DOI: 10.7860/JCDR/2018/36961.11741

Cepstral Peak Prominence of Blom-Singer 'Non-Indwelling Low Pressure' and 'Classic Indwelling' Voice Prosthesis in Male Tracheo-oesophageal Speakers

SHEELA SHEKARAIAH[1], VENKATARAJA U AITHAL[2], BELLUR RAJASHEKHAR[3], VASUDEVA GUDDATTU[4]

ABSTRACT

Introduction: Tracheo-oesophageal (TE) speech is one of the most commonly used alaryngeal voice restoration options for individuals who have undergone a total laryngectomy. Cepstral Peak Prominence (CPP) implies the overall voice quality objectively.

Aim: To compare the CPP between Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis across vowel phonation /a/ and text-reading tasks in male TE speakers.

Materials and Methods: The study included 10 male TE speakers in the age range of 45-75 years. Dr. Hillenbrand's 'Speech Tool' software was used to measure CPP. Two types of speech prosthesis used were Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis (both of 16 French diameter) and the tasks vowel phonation and text-reading were considered. The SPSS software, version 15.0 (SPSS Inc., Chicago, IL) was used for data analysis. The parametric paired t-test was applied to compare CPP between two types of voice prostheses across speech tasks.

Results: There was no significant difference between Blom-Singer non-indwelling low pressure voice prosthesis and classic indwelling voice prosthesis for CPP during vowel phonation/a/ task {t(9)=0.74, p=0.516} and text-reading task {t(9)=0.72, p=0.947)}. Similarly, there was no significant difference for CPP between speech tasks for Blom-Singer non-indwelling low pressure voice prosthesis {t(9)=1.11, p=0.347)} and Blom-Singer classic indwelling voice prosthesis {t(9)=0.51, p=0.644)}.

Conclusion: In TE speakers, CPP is not influenced by prosthesis type in terms of mode of fitting, prosthesis design and valve opening pressure across speech tasks, implying no difference in the objective measure of overall voice quality.

Keywords: Alaryngeal speech, Cepstral analysis, Neoglottis, Prosthetic voice

INTRODUCTION

Tracheo-oesophageal speech is one of the most commonly used alaryngeal voice modes for individuals undergoing total laryngectomy due to laryngeal cancer. Laryngeal cancer contributes to approximately 3-6% of all cancer in Indian males and only about 0.2-1% of all cancers in females (ICMR, 2013) [1]. In TE speakers, appropriate sized voice prosthesis is fitted into TE Puncture (TEP) for voice production [2]. The voice source (neoglottis) characteristics of TE speakers have been effectively studied and there is an increasing focus towards cepstral analysis of TE voice. Cepstral measure i.e., CPP is a more reliable measure of periodicity and implies on voice quality in an objective and quantifiable way [3-6]. In TE speakers, due to a periodic nature of the neoglottic vibration, CPP is more preferred to objectively quantify the overall voice quality [7,8].

There are differences present between Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis in terms of mode of fitting (non-indwelling versus indwelling), prosthesis design (size of the tracheal and oesophageal flanges, prosthesis tip and silicon grade used), valve opening pressure and cost of the voice prostheses. Valve opening pressure is reported to influence the device life [9]. In developing countries like India, cost is a major factor in patient decision of prosthesis selection, as most often, it is not covered under health insurance. Findings on voice quality differences in same TE speaker with two different voice prosthesis would substantiate the decision making during the selection of the voice prostheses for both clinicians and patients. It would provide additional information to the clinician justifying the cost of the device especially at the time of counselling. Hence, the objective of the current study was to compare the CPP between Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis across vowel phonation /a/ and text-reading tasks in male TE speakers.

MATERIALS AND METHODS

This is a comparative study with cross-over study design carried out at the Department of Speech and Hearing, School of Allied Health Sciences (SOAHS), Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India, between May 2014 to February 2018.

Participants: The clinical group consisted of 10 male TE speakers in the age range of 45-75 years. Participants who were referred from the head and neck surgeon were considered for the study. Participants who underwent total laryngectomy during two years time frame before the commencement of the study were considered. Participants who had undergone secondary TEP following total laryngectomy with catheter in place for 48 hours were referred to Speech-Language Pathologist (SLP). Patients with respiratory complaints, hearing problem, psychological and neurological problems, habits of cigarette smoking, alcohol intake, heavy exposure to chemicals and who have undergone chemotherapy/radiation post-total laryngectomy were excluded. The study was commenced after approval from Institutional Ethics Committee (IEC 245/2014). Prior to data collection, a signed 'consent form' and 'subject information sheet' were obtained.

Data collection: Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis were used in the study. Non-indwelling type requires removal, cleaning and re-insertion on a daily basis or at least every 3-4 days, by the patient or caregiver. The indwelling type of voice prosthesis is generally fitted by a trained SLP and does not require frequent removal and insertion, can stay in place for 3-6 months [10]. For all participants, first prosthesis to be fitted was selected on the basis of randomisation using 'research randomizer' programme. Participants could see the voice prosthesis, however they were blinded to the name, type of the prosthesis and the manufacturer. Each prostheses was fitted according to the manufacturer's instruction. Each participant made four visits. After 48 hours of secondary TEP with catheter, participants were subjected for data collection. One of the prostheses was fitted and after two days of speech stabilisation period, recording of vowel phonation/a/and text-reading were done. Following recording, the first prosthesis was removed and the catheter was inserted which was retained for next 48 hours. Second prosthesis was then fitted and again after two days of speech stabilisation period, recordings were repeated. Subsequent to each prosthesis fitting, participants were given a practice session on optimal digital occlusion of stoma for speaking. Two days of stabilisation period was given to ensure all participants who have acquired optimal occlusion to produce intelligible speech. Researcher who recorded the speech samples was blinded to both participant details and prosthesis.

Instrument used for recording: The participants were made to sit comfortably in a sound-treated room. Dr. Hillenbrand's 'Speech Tool' software (James Hillenbrand, Kalamazoo, MI) was used to measure CPP. It works on linear regression analysis which is an important step in calculating the CPP, thus limiting the usefulness of software programs that do not employ this step [4]. Unlike Computerised Speech Lab (CSL) algorithm, this software provides information on CPP with less time and cost. Speech samples were recorded with the sampling frequency of 44 kHz and 16-bits amplitude resolution by a Dynamic 'Shure' microphone, placed at a constant distance of 15 cm from the participants' mouth.

Speech tasks used for recording: Two speech tasks namely the vowel phonation and text-reading were considered. The vowel phonation task was selected as the software based automatic analysis relies mostly on sustained vowels and text recording was opted considering its relevance in terms of onsets, variation in the fundamental frequency (f0) and pauses. In vowel phonation task, each participant was instructed to sustain the vowel /a/ at a comfortable pitch and loudness for at least five seconds. Three such trials were recorded and the best trial was selected based on the visual inspection of waveform considering relatively steady phonation and no unvoiced component. For the analysis of data on vowel phonation /a/, stable middle 3 second of best three samples was considered. In text-reading task, the patient was instructed to read the English version of the Rainbow passage [11]. The second sentence of the rainbow passage "The rainbow is a division of white light into many beautiful colours" was considered for the analysis of text-reading data. This text was selected as it is commonly used in other studies and hence, makes the findings of the current study comparable to that of other studies [7,8].

STATISTICAL ANALYSIS

The SPSS software, version 15.0 (SPSS Inc., Chicago, IL) was used for data analysis. The CPP values of all 10 participants were averaged to calculate mean and SD with respect to type of voice prosthesis and task. Shapiro-Wilk test for normality revealed data to be normally distributed ($p \geq 0.05$). Therefore, the parametric 'Paired t-test' was used to compare CPP between two types of voice prostheses across speech tasks. Similarly, paired t-test was used to compare CPP between two types of speech tasks across voice prosthesis.

RESULTS

The results depicted in [Table/Fig-1] reveals Mean (M) and Standard Deviation (SD) of cepstral measure CPP for TE speakers' with Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis across speech tasks. There was no significant difference between Blom-Singer non-indwelling low pressure voice prosthesis and classic indwelling voice prosthesis for CPP {t(9)=0.74, p=0.516} during vowel phonation /a/ task. Similarly, there was no significant difference for CPP {t(9)=0.72, p=0.947)} between Blom-Singer non-indwelling low pressure voice prosthesis and classic indwelling voice prosthesis during text-reading task.

The results in [Table/Fig-2] reveal mean and SD of CPP for TE speakers' during vowel phonation /a/ and text-reading task across voice prostheses. For Blom-Singer non-indwelling low pressure voice prosthesis, there was no significant difference for CPP {t(9)=1.11, p=0.347)} between vowel phonation /a/ and text-reading task. Similarly, for Blom-Singer classic indwelling voice prosthesis, there was no significant difference for CPP {t(9)=0.51, p=0.644)} between speech tasks.

Cepstral parameter	Sample size (n)	Non-indwelling low pressure voice prosthesis (M±SD)	Classic indwelling voice prosthesis (M±SD)	t-value, (dof)	95% Confidence interval (CI) of the difference	p-value*
Vowel phonation /a/ task						
CPP (dB)	10	10.19±1.74	9.77±0.66	0.74 (9)	-1.40, 2.24	0.516
Text-reading task						
CPP (dB)	10	9.70±0.88	9.68±0.35	0.72 (9)	-0.97, 1.01	0.947

[Table/Fig-1]: Mean (M) and standard deviation (SD) of CPP for TE speakers with Blom-Singer non-indwelling low pressure voice prosthesis and classic indwelling voice prosthesis across vowel phonation /a/ and text-reading task.
Statistical test administered: Paired t-test; *p≥0.05

Cepstral parameter	Sample size (n)	Vowel phonation /a/ task (M±SD)	Text-reading task (M±SD)	t-value, (dof)	Confidence interval (CI) of the difference	p-value*
Non indwelling low pressure voice prosthesis						
CPP (dB)	10	10.19±1.74	9.70±0.88	1.11 (9)	-0.91, 1.88	0.347
Classic indwelling voice prosthesis						
CPP (dB)	10	9.77±0.66	9.68±0.35	0.51(9)	-0.47, 0.65	0.644

[Table/Fig-2]: Mean (M) and standard deviation (SD) of CPP for TE speakers during vowel phonation /a/ and text-reading task across Blom-Singer non-indwelling low pressure voice prosthesis and classic indwelling voice prosthesis.
Statistical test administered: Paired t-test; *p≥0.05

DISCUSSION

The objective of the study was to compare the CPP between Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis across vowel phonation /a/ and text-reading tasks in male TE speakers. The cepstral analysis revealed no significant differences between Blom singer non-indwelling low pressure voice prosthesis and classic indwelling type across vowel phonation and text-reading task. This can be attributed to the controlled conditions of shaft diameter of the voice prosthesis (air entry effect) and its 'valve type' (air exit effect) [7,12]. In 'air entry' effect, as airflow increases, resistance of the voice prosthesis also increases. The shaft diameter of the voice prosthesis {16 French (Fr)} remained same in all participants. In 'air exit effect', as airflow increases, resistance decreases and valve opens suddenly. In addition to the 'air entry' and 'air exit effect', both the types of voice prosthesis consisted of 'hinged (flapper) valve' design offering less resistance to the flow of air.

With the purview of CPP being the robust indicator of perceptual voice quality [4,13-15], current findings suggest that in spite of the difference in the design of the selected voice prosthesis (size of the tracheal and oesophageal flanges; prosthesis tip; silicon grade used; valve opening pressure), has no influence on the overall voice quality. In addition, the voice quality differences were not evident when TE speakers changed voice prosthesis alternatively from low pressure to classic indwelling voice prosthesis and vice versa.

In a study with subjects having various organic and functional dysphonia, text-reading task was recommended, as it outperforms both perturbation and cepstral based vowel analysis [16]. The findings of the current study suggest that TE speakers show no significant differences in voice quality between speech tasks across prostheses. CPP being the reliable measure of periodicity [3-6], the lack of difference between speech tasks across prostheses in the current study can be attributed to the inherent instability of PE segment [17]. This is also in line with the findings of the previous study [18] reporting TE speakers to produce only small variations in fundamental frequency (f0), with difficulty to control their speech.

The mean CPP value of 16 Fr Blom-Singer classic indwelling voice prosthesis for text-reading task is 9.68±0.35 and it is in consonance with the previous study [7] where CPP of continuous speech (text-reading) using indwelling Provox voice prosthesis was 11.45±2.08. Similarly, another study [4] reported CPP of continuous speech text-reading as 10.48±1.40. This reflects similar CPP in indwelling voice prostheses for continuous speech (text-reading) task irrespective of change in shaft diameter. Thus, it can be speculated that the dynamic changes in vocal tract configuration in continuous speech task (text-reading) obscures the effect of shaft diameter on the PE segment vibration, resulting in similar CPP values.

In vowel phonation task, median CPP value of 21.40 for vowel /a/ is reported to be highest in TE speakers using Provox voice prosthesis (Atos Medical, Sweden) [19]. In contrast, mean CPP value of 9.77±0.66 was observed for Blom-Singer classic indwelling voice prosthesis in the current study and found to be lowest. This can be attributed to the amount of airflow passing through the voice prosthesis of 16 Fr diameter as against ≥20 Fr in an earlier study [19]. The 16 Fr shaft relatively restricts the amount of airflow passed through the voice prosthesis, adding onto the inherent instability of the PE segment vibration during voicing, resulting in lesser CPP values. This reflects that the difference in cepstral measure in Blom-Singer classic indwelling voice prostheses for vowel task primarily depends on the shaft diameter of the voice prosthesis.

SD of the CPP values of Blom-Singer low pressure and classic indwelling voice prosthesis for vowel phonation and text-reading task are lesser than half of the mean indicating lesser variation from the measured mean. These findings are in consonance with the previous studies [7,8] where small variation is observed for SD of CPP of indwelling voice prosthesis using text-reading task. This can be attributed to the common factors in terms of recording procedure (Speech tool), type of speech task (text-reading), type of voice prosthesis (indwelling) and type of valves (flapper valve) used across the studies. This suggests a consistency in the measured mean CPP values of selected voice prostheses across selected speech tasks in the current study.

In the current study, the difference in CPP is not observed between voice prosthesis of 16 Fr across vowel phonation /a/ and text-reading tasks. Unlike in laryngeal speakers, the vowel as well as text-based cepstral analysis provides similar findings of voice quality in TE speakers using 16 Fr selected voice prostheses. The vowel phonation task or text-reading task is equally recommended in objectively evaluating the overall voice quality using CPP. However, more studies are needed to investigate task specific differences in laryngectomees using prosthesis of larger diameter.

Further, the influence of the software used to record and analyse CPP values, and their reliability cannot be overlooked, with its limitations of not being based on linear regression algorithms. This algorithm allows the researcher to compare between speech signals without any artefacts that are caused by differences in recording levels, instrumentation or loudness of the native voice signals [14,15]. In comparison to earlier findings in TE speakers [7,19], reduced standard deviation of CPP in the current study reflects the good stability of voice production with Blom-Singer low pressure and classic indwelling voice prosthesis during both the speech tasks.

LIMITATION AND FUTURE DIRECTIONS

In the current study only phonation of /a/ with comfortable loudness was considered. Hence, to validate the findings, future research is recommended to consider other vowels. Objective monitoring of vocal intensity in TE speakers was also not part of the current study. This however, could offer insights into the influence of intensity on CPP.

Other measures such as perceptual evaluation, maximum vowel phonation duration, speaking rate, voice onset time and long-term average spectrum need to be investigated for the comprehensive understanding of TE speech characteristics associated with different types of voice prosthesis. Future research could attempt a similar investigation of cepstral measure using other brands of prosthesis having different design, diameter and defined valve opening pressure to identify any task specific differences among TE speakers.

CONCLUSION

The current study is first of its kind to compare the CPP between Blom-Singer non-indwelling low pressure and classic indwelling voice prosthesis across speech tasks in male TE speakers. Findings suggest that the voice quality implied using CPP is not influenced by type of voice prosthesis in terms of mode of fitting (non-indwelling versus indwelling), prosthesis design (size of the tracheal and oesophageal flanges, prosthesis tip and silicon grade used) and valve opening pressure across speech tasks. Hence, clinician can suggest either of the prostheses of similar diameter for laryngectomees opting the TE mode of communication considering their affordability, where Blom-Singer classic indwelling type is triple the cost of non-indwelling low pressure type of voice prosthesis.

ACKNOWLEDGEMENTS

Sincere thanks to Cancer Aid and Research Foundation (CARF) and 'Prof. Bellur Rajashekhar, Chair in Head and Neck Cancer Research and Voice Restoration', Manipal Academy of Higher Education (MAHE), Manipal, India for financial support.

REFERENCES

[1] ICMR. Individual registry-wise annexure. In: Three-year Report of Population Based Cancer Registries 2009-2011. Bangalore: National Cancer Registry Program (ICMR); 2013. pp. 92-150.
[2] Hillman RE, Walsh MJ, Wolf GT, Fisher SG, Hong WK. Functional outcomes following treatment for advanced laryngeal cancer. Part I-Voice preservation in advanced laryngeal cancer. Part II-Laryngectomy rehabilitation: the state of the art in the VA System. Annals Oto, Rhin & Lary, Supp. 1998;172:01-27.
[3] Hillenbrand JM. A methodological study of perturbation and additive noise in synthetically generated voice signals. J Speech Hear Res. 1987;30:448-61.
[4] Heman-Ackah YD. Reliability of calculating the cepstral peak without linear regression analysis. J Voice. 2004;18:203-08.
[5] Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. J Sp Hrg Res. 1994;37(4):769-78.
[6] Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. J Speech Hear Res. 1996;39:311-21.
[7] Maryn Y, Dick C, Vandenbruaene C, Vauterin T, Jacobs T. Spectral, cepstral, and multivariate exploration of tracheoesophageal voice quality in continuous speech and sustained vowels. The Laryngoscope. 2009;119(12):2384-94.
[8] Nagle K, Eadie T. Determining time-and frequency-based acoustic correlates of listener effort in tracheoesophageal speech. In 10th International Conference Advances in Quantitative Laryngology. 2013; pp. 99.
[9] Kress P, Schäfer P, Schwerdtfeger FP, Rösler S. Are modern voice prostheses better? A lifetime comparison of 749 voice prostheses. European Archives of Oto-Rhino-Laryngology. 2014;271(1):133-40.
[10] Rajashekhar B, Harrington KJ, Kazi R. Voice Restoration after total laryngectomy. Delhi, India: Byword Books private limited, 2009.
[11] Fairbanks G. Voice and Articulation Drillbook. 1960.
[12] Weinberg B, Moon J. Aerodynamic properties of four tracheoesophageal puncture prostheses. Archives of Otolaryngology. 1984;110(10):673-75.
[13] Heman-Ackah YD, Michael DD, Baroody MM, Ostrowski R, Hillenbrand J, Heuer RJ, et al. Cepstral peak prominence: A more reliable measure of dysphonia. Annals of Otology, Rhinology & Laryngology. 2003;112(4):324-33.
[14] Heman-Ackah YD, Michael DD, Goding GS. The relationship between cepstral peak prominence and selected parameters of dysphonia. Journal of Voice. 2002;16(1):20-27.
[15] Zieger K, Schneider C, Gerull G, Mrowinski D. Cepstrum analysis in voice disorders. Folia Phoniatr Logop. 1995;47:210-17.

[16] Moers C, Möbius B, Rosanowski F, Nöth E, Eysholdt U, Haderlein T. Vowel- and text-based cepstral analysis of chronic hoarseness. Journal of Voice. 2012;26(4):416-24.
[17] Yan N, Ng ML, Wang D, Zhang L, Chan V, Ho RS. Nonlinear dynamical analysis of laryngeal, esophageal, and tracheoesophageal speech of Cantonese. Journal of Voice. 2013;27(1):101-10.
[18] Schindler A, Canale A, Cavalot AL, Albera R, Capaccio P, Ottaviani F, et al. Intensity and fundamental frequency control in tracheoesophageal voice. Acta Otorhinolaryngologica Italica. 2005;25(4):240.
[19] Clapham RP, As-Brooks CJ, Brekel MW, Hilgers FJ, Son RJ. Automatic tracheoesophageal voice typing using acoustic parameters. In Interspeech. ISCA. 2013;14:2162-66.

PARTICULARS OF CONTRIBUTORS:
1. Assistant Professor, Department of Speech and Hearing, School of Allied Health Sciences (SOAHS), Manipal Academy of Higher Education, Manipal, Karnataka, India.
2. Professor, Department of Speech and Hearing, School of Allied Health Sciences (SOAHS), Manipal Academy of Higher Education, Manipal, Karnataka, India.
3. Professor, Department of Speech and Hearing, School of Allied Health Sciences (SOAHS), Manipal Academy of Higher Education, Manipal, Karnataka, India.
4. Associate Professor, Department of Statistics, Manipal Academy of Higher Education, Manipal, Karnataka, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:
Ms. Sheela Shekaraiah,
Assistant Professor, Department of Speech and Hearing, School of Allied Health Sciences (SOAHS),
Manipal Academy of Higher Education, Manipal-576104, Karnataka, India.
E-mail: sheela.mslp@gmail.com; sheela.s@manipal.edu

FINANCIAL OR OTHER COMPETING INTERESTS: As declared above.

Date of Submission: **Apr 28, 2018**
Date of Peer Review: **May 18, 2018**
Date of Acceptance: **May 25, 2018**
Date of Publishing: **Jun 01, 2018**