Evaluating the effect of different voice prostheses on alaryngeal voice quality

Coffey, MM

Evaluating the effect of different voice prostheses on alaryngeal voice quality
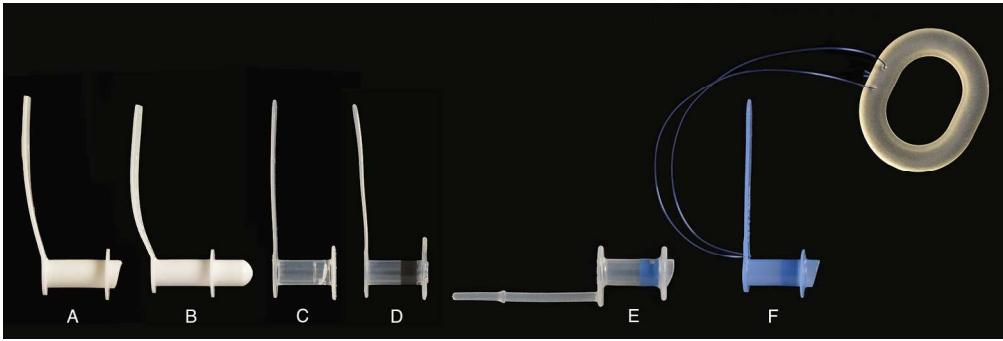
SCHOLARONE™
Manuscripts

Fig 1: Illustration of voice prostheses used in this study
(A) Blom Singer Low Pressure (B) Blom Singer Duckbill (C) Blom Singer Classic Indwelling (D) Blom Singer Advantage (E) Provox Vega (F) Provox NID

170x56mm (300 x 300 DPI)

Effect of voice prostheses on voice

Introduction

Functional communication with a voice prosthesis after laryngectomy is currently considered the gold standard of rehabilitation for post laryngectomy patients. Since the introduction of the Blom Singer duckbill voice prosthesis [1], an array of voice prostheses have been developed [2], [3,4] [5]. This study uses expert auditory perceptual ratings and patient self-evaluation to compare voice quality amongst a range of commonly available voice prostheses.

Methods of analysis of voice after laryngectomy

Analysis of voice post laryngectomy has been influenced by comparisons with normal laryngeal speakers [6]. Alaryngeal voice analysis has focused on three main areas; acoustic analysis [7-13], auditory perceptual analysis [14-16] and patient self evaluation [17] [18] [19] [20] [21], [22] [23] [24] [21] [25] [26] [27], [28] [29]. A smaller number of studies [14] [15] have combined acoustic and auditory perceptual analysis of alaryngeal voice.

Acoustic analysis of the voice signal in patients with a larynx involves computerised measurement of specific properties of the sound wave form [30] such as jitter, shimmer, and harmonics to noise ratio [31]. Auditory perceptual rating by an expert listener judging a voice sample [30] may include parameters such as intelligibility, voice quality and acceptability [6]. [32]. Patient self evaluation analysis is based on patient's perception of voice [33].

Effect of voice prostheses on voice

Previous studies comparing voice prostheses

Several studies have compared a limited number of voice prostheses using a variety of alaryngeal voice analysis methods. The first study to do so, [34] used acoustic measures and found that voice produced with a Blom Singer Low pressure prosthesis had greater amounts of periodic phonation compared with a Blom Singer duckbill. A larger study [35] compared the Blom Singer Classic Indwelling voice prosthesis with the Provox 1 using auditory perceptual analysis, acoustic measures and self evaluation by patients. While both prostheses were judged to be similar overall for voice quality, lifespan and patient satisfaction, the Blom Singer Classic indwelling was better than the Provox on a single acoustic measure of minimal loudness.   Self-assessment by patients also indicated a better voice quality with the Blom Singer Classic indwelling. A comparison [5] of the Provox NID and Blom Singer low pressure found that subjects preferred the Provox NID because of its safety and less effortful speech. A further comparison [36] of the indwelling Provox Vega and the indwelling Blom Singer Classic found that both subject self-assessment and perceptual analysis favoured the Provox Vega. A study focusing on patient perception [36] found a preference for the Provox Vega when compared with the Blom Singer Classic indwelling. Taken as a whole, these studies do not indicate a consensus on which prosthesis is best for voice.

2

Effect of voice prostheses on voice

Objectives

- To investigate expert raters' preference for best voice prosthesis using auditory perceptual analysis (STOPS)

- To investigate expert raters' subjective preference for best overall voice prosthesis.

- To investigate patient preference for best voice prosthesis for voice using self-evaluation.

- To investigate patients' subjective preference for best overall voice prosthesis.

Materials and Methods

STOPS

The Sunderland Tracheosophageal Voice Perceptual Scale (STOPS) for professional raters is a 14-item questionnaire divided into two domains: (i) Voice quality parameters (6 items), and (ii) Parameters not related to voice quality (7 items) together with an overall score voice rating. Voice quality parameters include perceptual voice tonicity, strain, wetness, impairment of volume, impairment of social acceptability of voice and whisper. The perceptual voice tonicity parameter is measured on an 11 point bipolar semantic scale reflecting the continuum of tone [37] from hypotonic to hypertonic [38] with the absence or presence of stenotic voice measured on a separate arm to the tone scale. Each of the remaining 5 items in the voice quality parameters domain are measured on a 4 point equally appearing interval scale 0 (optimal tracheosophageal voice quality), 1 (mild), 2 (moderate) and 3 (severe).

Effect of voice prostheses on voice

Parameters not related to voice quality include impression of intelligibility, stoma noise (inhalation or exhalation noise), impairment of fluency, impairment of articulatory precision, positive features of articulation, accent and poor reader (poor literacy). Each of these parameters, with the exception of positive features of articulation is measured on a 4 point equally appearing interval scale 0 (optimal tracheosophageal voice quality), 1 (mild), 2 (moderate) and 3 (severe). Positive features of articulation are measured on an alternatively worded 4-point equally appearing interval scale 0 (neutral), 1 (good), 3 (excellent), and 4 (outstanding).

Patient self-evaluation questionnaire

A patient self-evaluation questionnaire was developed for this study based on sampling opinions from 20 patients on what made a prosthesis work well for voice. Thematic analysis of responses identified effort, understandability, consistency, loudness and appearance as important factors. These themes led to the development of a Communication with a Voice Prosthesis questionnaire with six questions (see appendix). Responses were measured on a 5-point Likert scale giving a maximal score of 30. The questionnaire contained an additional blank sheet for any further comments.

Data were derived from consecutive sampling of laryngectomy patients attending an outpatient head and neck clinic. Each patient attended two visits.

Effect of voice prostheses on voice

Visit 1

The tracheoesophageal puncture (TEP) was initially measured. Three prostheses in the appropriate size were then randomly selected from the following: Provox NID *(Atos Medical, Horby, Sweden)*, Provox Vega *(Atos Medical, Horby, Sweden)*, Blom Singer Duckbill *(InHealth Technologies, California, USA)* (not available in 20fg), Blom Singer Low Pressure *(InHealth Technologies, California, USA)*, Blom Singer Classic Indwelling *(InHealth Technologies, California, USA)* and Blom Singer Advantage *(InHealth Technologies, California, USA)* see figure 1. Randomisation was achieved using the "Research Randomizer" programme on the website http://www.randomizer.org/. Patients could see the voice prosthesis but they were not told the name of the prosthesis or the manufacturer. Each prosthesis was placed according to individual manufacturer's instructions including use of the gel cap insertion system for Blom Singer prostheses. The absence of central and peripheral leakage for each individual prosthesis was confirmed by asking each subject to take three sips of a cup of 200ml water coloured with 2ml of Silver Spoon blue food colouring (British Sugar PLC). The following protocol was then followed:

1. For each prosthesis trial, subjects had a Speedlink SL-8691-SBK spes clip-on metal microphone (Speedlink, Weertzen, Germany) attached to their clothing, 10 cm lateral to the stoma, on the opposite side to the hand used to occlude the stoma during voicing. All subjects produced voice by occluding their stoma rather than depressing a humidification exchange device or using a hands-free attachment. Subjects read a short

Effect of voice prostheses on voice

version of the Rainbow passage [39]. This was recorded onto a Sony ICD-PX820 Digital Voice Recorder with flash 2 GB (Sony, Weybridge, UK) in MP3 format to be rated later by experts.

2. After each prosthesis trial, the patient was asked to rate their experience of voice quality using a 5 point Likert Self Evaluation questionnaire with 5 questions relating to swallow, (see appendix)

   Both steps were then repeated with the remaining 2 prostheses. Once the last prosthesis has been removed, the prosthesis that the subject had in situ on arrival was replaced.

Visit 2

Each patient's voice prosthesis was removed and the first of the remaining 2 prostheses in the randomisation sequence (if the subject had a 20fg prosthesis diameter measurement) or the remaining 3 prostheses in the randomisation sequence (if the patient had a 16fg prosthesis diameter measurement) was placed. The procedure described in visit 1 was then repeated. At the end of visit 2, the patient was then asked to indicate their best overall voice prosthesis.

Data analysis

Recordings of voice samples with individual voice prostheses were extracted in MP3 format and transferred to Final Cut Pro (Apple, California, USA) to allow titles to be added to indicate anonymised subject number and anonymised voice

Effect of voice prostheses on voice

prosthesis letter. Voice samples were then exported to 3 Verbatim 4GB pinstripe USB memory sticks (Verbatim, Surrey, UK). Raters were blinded to patient, prosthesis type, gender, type of laryngectomy surgery (extended laryngectomy or standard total laryngectomy) and history of radiotherapy and chemotherapy. Voice samples were sent to three expert Speech and Language Therapy raters along with blank numbered and lettered STOPS forms which corresponded to each voice sample for each subject. Raters also received and were asked to complete a form indicating which prosthesis in their subjective opinion was "best" overall for voice for each subject.

Training of raters

Each Speech and Language Therapy rater had at least 5 years experience specialising in the communication and swallowing rehabilitation of post laryngectomy and other head and neck cancer patients and had completed advanced training in the field. Each rater participated in 3 hours of training with the investigator in the use of the STOPS.

Statistical analysis

Data was entered and analysed in IBM SPSS (Statistical Product and Service Solutions) version 23 (IBM Armonk, New York) for both the STOPS scale and the Self Evaluation Patient Questionnaire for Voice and Swallow.

Effect of voice prostheses on voice

Results

40 laryngectomy patients attended 2 appointments within a 48-hour period with one further patient attending 2 appointments within a 72-hour period. 1 patient was excluded due to failure to attend the second appointment.

Expert rater's preference for best voice prosthesis using STOPS

The mean of the three raters' scores was calculated for questions 1 and 3-10. Questions 2 (perceptual voice tonicity), 11 (impairment of articulatory precision) and 12 (positive features of articulation) were excluded because agreement between raters did not reach ICC > 0.6. Data was then unblinded so prostheses could be compared for each question. As data was not normally distributed, median and inter quartile ranges were used to compare prostheses to each other. Results indicated no difference between average ratings for each prosthesis on each question analysed on the STOPS (p>0.05), see table 1.

Effect of voice prostheses on voice

Expert raters subjective preference for best overall voice prosthesis

A consensus score for best voice prosthesis for voice for each patient was calculated from the ratings of each expert based on the prosthesis that two or more raters considered best for voice. When no consensus was achieved, this was indicated within the data. Frequency analysis for choice of "best" prosthesis based on consensus scores from expert raters was then undertaken. Results indicate the Blom Singer Classic Indwelling was most frequently chosen as best prosthesis for voice by clinicians, with the Provox Vega least frequently chosen, table 2. However, there was considerable variation between patients and for eight patients, there was no consensus between clinicians regarding best prosthesis for voice.

Patient preference for best prosthesis for voice using self-evaluation

Six questions relating to voice from the "Communication with voice prostheses self-evaluation questionnaire" were scored. The sum of the scores from each question provided a total voice score for each prosthesis for each individual patient (max=30). The higher the score achieved, the more negatively patients evaluated voice with that prosthesis. This data represented a single factor, repeated measures design with six experimental conditions. Analysis was therefore undertaken using Friedman Two Way Analysis of Variance by Ranks. Median total scores were elicited for each prosthesis, see table 3.

Effect of voice prostheses on voice

Results indicated a significant difference between the six prosthesis (Friedman X^2 (5) = 23.57, p < 0.001). Dunn-Bonferroni post hoc tests indicated that there were significant differences between the Blom Singer Duckbill prosthesis and each of the other prostheses; Blom Singer Low pressure (p <0.017), Blom Singer Classic Indwelling (p <0.001), Blom Singer Advantage (p <0.024) Provox NID (p <0.001), Provox Vega (p <0.003) after Bonferroni adjustments. There were no significant differences between any other methods, indicating that the Duckbill appears to be the least preferred prosthesis for voice, whereas all other prostheses score equally well.

Patients' subjective preference for best overall voice prosthesis

Frequency analysis of patient self-evaluation of best overall prosthesis indicated that subjects most frequently chose the Provox NID and least frequently the Duckbill (Table 4).

Effect of voice prostheses on voice

Discussion

Expert raters' preference for a voice prosthesis for voice on perceptual evaluation

The first objective was to investigate whether expert raters consider one voice prosthesis as best for voice for patients on auditory perceptual evaluation using the STOPS rating scale. Raters were unable to identify one prosthesis as best for voice despite using a laryngectomy specific auditory perceptual scale for which they had already demonstrated good agreement. However, several limitations are acknowledged. Firstly, experts as opposed to naïve listeners provided ratings. Experts may become desensitised to the abnormalities of the alaryngeal voice. Secondly, ratings were made on a standard reading passage as opposed to a more functional conversational speech task. Furthermore results may have been influenced by the expert raters lack of information about non-verbal communication including facial expression and gesture.

Expert rater preference for a "best" overall voice prosthesis

The next objective of this study was to investigate whether raters' choice of best prosthesis for each subject identified an overall best prosthesis for voice. For a large proportion of the sample, there was no consensus between clinicians regarding best prosthesis for voice. The Blom Singer Classic Indwelling was most frequently chosen as best prosthesis for voice by expert raters. However, this question represents a subjective judgment of voice quality reflecting individual raters bias and opinion and results should be interpreted cautiously.

Effect of voice prostheses on voice

Patient preference for best voice prosthesis for voice

The next objective was to investigate whether patients have a preference for a prosthesis for voice production based on self-evaluation.

Using the criteria in the self-evaluation of voice questionnaire, results indicated that the least preferred prosthesis was the Blom Singer Duckbill. The Blom Singer Duckbill as the first widely commercially available voice prosthesis requires a higher air pressure to produce voice. Since the inception of this prosthesis, the design of voice prostheses has become more refined to facilitate voice production with less pressure or effort. The increase in effort required to speak with a duckbill voice prosthesis compared with others may account for it being least preferred by patients. There was no difference in preferences between the other five prostheses.

Patient preference for best overall voice prosthesis

Finally patients indicated a preference when simply asked to identify a best overall voice prosthesis. The most frequently cited was the Provox NID, and the least preferred being the Blom Singer Duckbill. The Provox NID is an ex-dwelling voice prosthesis, which patients can learn to change themselves. It is unique amongst the prostheses trialed in that it is blue in color making it more visible in situ than clear prostheses such as the Blom Singer Classic Indwelling, the Blom Singer Advantage and the Provox Vega. Several patients also commented that this prosthesis would be easier to clean because it was more visible, concurring with a previous study [5] which found that that the Provox NID was easier to clean

Effect of voice prostheses on voice

than the Blom Singer Low Pressure. It is likely that a number of factors, other than voice, influence patients' choice of a "best" prosthesis. Further identification of these factors could lead to both improved patient choice and enhanced design of future voice prostheses.

Conclusions

Expert raters using the STOPS identified no differences between prostheses. However, subjective ratings by experts indicated that the Blom Singer Classic Indwelling may be preferred for voice. However, there was a lack of consensus amongst raters for a large proportion of the sample.

Patients' had no consistent preference of prothesis for the purposes of voice based on self-evaluation, although the Blom Singer Duckbill was least preferred. The Provox NID was most frequently chosen by patients as the best voice prosthesis overall with several factors other than voice influencing this choice

Most patients are currently not provided with a choice of voice prosthesis. They use the prosthesis recommended by their clinician or which their hospital supplies. The finding that subjects, when faced with a choice of voice prostheses, have preferences is important. It signifies that patients as end users of these devices should have access to a choice of prostheses and be involved in making that choice. This study is the first to investigate a broad range of voice prostheses. This study is also unique in that subjects themselves were instrumental in deciding which criteria would be used to self-evaluate voice prostheses. Future research is required to further analyze which properties of

Effect of voice prostheses on voice

individual voice prostheses are most influential in determining preference for a

voice prosthesis

Effect of voice prostheses on voice

References

1. Blom E, Singer M. Surgical-Prosthetic Approaches for Post-laryngectomy Voice Restoration *Laryngectomy Rehabilitation*. Houston: College Hill Press, 1979.
2. Leder S, Acton L, Kmiecik J, Ganz C, Blom E. Voice restoration with the Advantage tracheoesophageal voice prosthesis. Otolaryngology Head and Neck Surgery 2005; 133:681-684.
3. Hilgers F, Ackerstaff A, Balm A, VandenBrekel M, Tan B, Persson J. A New Problem-solving Indwelling Voice Prosthesis, Eliminating the Need for Frequent Candida- and ''Underpressure''-related Replacements: Provox ActiValve. Acta Otolaryngol 2003; 123:972-979.
4. Op de Coul BM, Hilgers FJ, Balm AJ, Tan IB, van den Hoogen FJ, van Tinteren H. A decade of postlaryngectomy vocal rehabilitation in 318 patients: a single Institution's experience with consistent application of provox indwelling voice prostheses. Archives of otolaryngology--head & neck surgery 2000; 126:1320-1328.
5. Hancock K, Houghton B, Van As-Brooks CJ, Coman W. First clinical experience with a new non-indwelling voice prosthesis (Provox NID) for voice rehabilitation after total laryngectomy. Acta Otolaryngol 2005; 125:981-990.
6. Doyle P, Eadie T. The perceptual nature of alaryngeal voice and speech. In: Doyle P, Keith R, eds. *Contemporary Considerations In The Treatment And Rehabilitation Of Head And Neck Cancer: Voice, Speech, And Swallowing*. Austin, Texas: Pro Ed, 2005:113-139.
7. Robbins J, Fisher HB, Blom EC, Singer MI. A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. J Speech Hear Disord 1984; 49:202-210.
8. Blood GW. Fundamental frequency and intensity measurements in laryngeal and alaryngeal speakers. J Commun Disord 1984; 17:319-324.
9. Pindzola RH, Cain BH. Duration and frequency characteristics of tracheoesophageal speech. The Annals of otology, rhinology, and laryngology 1989; 98:960-964.
10. Qi Y, Weinberg B. Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers. Journal of speech and hearing research 1995; 38:536.
11. Moerman M, Pieters G, Martens J, Borgt MVd, Dejonckere P. Objective evaluation of the quality of substitution voices. European archives of oto-rhino-laryngology 2004; 261:541.
12. vanGogh C, Joost M, Verdonck-deLeeuw Iet al. Acoustical analysis of tracheoesophageal voice. Speech communication 2005; 47:160-168.
13. van As-Brooks CJ, Koopmans-van Beinum FJ, Pols LC, Hilgers FJ. Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech. J Voice 2006; 20:355-368.
14. van As CJ, Hilgers FJ, Verdonck-de Leeuw IM, Koopmans-van Beinum F. Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. Journal of voice 1998; 12:239.

Effect of voice prostheses on voice

15.   Finizia C, Dotevall H, Lundstrom E, Lindstrom J. Acoustic and perceptual evaluation of voice and speech quality: a study of patients with laryngeal cancer treated with laryngectomy vs irradiation. Archives of otolaryngology--head & neck surgery 1999; 125:157-163.

16.   Omori K, Kojima H. Neoglottic vibration in tracheoesophageal shunt phonation. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 1999; 256:501-505.

17.   Stewart MG, Chen AY, Stach CB. Outcomes analysis of voice and quality of life in patients with laryngeal cancer. Archives of otolaryngology--head & neck surgery 1998; 124:143-148.

18.   Moerman M, Martens J, Dejonckere P. Application of the Voice Handicap Index in 45 patients with substitution voicing after total laryngectomy. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 2004; 261:423-428.

19.   Sewnaik A, van den Brink JL, Wieringa MH, Meeuwis CA, Kerrebijn JD. Surgery for recurrent laryngeal carcinoma after radiotherapy: partial laryngectomy or total laryngectomy for a better quality of life? Otolaryngol Head Neck Surg 2005; 132:95-98.

20.   Schuster M, Lohscheller J, Hoppe U, Kummer P, Eysholdt U, Rosanowski F. Voice handicap of laryngectomees with tracheoesophageal speech. Folia Phoniatr Logop 2004; 56:62-67.

21.   Kazi R, Cordova JD, Singh A et al. Voice-related Quality of Life in laryngectomees: assessment using the VHI and V-RQOL symptom scales. J Voice 2007; 21:728-734.

22.   Fung K, Lyden T, Lee J et al. Voice and swallowing outcomes of an organ-preservation trial for advanced laryngeal cancer. International Journal of Radiat Oncol Biol Phys 2005; 63:1395-1399.

23.   Evans E, Carding P, Drinnan M. The voice handicap index with post-laryngectomy male voices. International journal of language and communication disorders 2009; 44:575-586.

24.   Moukarbel RV, Doyle PC, Yoo JH, Franklin JH, Day AM, Fung K. Voice-related quality of life (V-RQOL) outcomes in laryngectomees. Head & neck 2011; 33:31-36.

25.   Oridate N, Homma A, Nakamaru Y et al. Voice-Related Quality of Life After Treatment of Laryngeal Cancer. Archives of otolaryngology--head & neck surgery 2009; 135:363.

26.   Blood GW. Development and Assessment of a Scale Addressing Communication Needs of Patients With Laryngectomies. Am J Speech Lang Pathol 1993; 2:82-90.

27.   Johansson M, Ryden A, Finizia C. Self evaluation of communication experiences after laryngeal cancer - a longitudinal questionnaire study in patients with laryngeal cancer. BMC Cancer 2008; 8:80.

28.   Evitts P, Kasapoglu F, Demirci U, SterneMiller J. Communication adjustment of patients with a laryngectomy in Turkey: Analysis by type of

Effect of voice prostheses on voice

surgery and mode of speech. Psychology Health and Medicine 2011; 16:650-660.

29. Kazi R, Singh A, De Cordova J, Clarke P, Harrington K, Rhys-Evans P. A new self-administered questionnaire to determine patient experience with voice prostheses (Blom-Singer valves). J Postgrad Med 2005; 51:253-258; discussion 258-259.

30. Carding P, Carlson E, Epstein R, Mathieson L, Shewell C. Formal perceptual evaluation of voice quality in the United Kingdom. Logoped Phoniatr Vocol 2000; 25:133-138.

31. Carding P, Wilson J, MacKenzie K, Deary I. Measuring voice outcomes: state of the science review. The Journal of laryngology and otology 2009; 123:823-829.

32. Eadie TL, Doyle PC. Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. Laryngoscope 2004; 114:753-759.

33. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. Laryngoscope 2004; 114:1549-1556.

34. Pauloski BR, Fisher HB, Kempster GB, Blom ED. Statistical Differentiation Of Tracheoesophageal Speech Produced under Four Prosthetic/Occlusion Speaking Conditions. J Speech Hear Res 1989; 32:591-599.

35. Delsupehe K, Zink I, St Lejaegere M, Delaere PMDP. Prospective Randomized Comparative Study of Tracheoesophageal Voice Prosthesis: Blom-Singer Versus Provox. The Laryngoscope 1998; 108:1561-1565.

36. Ward E, Hancock K, Lawson N, van As-Brooks CJ. Perceptual characteristics of tracheoesophageal speech production using the new indwelling Provox Vega voice prosthesis: a randomized controlled crossover trial. Head & neck 2011; 33:13.

37. Perry A. Vocal rehabilitation after total laryngectomy *Leicester School of Speech Pathology*. Leicester: De Montfort University, 1989:176.

38. Hurren A, Hildreth A, Carding P. Can we perceptually rate alaryngeal voice? Developing the Sunderland Tracheoesophageal Voice Perceptual Scale. Clinical otolaryngology : official journal of ENT-UK ; official journal of Netherlands Society for Oto-Rhino-Laryngology & Cervico-Facial Surgery 2009; 34:533-538.

39. Fairbanks D. Voice and Articulation Drill book. New York: Harper & Brothers, 1960.

Table 1: Differences between prostheses based on the median expert rating of each STOPS question.

	Blom Singer Duckbill	Blom Singer Low Pressure	Blom Singer Classic Indwelling	Blom Singer Advantage	Provox NID	Provox Vega	P (p <0.05)
Q1 Overall severity	1.33 (1.00-2.00)	1.67 (0.83-2.00)	1.33 (1.0-1.67)	1.33 (1.00-2.00)	1.33 (1.00-2.00)	1.33 (1.00-2.00)	0.97
Q3 Strain	1.16 (0.66-1.66)	1.0 (0.66-1.33)	1.0 (0.33-1.33)	1.00 (0.42-1.33)	1.0 (0.66-1.66)	1.00 (0.66-1.58)	0.72
Q4 Wetness/ Gurgliness	0.83 (0.33-1.66)	0.66 (0.33-1.33)	0.66 (0.33-1.33)	0.66 (0.33-1.33)	1.0 (0.33-1.66)	1.0 (0.66-1.58)	0.85
Q5 Impairment of volume	0.66 (0.33-1.0)	0.33 (0.0-1.33)	0.50 (0.00-1.33)	0.66 (0.08-1.25)	0.33 (0.00-1.00)	0.50 (0.33-1.25)	0.99
Q6 Social acceptability	1.33 (1.0-2.91)	1.33 (0.66-1.83)	1.00 (1.00-1.66)	1.33 (1.0-1.66)	1.33 (1.0-2.25)	1.33 (1.0-2.25)	0.66
Q7 Whisper	1.66 (0.00-0.66)	0.33 (0.00-1.00)	0.33 (0.00-0.66)	0.33 (0.00-0.92)	0.33 (0.00-1.00)	0.33 (0.00-0.66)	0.99
Q8 Intelligibility	0.66 (0.33-1.33)	0.33 (0.33-1.0)	0.33 (0.33-0.92)	0.66 (0.33-1.0)	0.66 (0.33-1.25)	0.66 (0.33-1.0)	0.83
Q9 Stoma noise	1.33 (1.0-2.0)	1.33 (1.0-1.66)	1.0 (1.0-1.66)	1.33 (!.0-2.0)	1.66 (1.0-2.0	1.33 (1.0-1.66)	0.72
Q10 Fluency	0.66 (0.33-1.0)	0.66 (0.16-1.00)	0.66 (0.08-1.0)	0.66 (0.83-1.0)	0.66 (0.33-1.25)	0.66 (0.33-1.33)	0.82
Q13 Accent	0.33 (0.83-1.25)	0.33 (0.33-1.5)	0.66 (0.33-1.50)	0.5 (0.33-1.0)	0.66 (0.33-1.58)	0.66 (0.33-1.66)	0.98
Q14 Poor reader	0.00 (0.00-0.33)	0.00 (0.00-0.33)	0.33 (0.00-0.33)	0.33 (0.00-0.33)	0.00 (0.00-0.33)	0.16 (0.00-0.33)	0.98

Table 2: **Frequency analysis of expert raters consensus of best prosthesis for voice**

Prosthesis	n	Frequency	Percentage of sample
BS Duckbill*	32	4	12.5%
BS Low pressure	41	5	12.1%
BS Classic indwelling	41	9	21.9%
BS Advantage	41	7	17%
Provox NID	40	5	8%
Provox Vega	35	3	8.5%
No consensus on best prosthesis	41	8	19.5%

* Not available in 20FG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 3: Median total scores for each prosthesis

Prosthesis	Median
BS Duckbill	20.5 (17.2-24.0)
BS Low pressure	15.0 (13.0-17.3)
BS Classic indwelling	13.0 (12.0-16.3)
BS Advantage	14.0 (13.0-17.3)
Provox NID	14.0 (10.0-16.0)
Provox Vega	14.0 (12.0-16.5)

Table 4: Frequency analysis of self-evaluation of best overall voice prosthesis

Prosthesis	n	Frequency	Percentage of sample
BS Duckbill	32	2	6.2%
BS Low pressure	41	6	14.6%
BS Classic indwelling	41	5	12.2 %
BS Advantage	41	5	12.2%
Provox NID	40	13	32.5%
Provox Vega	35	10	28.5%