voice: not being understood, not being able to make oneself heard in noisy rooms because of lack of volume, feeling embarrassed by their substitute voice, and feeling inhibited in expressing emotions such as anger and joy. Accordingly, rapid and effective voice restoration is critical to the successful prevention of psychological, social, and economic setbacks resulting from post-laryngectomy aphonia.

Numerous attempts to obtain or improve postlaryngeal voice by creating a pneumatic artificial source of voice production have been made since laryngectomies have been performed. Esophageal injection speech and electrolarynx devices, however, formed the standard approaches to alaryngeal voice rehabilitation until tracheoesophageal (TE) puncture incorporating a silicone shunt valve prosthesis evolved worldwide as an established technique for postlaryngeal voice restoration. The advantages of this method over esophageal injection speech are louder phonation, better intelligibility, usually quick and trouble-free voice acquisition, higher speech rate, and more sustained phrasing resulting from a larger available air reservoir.

The term “voice prosthesis” is widely used in the literature when referring to TE shunt valves, although these devices do not actually produce sound. In both esophageal injection voice and TE shunt voice, the passage of air through the pharyngoesophageal (PE) segment sets the closely approximated mucosal surfaces of this structure into vibration, producing a low-pitched sound that can be used as a substitute voice. If, however, the toxicity of the PE segment is too low to attain sufficient mucosal apposition, the resulting voice will be weak and breathy, or merely a coarse whisper. Female laryngectomees especially often have severe problems accepting their low-pitched alaryngeal voice. The mean fundamental frequency of laryngeal female voice ranges from 180 to 240 Hz, which decreases after laryngectomy and current voice rehabilitation to an unnaturally low pitch of 60 to 120 Hz.
To improve voice quality of these two groups of laryngectomees (females and those with a hypotonic PE segment), a small pneumatic sound source to be incorporated in a regular TE shunt valve was designed in cooperation with the Department of Biomedical Engineering at the University of Groningen. In a preceding pilot study, four different prototypes were tested, consisting of either one or two bent silicone lips, based on the outward-striking reed principle. The pilot study proved the feasibility of the mechanism and provided us with directions for the development of the current sound-producing shunt valve. This report describes our results with a newly developed, single-lipped vibrating element placed in a Groningen ultra-low-resistance shunt valve in vitro and in a group of six laryngectomized patients.

MATERIALS AND METHODS

Sound-Producing Voice Prosthesis

The new sound source consists of a single bent silicone lip (10.0 × 5.5 × 0.25 mm), which performs self-sustaining oscillations driven by the expired pulmonary air that flows along the outward-striking lip through the TE shunt valve. The resulting frequency of oscillation and sound intensity can be modified by altering the airflow along the lip. To allow easy placement and removal of the sound source for speech evaluation purposes in our experimental setting, the bent silicone lip was fitted in a small rectangular metal container (Fig. 1), which can be partly inserted in a patient’s Groningen shunt valve (Fig. 2). Thus, we were able to evaluate the sound-producing voice prosthesis (SPVP) without the burden of replacing the entire shunt valve. After optimization of frequency, timbre, and sound intensity using these replaceable sound-producing modules, the intention is to eventually integrate the bent silicone lip in the design of a regular TE shunt valve, e.g., the Groningen ultra-low-resistance (Gr.ULR).

In Vitro Measurements

In vitro measurements were performed with the experimental set-up shown in Figure 3. A pressure transducer (DT-2–14P–0–10L, Modus Instruments, Clinton, MA) measured the air pressure in front of the tested voice prosthesis. The transdevice airflow was determined using a respiratory flowhead (F300 L, Mercury Electronics, Glasgow, Scotland) connected to another pressure transducer (DT-2–02P–0–10L, Modus Instruments). These transducers were accurately calibrated by means of a calibration analyzer (Timeter RT-200, Allied Health Care, St. Louis, Missouri).
MO). A miniature electret condenser microphone (9468, Microtron, Amsterdam, The Netherlands) with a hybrid integrated amplifier was attached to the flowhead. The sensor signals were preprocessed using National Instruments SCXI signal-conditioning hardware, sampled at 10000 samples per second by means of a PC-based 16-bit analog–digital converter (PCI-MIO-16xe-10, National Instruments, Austin, TX) and digitally processed by a custom-built LabVIEW® software application (National Instruments, Austin, TX). Using compressed air, the pressure–airflow characteristics were determined for a Provox-II, a regular Groningen ULR, and a Groningen ULR shunt valve incorporating the sound-producing module. For every 100 ms the fundamental frequency was calculated off-line from the microphone signal by another custom-built LabVIEW® application using a harmonic product spectrum algorithm.13

Patients
Subjects were 5 females and 1 male with a mean age of 64 years (range, 51–75 y). All underwent total laryngectomy with primary tracheoesophageal puncture 13 to 84 months before this study. The male subject underwent laryngectomy for intractable aspiration causing recurrent pneumonia resulting from progressive oculopharyngeal muscular dystrophy. The five females underwent laryngectomy for laryngeal carcinomas, four of them primary, one of them a recurrence after radiotherapy. Three subjects received radiotherapy postlaryngectomy, three subjects underwent uni- or bilateral neck dissections, one patient required a pectoralis major myocutaneous flap reconstruction, and four patients had a primary pharyngeal myotomy. PE segment tonicity judgments were based on both postoperative videofluoroscopic observations and voice characteristics.14 Clinical data are summarized in Table I. Spoken and written informed consent was obtained from all patients in this study and their travel expenses were reimbursed. The medical ethics committee of the Vrije Universiteit Medical Center, Amsterdam, approved the research protocol.

Statistical Evaluation
Two block-randomized groups of patients (block size 2) were formed for this crossover trial. One group started all described vocal tests with their own Gr.ULR shunt valve, while the other group first performed all measurements with the new sound-producing module inserted in their TE shunt valve. Results for the various test parameters, obtained in both situations, were compared and statistically analyzed using repeated-measures analysis of variance and paired Student t test or corresponding nonparametric tests.

Acoustical and Aerodynamic In Vivo Registrations
In vivo registrations were performed with the experimental set-up shown in Figure 4. For this purpose we used the same hard- and software as for the in vitro measurements: the DT-2–14P–0–10L pressure transducer was connected to a tracheostoma adapter (BE-6040, Inhealth Technologies, Carpinteria, CA) and the flowhead with the condenser microphone was fitted on a facemask (King Systems Corp., Noblesville, IN). In

<table>
<thead>
<tr>
<th>Patients (n = 6)</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>M1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>75</td>
<td>67</td>
<td>51</td>
<td>63</td>
<td>52</td>
<td>75</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Female</td>
<td>Female</td>
<td>Female</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Months since laryngectomy</td>
<td>72</td>
<td>84</td>
<td>35</td>
<td>13</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>pTNM stage</td>
<td>Recurrence after RT</td>
<td>T3N0M0</td>
<td>T3N2bM0</td>
<td>T4N0M0</td>
<td>T4N0M0</td>
<td>NA</td>
</tr>
<tr>
<td>Neck dissection</td>
<td>Bilateral</td>
<td>Bilateral</td>
<td>Unilateral</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Pharyngeal closure method</td>
<td>Vertical</td>
<td>T-shaped</td>
<td>PM-flap</td>
<td>T-shaped</td>
<td>T-shaped</td>
<td>T-shaped</td>
</tr>
<tr>
<td>Pharyngeal myotomy</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>18 mo preop</td>
<td>None</td>
<td>Postop</td>
<td>Postop</td>
<td>Postop</td>
<td>None</td>
</tr>
<tr>
<td>Tonicity PE segment</td>
<td>Hypotonic</td>
<td>Hypotonic</td>
<td>Extremely hypotonic</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
</tbody>
</table>

pTNM stage = pathological TNM stage; RT = radiotherapy; PE = pharyngoesophageal; PM = pectoralis major myocutaneous; NA = not applicable.
addition, an electro-glottograph (EG-830, Froekjaer-Jensen Electronics, Holte, Denmark) was connected to the patient through a set of Ag/AgCl electrodes. Moreover, an 8F catheter with three pressure transducers (SSD-915, Millar Instruments, Houston, TX) was inserted through the nose until the distal sensor was located at the level of the esophageal flange of the TE shunt valve. This location was checked with a small flexible endoscope. Because all channels were sampled simultaneously with 16-bit resolution at 10,000 samples per second, both static pressure and pitch-synchronous pressure variations are available for three pharyngeal levels (as described by Cranen and Boves for similar recordings of laryngeal voice). Recordings were made during phonation of a sustained vowel /a/ at comfortable loudness. Vocal intensity at 30 cm from the facemask was obtained from the root mean square power level (window width 500 ms) of the microphone signals after comparison to a fixed-level reference tone. Intratracheal pressure, airflow and absolute esophageal pressure were read off-line from the raw data, whereas the acoustic signals were converted to .WAV files for further software analysis. The mean fundamental frequency during a period of stable phonation was estimated from the spectrogram.

Speech Rate and Maximal Phonation Time

All subjects’ responses were recorded in a sound-treated room using a microphone (MKE 212–3, Sennheiser, Germany) and a DAT recorder (DA-7, Casio Computers, Japan). To establish speech rate, determined as the number of words per minute, each subject was asked to read the first paragraphs of the Dutch prose “De Vijvervrouw” in a normal conversational manner. The maximal phonation time in seconds was obtained by sustaining for as long as possible the vowel /a/ at a comfortable pitch and loudness on one deep breath. The best attempt out of three was taken for statistical evaluation. The average speech rate of 52 laryngeal reference speakers from our institute was 181 words per minute (standard deviation [SD] = 24) for this standard text.

Intelligibility

Each of the six laryngectomies read aloud two lists of 13 short, common Dutch sentences, one list while using regular TE shunt voice and one with the new SPVP. The 156 sentences were sampled at 44,1 kHz and 16-bit resolution. For each speech recording, an additional fixed-level reference tone was recorded. One factor in speech communication, the long-term average speech level, was determined by measuring the average speech level for each of the 12 lists from the average RMS levels (window width 100 ms) using the recorded tones as reference. A second factor, intelligibility, was measured in terms of the speech reception threshold (SRT) in competing noise, i.e., the signal to noise ratio at which 50% of the sentences were reproduced without a single error by listeners with normal hearing. Because speech from multiple speakers is the most frequent disturbing sound in everyday situations, SRT was determined in interfering noise of 60 dBA with a spectrum equal to the long-term average spectrum of normal speech. An adaptive up-and-down procedure for accurately measuring the SRT with these lists of 13 sentences was developed for evaluating hearing impairment and later adapted for intelligibility assessment of impaired speech. The 12 lists were presented binaurally to 12 listeners according to a 12 x 12 balanced Latin square. Tucker Davis hardware (System II, Tucker Davis Technologies, Gainesville, FL) and headphones (DT-48, Beyer, Germany) were used to play out and attenuate the samples.

Perceptual Voice Evaluation by a Professional Listener

Digital recordings of approximately 90 seconds, also used to establish speech rate, were made for each laryngectomee with both types of voice prosthesis. The speech material consisted of the first paragraphs of the Dutch prose “De Vijvervrouw.” A subset of seven-point bipolar semantic scales, developed by Nie-
boer\textsuperscript{18} and later modified\textsuperscript{19,20} were adapted for our purpose. This subset included the five scales “high pitch–low pitch,” “weak–powerful,” “tense–nontense,” “gurgling–nongurgling,” and “melodious–monotonous.” The 12 numbered recordings were presented twice in random order to a voice-and-speech pathologist experienced with laryngectomized speakers. To play out the stored samples directly from the host computer, a Kay CSL unit (CSL 4300B, Kay Elemetrics, Pine Brooks, NJ) and headphones (MDR-V900, Sony, Japan) were used. Intrarater reliability was calculated from the differences between test and retest. For comparison with literature data, the reliability was defined as the percentage of test–retest differences smaller than or equal to one scale value.

**Self-Assessment by Patients**

Each patient only had between 2 and 4 hours of experience with the new sound-producing voice prosthesis, during which all measurements were performed. We deemed this period too short for a patient to estimate the impact of the SPVP on daily functioning and quality of life by means of an extensive questionnaire, like the VHI\textsuperscript{21} or the EORTC QLQ-H&N\textsuperscript{35}. Therefore, a short questionnaire was designed to assess the patients’ primary judgment of voice quality attained with the SPVP compared with their regular TE shunt voice. The items were effort required for speaking, pitch of voice, vocal control, vocal intensity, and intelligibility of speech. After the last measurement, patients were asked to fill out a form indicating whether these five aspects of their voice were influenced positively, not at all, or negatively by the SPVP.

**RESULTS**

**In Vitro Measurements**

Figure 5 shows the in vitro relationship between the airflow through and the pressure over the voice prostheses: Provox-II, the Gr.ULR, and the SPVP. Each datapoint represents the average of 1000 consecutive samples. At airflow rates above 0.15 L/s, the Gr.ULR incorporating the sound-producing module requires increasingly more pressure than the regular Gr.ULR shunt valve, which requires a little more than the Provox-II. At regular airflow rates for alaryngeal speech, which hardly exceed 0.30 L/s,\textsuperscript{22} the sound-producing module will add only little to the total pressure needed for phonation in vivo (see Table II).

Figure 6 shows the relationship between pressure over the SPVP and fundamental frequency of the tone.

<table>
<thead>
<tr>
<th>Patients</th>
<th>$f_0$ (Hz)</th>
<th>Sound Pressure Level (dB)</th>
<th>Tracheal Pressure (kPa)</th>
<th>Airflow Rate (L/s)</th>
<th>$f_0$ (Hz)</th>
<th>Sound Pressure Level (dB)</th>
<th>Tracheal Pressure (kPa)</th>
<th>Airflow Rate (L/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>75</td>
<td>57–70</td>
<td>2.5–3.0</td>
<td>0.11–0.24</td>
<td>210</td>
<td>55–69</td>
<td>2.1–3.0</td>
<td>0.08–0.25</td>
</tr>
<tr>
<td>F2</td>
<td>95</td>
<td>53–77</td>
<td>1.3–5.5</td>
<td>0.05–0.17</td>
<td>255</td>
<td>47–58</td>
<td>1.8–3.3</td>
<td>0.06–0.11</td>
</tr>
<tr>
<td>F3</td>
<td>32</td>
<td>50–57</td>
<td>1.9–2.5</td>
<td>0.14–0.27</td>
<td>200</td>
<td>56–64</td>
<td>2.1–2.4</td>
<td>0.14–0.23</td>
</tr>
<tr>
<td>F4</td>
<td>70</td>
<td>58–72</td>
<td>4.0–4.5</td>
<td>0.08–0.25</td>
<td>55</td>
<td>56–65</td>
<td>2.4–3.8</td>
<td>0.04–0.11</td>
</tr>
<tr>
<td>F5</td>
<td>130</td>
<td>66–77</td>
<td>4.9–8.1</td>
<td>0.24–0.57</td>
<td>250</td>
<td>45–63</td>
<td>4.2–5.6</td>
<td>0.04–0.09</td>
</tr>
<tr>
<td>M1</td>
<td>45</td>
<td>53–68</td>
<td>2.4–3.1</td>
<td>0.01–0.06</td>
<td>55</td>
<td>53–61</td>
<td>2.1–3.1</td>
<td>0.01–0.06</td>
</tr>
</tbody>
</table>

$\text{TE} = \text{tracheoesophageal}$.
Fig. 6. In vitro measurement of fundamental frequency versus pressure for the newly developed sound-producing voice prosthesis.

Fig. 7. Short disturbance of the tone (approximately 200 Hz) of the sound-producing module by low-frequency vibrations (approximately 35 Hz) of the patient’s PE segment during sustained phonation of a vowel /a/.
produced by the sound-producing module in vitro. Below 1 kPa the tone of the SPVP is too weak for the algorithm to discern the fundamental frequency from the noise. Between 1 and 2.2 kPa the fundamental frequency decreases from 265 to 210 Hz, after which it increases gradually to 315 Hz at 5.4 kPa.

**Acoustical and Aerodynamic In Vivo Registrations**

For all patients the fundamental frequency, estimated for a period of stable phonation, the vocal intensity range, the intratracheal pressure range and airflow range are shown in Table II for both types of voice. Using the SPVP, four female patients (F1, F2, F3, and F5) attained mean fundamental frequencies between 200 and 255 Hz, comparable with laryngeal female voices.

Two patients (F4 and F5), classified as having a normotonic PE segment, found it difficult not to evoke the low-frequency mucosal vibrations that form the regular TE shunt voice and used the SPVP for voice production instead. Apparently, the pulmonary airflow through the PE segment already passively evoked these vibrations. By voluntary reduction of intratracheal pressure and airflow at the expense of vocal intensity, subject F5 did manage to restrain her PE segment and make the SPVP audible, whereas subject F4 did not. Sound generated by mucosal vibrations of the PE segment is identified by a corresponding electro-glottographic signal, whereas sound produced by the SPVP does not result in electro-glottographic activity. This detail is illustrated by Figure 7, showing a brief disruption of the SPVP tone by the mucosal vibrations of the PE segment. For subjects F4 and M1 the unavoidable vibrations of the PE segment dominated, thereby masking the SPVP voice.

Most subjects experienced a paradoxically rising pitch at the end of each spoken sentence. The usual reduction of airflow and intratracheal pressure at the end of phonation caused vocal intensity to decrease and a gradual rise of fundamental frequency from approximately 200 Hz to 240 Hz, as shown in Figure 8, for sustained phona-
tion of a vowel /a/. This finding corresponds to the results of the in vitro measurements shown in Figure 6.

**Speech Rate and Maximal Phonation Time**

For all patients the average number of words per minute is shown in Table III. The difference between SPVP speech (mean, 116 wpm) and regular TE shunt speech (mean, 119 wpm) was not significant (t5 = 0.392; P = .71).

The maximal phonation time was obtained for only four subjects, as shown in Table III. The difference between the SPVP voice (mean, 16.3 s) and the regular TE shunt voice (mean, 14.3 s) was not significant (t5 = 0.730; P = .52).

**Intelligibility**

Table IV presents the long-term average speech levels of each patient for both voice sources. For three patients (F1, F4, and F5) speech levels were reduced using the SPVP compared with levels obtained for regular TE shunt speech. Subject F3, however, gained 7 dBA. For the total group of patients, the difference between SPVP speech (mean, 70 dBA) and regular TE shunt speech (mean, 73 dBA) was not significant (t5 = 1.026; P = .35).

Table IV also shows the speech reception threshold in noise for the five female patients and both voice prostheses, averaged over the 12 listeners. A lower S/N ratio at the SRT indicates a better intelligibility of the speech. Using the SPVP, intelligibility in noise of two laryngectomees (F4 and F5) was reduced by 5.7 dB and 3.3 dB, respectively, compared with their regular TE shunt voices. For patients F1 and F2 the SRT was virtually unchanged by the SPVP. A repeated-measures analysis of variance (5 patients × 2 sound sources × 12 listeners) showed the SRT differences between both voice sources to be not significant (F14 = 3.591; P = .131), nor the differences between the 12 listeners (F11,44 = 1.849; P = .074). The differences between the five patients were significant (F14 = 20.318; P = .011).

Because of muscular dystrophy of his articulators, the intelligibility of patient M1 was too poor to determine the SRT using the adaptive up-and-down procedure. For this patient the number of words repeated correctly by the 12 listeners at an S/N ratio of +20 dB (headphones speech level: approximately 80 dBA) was counted in both situations. Using the SPVP, an average of 29% (SD, 11.4%) out of 85 words was repeated correctly, whereas 55% (SD, 9.0%) out of 81 words was repeated correctly using his regular TE shunt voice. This difference was significant (t11 = 8.360; P < .0005).

**Perceptual Voice Evaluation by a Professional Listener**

For all 60 judgments (6 patients × 2 voice prostheses × 5 scales), the first rating was within 1 scale value of the second rating, rendering an intrarater reliability of 100%. For 76.7% of the judgments, test and retest were even exactly identical.

Table V shows for each patient and each scale, after averaging test and retest, the rating difference between SPVP voice and the patient’s regular TE shunt voice.

Positive values obtained in this way indicate improvement of voice quality by the SPVP: less gurgling, higher pitch, more strength, more melodic, or less tense. According to separate Wilcoxon signed rank tests, none of the scales significantly differentiated between regular TE shunt voices and SPVP voices, although the SPVP tended to cause a higher pitched, yet monotonous voice. In fact, the effects of the SPVP varied considerably between patients. The effect of the SPVP was almost absent in patients F4 and M1, because they were unable to prevent the mucosal vibrations of their PE segment when using the SPVP. One patient’s voice (F5) deteriorated on nearly every scale: weaker and more monotonous, tense, and gurgling. Another patient’s voice (F3), however, became more powerful and less gurgling using the SPVP, whereas her melodiousness and tenseness remained unaltered. Because her regular TE shunt voice is just an aphone whisper, pitch could not be judged in that situation. In patients F1 and F2 pitch increased drastically and gurgling decreased, whereas melodiousness and vocal strength deteriorated.

**Self-Assessment by Patients**

Table VI shows for each patient the judgment for each of the five questionnaire items. Five patients (83%) judged the pitch of their voice to be higher. Some (F2, F3, and M1) judged the SPVP to improve vocal intensity and intelligibility. Patient F5 judged the SPVP inferior to her own TE shunt speech on all items except pitch. Three laryngectomees (F1, F4, and F5) judged vocal intensity to decrease using the SPVP. Patient F3 was not able to compare pitch of the SPVP voice to her regular TE shunt voice, because the latter is an aphone whisper.

**DISCUSSION**

This report is presented in the nature of initial observations on a new mechanism of alaryngeal voice by means of a single vibrating silicone lip. Previous attempts to create an alaryngeal pseudovoice by means of an inward-striking metal reed invariably led to an unnatural monotonous voice.23,24 In our preceding pilot study we found that outward-striking 12 silicone lips generated a tone of variable pitch, although in the initial prototypes

---

**TABLE IV.**

<table>
<thead>
<tr>
<th>Patients</th>
<th>Average RMS Power [dBA]</th>
<th>S/N ratio at SRT [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPVP</td>
<td>Gr.ULR</td>
<td>SPVP</td>
</tr>
<tr>
<td>F1</td>
<td>68</td>
<td>73</td>
</tr>
<tr>
<td>F2</td>
<td>73</td>
<td>74</td>
</tr>
<tr>
<td>F3</td>
<td>74</td>
<td>67</td>
</tr>
<tr>
<td>F4</td>
<td>66</td>
<td>73</td>
</tr>
<tr>
<td>F5</td>
<td>65</td>
<td>74</td>
</tr>
<tr>
<td>M1</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>70 (4.7)</td>
<td>73 (3.0)</td>
</tr>
</tbody>
</table>

SPVP = sound-producing voice prosthesis; Gr. ULR = Groningen ultra-low-resistance; SD = standard deviation; NA = not applicable.
the resulting voice was unnaturally high-pitched and not loud enough. The current prototype has a single lip which is broader than before, and a metal housing which is wider and higher than before, while protruding just 7 mm from the tracheal flange of the voice prosthesis.

The present findings in vitro indicate that airflow resistance at regular airflow rates for alaryngeal speech (0.05–0.30 L/s) is marginally increased by the new SPVP prototype compared with two commonly applied types of voice prosthesis. Moreover, the fundamental frequency of the tone generated by the single outward-striking silicone lip could be varied between 210 and 315 Hz by altering the airflow rate between 0.05 and 0.75 L/s. In vivo, however, airflow rates vary from 0.05 to just 0.30 L/s, thereby addressing only a small part of the SPVP’s melodic and dynamic range. The resulting narrow frequency range caused our voice-and-speech pathologist and several patients to complain about the monotonous nature of the voice generated by the SPVP.

To clarify several observations, such as the U-shaped graph in Figure 6 and the paradoxically rising pitch at the end of sentences spoken by the patients, we evaluated the SPVP’s silicone lip movement in vitro by means of video-stroboscopy. Through this informal evaluation we learned that for airflow rates between 0.05 and 0.45 L/s, the oscillation gradually involves a larger part of the lip, starting from its tip. This gradual involvement causes the mass and length of the oscillating part to increase with the airflow rate, thereby decreasing the frequency of oscillation. At higher airflow rates the entire lip is involved. For SPVP speech to acquire more natural prosody, the silicone lip of future prototypes will have to oscillate entirely at lower airflow rates, preferably less than 0.10 L/s.

In vivo the quality of voice attained through the SPVP mainly depended on the tonicity of the PE segment, i.e., the ability to let pulmonary air flow through the PE segment without evoking the mucosal vibrations that form the regular TE shunt voice. Three patients (F1, F2, and F3), classified as having a hypotonic PE segment, were able to do so within 10 minutes after insertion of the sound-producing module. For these subjects, the SPVP increased the average pitch of voice to meet the pitch of laryngeal female speakers without decreasing intelligibility or necessitating other pressure and airflow rates than regular TE shunt speech. While they succeeded in preventing the mucosal surfaces of their PE segment from

### TABLE VI.
Perceptual Evaluation of Voice Quality by an Expert Listener for Six Laryngectomees Using the SPVP, Compared With Their Regular TE Shunt Voice

<table>
<thead>
<tr>
<th>Patients</th>
<th>Low Pitch-High Pitch</th>
<th>Gurgling-Nongurgling</th>
<th>Weak-Powerful</th>
<th>Tense-Nontense</th>
<th>Monotonous-Melodious</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>2.5</td>
<td>3</td>
<td>-2</td>
<td>-1</td>
<td>-1.5</td>
</tr>
<tr>
<td>F2</td>
<td>3</td>
<td>1</td>
<td>-1.5</td>
<td>-2</td>
<td>-1.5</td>
</tr>
<tr>
<td>F3</td>
<td>NA</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F4</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F5</td>
<td>0</td>
<td>-1</td>
<td>-1.5</td>
<td>-2</td>
<td>3</td>
</tr>
<tr>
<td>M1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>+0.5</td>
<td>-1</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>+1.40 (1.29)</td>
<td>+0.67 (1.37)</td>
<td>-0.33 (1.83)</td>
<td>-0.75 (1.08)</td>
<td>-1.17 (1.13)</td>
</tr>
<tr>
<td>Significance</td>
<td>0.068</td>
<td>0.257</td>
<td>0.713</td>
<td>0.141</td>
<td>0.066</td>
</tr>
</tbody>
</table>

*Positive values indicate voice quality improvement: higher pitch, less gurgling, more strength, less tense, or more melodious.

SPVP = sound-producing voice prosthesis; TE = tracheoesophageal; SD = standard deviation; NA = not applicable.

### TABLE VII.
Self-Assessment of the Sound-Producing Voice Prosthesis by Six Patients, Compared With Their Regular TE Shunt Voice

<table>
<thead>
<tr>
<th>Patients</th>
<th>Effort for Speaking</th>
<th>Pitch of Voice</th>
<th>Vocal Control</th>
<th>Loudness of Voice</th>
<th>Intelligibility of Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>F2</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>F3</td>
<td>0</td>
<td>+</td>
<td>NA</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>F4</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>F5</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M1</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
</tr>
</tbody>
</table>

*A plus indicates improvement: less effort for speaking, higher pitch, better vocal control, louder voice and better intelligibility. A minus indicates the opposite.

TE = tracheoesophageal; NA = not applicable.
vibrating, the SPVP voice sounded in fact clear and less gurgling.

For subject F3 especially, the new SPVP was beneficial, because her regular TE shunt voice consists of aperiodic noise, whispering, resulting from severe hypotonicity of her PE segment after pectoralis major myocutaneous flap reconstruction. The SPVP provided her with a tonal voice that was markedly louder than her regular TE shunt voice.

Subjects F4, F5, and M1, classified as having a normotonic PE segment, were not able to let air flow through the TE fistula and the PE segment without provoking low-frequency mucosal vibrations. At physiological airflow rates, the typical PE segment voice and the newly introduced tone of the SPVP intermingled, thereby deteriorating the vocal intensity and intelligibility of both without attaining appreciable advantages such as a higher pitch or less gurgling. From this group, only patient F5 managed to restrain her normotonic PE segment on demand with difficulty by decimating the airflow rate at the expense of vocal intensity, intelligibility, and melodiousness.

To date, the aim has been to achieve the correct amount of sphincter tonicity for voice production in the PE segment, neither hypertonic nor hypotonic. Previous research and the current study, however, indicate that sound-producing voice prostheses yield optimal results in laryngectomees with a hypotonic PE segment. Four different methods to decrease tonicity of the PE segment are described in the literature: 1) myotomy of the PE segment muscles, 2) nonclosure of the pharyngeal musculature, 3) unilateral pharyngeal plexus neurectomy, and 4) botulinum toxin injections of the cricopharyngeal muscle complex. The residual tension in the PE segment after each of these procedures is hard to predict, giving rise to a certain risk of hypotonicity.

The present SPVP prototype was designed simply to test the feasibility and characteristics of the new mechanism of alaryngeal speech by a single silicone lip. Thence it temporarily turned a blind eye to important issues like durability, clogging of the device, and patient cleaning procedures. After optimization of melodiousness, timbre, and vocal intensity using replaceable sound-producing modules, the bent silicone lip will be integrated into the design of a regular TE shunt valve, and the durability issues can be assessed.

CONCLUSION

Alternative voice production after laryngectomy by means of a single-lipped sound-producing voice prosthesis is feasible and does not result in unacceptable airflow resistance. Pitch regulation of this prosthetic voice is possible, yet limited. The current prototype was beneficial in female laryngectomees with a hypotonic PE segment only. For them, the sound-producing voice prosthesis increased the average pitch of voice and reduced gurgling without decreasing intelligibility or necessitating other pressure and airflow rates than regular TE shunt speech.

For this new mechanism of alaryngeal voice to become an established technique for postlaryngectomy voice restoration, we will direct further development at: 1) generating a voice suitably pitched for male laryngectomees as well, 2) attaining a large part of the prosthesis’ melodic and dynamic range at physiological airflow rates, and 3) enabling convenient cleaning of the sound-producing voice prosthesis by the patient.

Acknowledgments

Grant GGN 55.3712 from the Dutch Technology Foundation STW supported this study. The authors acknowledge the valuable input provided by A. G. J. E. Peeters and A. van der Plaats.

BIBLIOGRAPHY


