

# Cepstral Analysis of Voice in Patients With Thyroidectomy

Yu Jeong Shin<sup>1</sup> · Ki Hwan Hong<sup>2</sup>

<sup>1</sup>Department of Speech-Language Therapy, Howon University, Gunsan; <sup>2</sup>Department of Otolaryngology-Head and Neck Surgery, Research Institute of Clinical Medicine-Biomedical Research Institute of Chonbuk National University Hospital, Jeonju, Korea

**Objectives.** The vocal changes after a thyroidectomy are temporary and nonsevere, therefore, obtaining accurate analytical results on the pathological vocal characteristics following such a procedure is difficult. For a more objective acoustic analysis, this study used the cepstral analysis method to examine changes in the patients' voices during the perioperative period regarding sustained vowel phonation.

**Methods.** The sustained phonation of the five vowels (i.e., /a/, /e/, /i/, /o/, and /u/) by 35 patients with thyroidectomy were recorded by using a Multi-Speech program. Of the 35 patients, 10 were men and 25 were women, with an average age of 51.5 years. Voice data were collected a total of 3 times (preoperatively, 5–7 days after the operation, and 6 weeks after the operation) and were edited according to each fragment (on-set, mid, and off-set) for cepstral analysis.

**Results.** The cepstral analysis on the patients' voices revealed no significant differences between the examination periods of all vowel phonations. However, analysis of the on-set fragment of the vowel /i/ revealed pathological characteristics in which the cepstral measurements of the voice were significantly lower after the operation than before the operation, with the cepstral measurements of the voice increasing further 6 weeks following surgery.

**Conclusion.** The results of the acoustic analysis on the on-set fragment of the vowel /i/ will be important data for characterizing the vocal changes during the perioperative period. This study contributes to future research on the mechanisms underlying changes in the voice of patients with a history of thyroid or neck surgery.

**Keywords.** *Thyroidectomy; Cepstral Analysis; Voice*

## INTRODUCTION

Recently, thyroidectomies are being performed more frequently and the importance of managing and treating vocal changes after thyroidectomy are being emphasized due to an increasing trend in the incidence of thyroid diseases. Thyroidectomy affects the steadiness of the vocal tract [1], which may lead to vocal

changes, and the characteristics of vocal changes in patients who have undergone thyroidectomy mainly include symptoms such as reduced vocal range, vocal fatigue, reduced speaking fundamental frequency, hoarseness, and reduced vocal strength [2].

In contrast to vocal cord disorders, such as vocal nodules or vocal polyps, vocal changes after an operation such as thyroidectomy occur because of a complex interaction between several factors including damage to the nerves and muscles during surgical procedures and psychological issues in individual patients [3]. Neck muscles may be damaged due to the lateral traction used during the thyroid surgery or cutting of these muscles, or wound contracture with surrounding structures after surgery [4,5]. Neck muscles participate indirectly or directly in the functioning of the larynx and their phonatory function is well-defined. The function of the extrinsic laryngeal muscles, the so-

• Received February 5, 2015

Revised March 4, 2015

Accepted March 18, 2015

• Corresponding author: **Ki Hwan Hong**

Department of Otolaryngology-Head and Neck Surgery, Chonbuk National University Hospital, Chonbuk National University Medical School, 20 Geonji-ro, Deokjin-gu, Jeonju 54907, Korea

Tel: +82-63-250-1990, Fax: +82-63-250-1986

E-mail: khhong@chonbuk.ac.kr

Copyright © 2016 by Korean Society of Otorhinolaryngology-Head and Neck Surgery.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

called external frame function, lengthens or shortens the vocal folds, changing the relationship of the thyroid to the cricoid cartilage. Hong et al. [4] reported that the contraction of the sternohyoid and sternothyroid muscles causes a laryngotracheal downward pull, changing air volume in the subglottic air space. This downward pull shortens the cricothyroid distance, and anterior downward bending shortens the anterior cricothyroid distance, which in turn lengthens the vocal folds and raises the frequency of the sound emitted. Hong et al. [5] suggested the importance of the extralaryngeal frame function to the pitch control in patients with total thyroidectomy. The authors reported that the phonation time and fundamental frequency were not changed after surgery, but that the speaking fundamental frequency (Sf0), range of Sf0, and vocal range might be diminished after surgery. They suggested that the cause of voice dysfunction is not seen in a neural lesion, but in a disturbance of the extralaryngeal frame function.

Therefore, objective and accurate evaluation and analysis of the patient's vocal state are required in patients who undergo thyroidectomy procedures. In addition, because vocal changes after thyroidectomy procedures tend to resolve over time in most cases, the performance of tracking studies on patients' vocal characteristics pre- and postoperation and over time is required. Numerous studies have investigated test methods for analyzing the voices of patients with voice disorders, including thyroidectomy patients; however, regardless of whether an objective or subjective test is used, producing reliable and consistent test results can be difficult [6-8]. Fundamental frequency changes must be accurately measured and estimated when obtaining jitter, shimmer, and noise-to-harmonic ratio measurements, which are acoustic vocal tests frequently used in clinical studies and research [6]. However, the periodicity becomes lower as the severity of the voice disorder increases, thereby making it more difficult to perform the measurement and lowering the reliability of the test results. The 'cepstrum' is a voice analysis method that has been proposed as an alternative to this problem. The 'cepstrum' is a Fourier transform of the logarithm power spectrum of an acoustic signal and was originally described by Noll (1964) as a procedure for extracting the fundamental frequency from spectrum of a sound wave [9,10].

In this study, the cepstral analysis method, which is not associated with the drawbacks of the conventional method and improves the reliability of analytic results, was utilized to analyze patients' voices. In addition, because the measurements can vary depending on the voice types of the subjects in objective voice analysis [6], we aimed to closely investigate the acoustic characteristics of the phonation of patients with voice disorders by analyzing subdivided vocal data via utilization of data from each section of sustained vowel phonation. The goal of this study was to provide useful data on voice management and analysis of thyroidectomy patients' phonation.

## MATERIALS AND METHODS

The study participants included 35 patients (10 men and 25 women) who had received a diagnosis of thyroid cancer or thyroid tumor and had undergone a thyroidectomy. All subjects had no laryngeal pathology as determined via laryngoscopy before surgery and no reported laryngeal nerves injury or damage due to intubation on the day of surgery. Every time the patients visit the clinic after discharge, laryngoscopy was performed and all patients were normal. The subjects' ages ranged from 19 to 74 years (average age, 51.5 years); 28 patients had undergone total a thyroidectomy and seven patients had undergone a lobectomy.

For the acoustic analysis, the Multi-Speech Model 3700 (Kay-pentax, Lincoln Park, NJ, USA) was used for collecting and analyzing voice data for cepstral analysis. The sampling rate was 11,025 Hz, and the AKG C420 Headset Condenser Microphone (AKG, Vienna, Austria) was used. A speech language pathologist instructed the patients to phonate as comfortably as possible in their normal voice. The patients subsequently followed the instructions of the pathologist to phonate the 5 vowels (/a/, /e/, /i/, /o/, and /u/) for approximately five seconds while sitting in a chair equipped with a headset microphone. The recordings were conducted three times (preoperatively, 5-7 days after the operation, and 6 weeks after the operation) in an identical manner. The cepstral measurements for the five second recordings of the entire fragment were obtained by analyzing all vowels. Each cepstral measurement was derived after the phonation of the vowels were edited into 1 second fragments of on-set, mid, and off-set in accordance with the analysis time (1 second) described in previous studies [7,8]. The cepstrum was calculated after converting the edited fragments into a power spectrum and subsequently processing the data through the fast-Fourier transform process [10,11] (Fig. 1).

A statistical program, IBM SPSS ver. 18.0 (IBM Co., Armonk, NY, USA), was used to perform repeated measures analysis of variance (RM-ANOVA) based on a 95% confidence level ( $P < 0.05$ ) in order to identify the difference between the examination periods of the voices collected from the 35 patients over three examinations. Bonferroni method was used to conduct *post hoc* analysis on results that showed a significant difference.

## RESULTS

The performance of RM-ANOVA by using the preoperative and postoperative cepstral measurements from all sustained vowels (/a/, /e/, /i/, /o/, and /u/) collected from the patients revealed no significant differences ( $P > 0.05$ ) in the vocal examination of all 5 vowels between the periods before and after the operation.

Further, RM-ANOVA was conducted on the 15 voice data points per patient (the 5 vowels collected during the 3 examinations) by subdividing the data into 3 fragments (1 second of

Table 1. Cepstral measurements of each vowel according to fragment by period

Vowel	On-set			Mid			Off-set		
	Preoperative	After 5-7 Days	After 6 Weeks	Preoperative	After 5-7 Days	After 6 Weeks	Preoperative	After 5-7 days	After 6 weeks
/a/	6.045±0.382	5.925±0.404	5.995±0.375	5.790±0.348	5.753±0.388	5.765±0.360	5.757±0.347	5.701±0.370	5.707±0.390
/e/	6.080±0.399	5.986±0.497	5.954±0.501	5.924±0.359	5.828±0.411	5.811±0.417	5.858±0.357	5.707±0.503	5.775±0.416
/i/	5.852±0.284*	5.583±0.465*	5.703±0.360*	5.661±0.299	5.537±0.334	5.547±0.387	5.628±0.314	5.463±0.432	5.583±0.352
/o/	5.455±0.292	5.355±0.492	5.378±0.354	5.299±0.321	5.218±0.413	5.169±0.548	5.312±0.347*	5.094±0.409*	5.211±0.329*
/u/	5.473±0.430	5.351±0.448	5.398±0.491	5.372±0.330	5.283±0.442	5.286±0.415	5.275±0.387	5.199±0.507	5.281±0.379

Values are presented as mean±SD.

\*Significant difference by period ( $P<0.05$ , repeated measures analysis of variance).

Table 2. Repeated measures analysis of variance results of cepstral measurements of on-set fragments of vowels by period

	Vowel				
	/a/	/e/	/i/	/o/	/u/
F	0.805	0.742	4.634	0.658	0.813
P	0.452	0.480	0.013*	0.521	0.448

\* $P<0.05$ .

Table 3. Cepstral measurements of on-set fragment on vowel /i/ by period

Period	Mean±SD
Preoperative	5.852±0.284
After 5-7 days	5.583±0.465
After 6 weeks	5.703±0.360

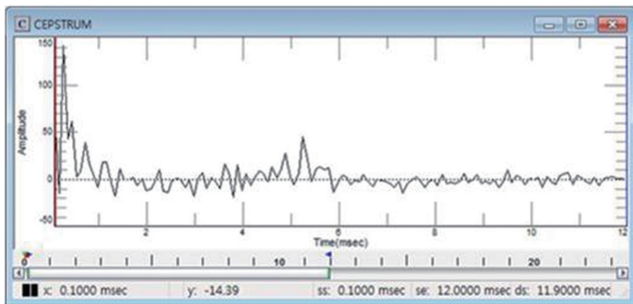


Fig. 1. An example of cepstral analysis by using the Multi-Speech program.

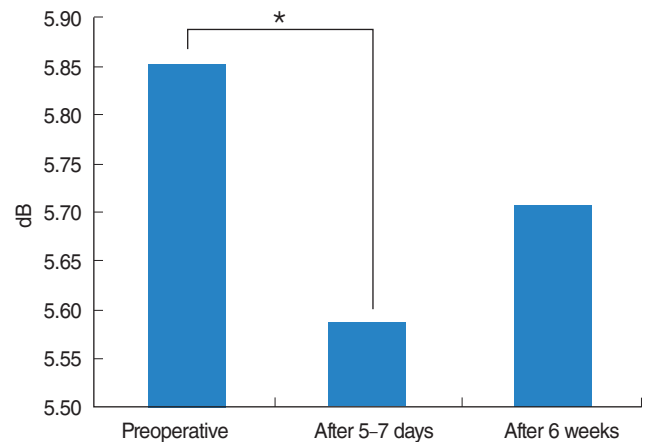


Fig. 3. Bonferroni *post hoc* result on on-set fragment of vowel /i/. \*Significant difference ( $P<0.05$ ).

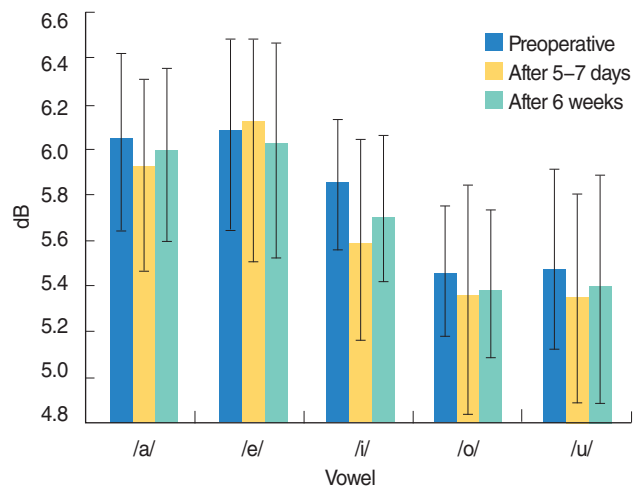


Fig. 2. Cepstral measurements of on-set fragment on each vowel by period.

each: on-set, mid, and off-set) in order to determine whether a difference existed between the cepstral measurements of the on-set, mid, and off-set phonation voice data for each vowel. The results revealed significant differences ( $P<0.05$ ) in the cepstral measurements for on-set, mid, and off-set fragments in 13 vowel phonations, excluding the vowel /i/ immediately after the operation and the vowel /u/ 6 weeks after the operation.

To examine the acoustic characteristics of the vowel phonation by each fragment according to the examination period based on the aforementioned results, the collected vowel phonation data were divided into the on-set, mid, and off-set fragments, and RM-ANOVA was performed on each set. Significant differences were found in the preoperative and postoperative periods for the on-set fragment of the vowel /i/ ( $F(2, 68)=4.634$ ,  $P<0.05$ ) (Tables 1, 2, and Fig. 2). *Post hoc* analysis with Bonferroni method to compare the examination periods revealed that

the average cepstral value immediately after the operation ( $5.583 \pm 0.079$ ) was significantly lower than that before the operation ( $5.852 \pm 0.048$ ) for the on-set fragment of the vowel /i/; further, although this value was slightly elevated 6 weeks after the operation ( $5.703 \pm 0.061$ ), a comparison of the preoperative and postoperative values showed no significant difference (Table 3 and Fig. 3).

The cepstral measurements for the mid fragment of vowel phonation indicated that no vowels were significantly different in relation to the examination period, whereas the vowel /o/ showed a significant difference in the off-set phonation fragment ( $F(2, 68) = 3.686, P < 0.05$ ).

## DISCUSSION

The results of cepstral analysis revealed no significant acoustic differences in the voices of the thyroidectomy patients overall between the preoperative and postoperative assessments. Many factors affect the voice of patients after a thyroidectomy, including damage to the recurrent laryngeal nerve and the superior laryngeal nerve, changes in muscle movement due to damage to the cricothyroid muscle or outer larynx, mucosal damage due to tracheal intubation, and fibrosis caused by the wound healing process after the operation [2,4]. If nerve damage does not occur during the thyroidectomy process, fibrosis or minor damage to the laryngeal muscles and mucous membrane have insignificant effects on the acoustic voice analysis; further, since these problems resolve naturally over time, the discomfort of the phonation gradually disappears.

The majority of previous studies focused on the entire fragment from on-set to off-set phonation, or on the mid fragment of /a/, as a means of analyzing the voice data. In contrast, the vowels were divided into different fragments in this study, and this resulted in the notable finding of a significant difference in the on-set fragment of the vowel /i/ ( $P < 0.05$ ) and the off-set fragment of the vowel /o/ ( $P < 0.05$ ) relative to the period of examination. The cepstral measurements for the immediate postoperative period for the on-set fragment of the vowel /i/ were significantly lower ( $P < 0.05$ ) as compared with the values before the operation, whereas the cepstrum showed a tendency to increase again at 6 weeks after the operation. This is consistent with the results of previous studies [2,12-16], which proposed that pathological changes in thyroid patients' voices primarily occur immediately after an operation and are characteristically temporary, with recovery over time. In particular, the fundamental frequency of the thyroidectomy patient's voice clearly declines during the immediate postoperative period [12,17], and the cepstral analysis, which is based on the exact calculation of the fundamental frequency, adequately reflects this characteristic. This suggests that this decline in the fundamental frequency after the operation is due to the inability of the cricothyroid

muscle, which affects the length adjustment of the vocal cords, to function properly with regard to muscle tension after surgery [18].

The cepstral measurement of the on-set of vowel /i/ 6 weeks after the operation showed a slight tendency to increase without significant differences in perioperative measurements, which is presumably related to the voice recovery period of the thyroidectomy patient. Several studies that analyzed the vocal conditions of patients at more than one month after the operation reported that the voice recovers its normal state within 3 months after the operation [14,15]. In contrast, other studies suggest that vocal changes remain for three months [19] or even six months [12] after the operation. In the present study, we focused on measuring patients immediately after and 6 weeks after the operation for postoperative voice analysis. Thus, we presumed that the cepstral measurements showed ascending values and that the statistical analysis did not indicate a significant difference between the examination periods because the recovery process was already present in the patients 6 weeks after the operation. Furthermore, vocal restoration after a thyroidectomy begins shortly after the operation, and rapid recovery at the initial stages after the operation is expected.

Moreover, the reason that the vowel /i/ was the only vowel to show a significant difference according to the examination periods in cepstral analysis is based on its acoustic characteristics. Differences in the height of the tongue and in the movement of the vocal organs exist when a person phonates a vowel, which in turn affects the location of the laryngeal cartilage, glottis, and vocal tract, and the tension of the vocal cords [20]; therefore, different acoustic characteristics can be present even in the same person. More noise is found on the spectrum in high vowels, such as /i/, compared to low vowels, such as /a/ [21], presumably because the location of the tongue and jaw are lowered during the phonation of a low vowel, such as /a/, which reduces oral airway resistance [22]. In contrast, the high vowel /i/ is a closed vowel that is pronounced in a state in which the surface of the tongue almost touches the roof of the mouth, increasing oral airway resistance, so it can be more difficult to phonate compared to the low vowels. Therefore, the significant difference in relation to the examination period during the cepstral analysis may have been caused by the relative nose output when the vowel /i/ is pronounced.

As shown in both previous studies and this study, the selection of vowels for the voice analysis affects the acoustic measurements [23,24]; therefore, a sufficient amount of information and planning is required when determining which vowels to use in research.

Targeting the stable period of the sustained vowel phonation is universally used for the efficiency and reliability of the analysis in clinical and research fields. When applying the acoustic analysis method on sustained vowel phonation, this study divided the vowel phonation into on-set, mid, and off-set fragments

for analysis, rather than focusing on the entire length of the vowel phonation data or the stable middle section. A significant difference between the on-set, mid, and off-set fragments within a single vowel phonation data set forms the basis for considering which section should be selected for the voice analysis. This is because the acoustic characteristics of vowel phonation continuously change during even a short period, and each period may have different characteristics. Dividing the sustained vowel phonation into three fragments showed a significant ( $P < 0.05$ ) difference in the cepstral measurement in relation to the examination period for the on-set fragment of the /i/ vowel phonation.

A previous study [25] that examined the relationship between acoustic analysis and auditory prospection analysis by dividing the vowel into 3 fragments (on-set, mid, and off-set; each for 200 msec) reported that roughness, a category for sound quality assessment, is related to the nonperiodicity and instability of the voice signal. Further, the study indicated that the on-set fragment of the vowel has a higher explanatory power compared to the midsection of the vowel or the entire vowel. Because the study analyzed nine different types of voice disorders rather than only one type of voice disorder, it can be said that the roughness characteristic occurs in the on-set of vowel articulation in terms of auditory-perceptual evaluation. In particular, because the phonation on-set of voice disorder patients may result in inappropriate vocal cord contact when compared to normal phonation, starting the phonation may become more difficult than retaining the phonation after articulation as the severity of the voice disorder increases. Thus, the on-set fragment of vowel phonation appears to reveal unstable vocal characteristics, and the characteristics differ according to the disorder.

In the case of vowel off-set fragments, the glottal fry is often computed as the vocal cord vibration slows down with the end of phonation, which affects the periodicity and fundamental frequency of the vibration movement [26,27]; therefore, this process requires attention during analysis. In this study, significant differences ( $P < 0.05$ ) were observed in the off-set of the vowel /o/; however, given that this is the end fragment of the 5 seconds of vowel phonation, determining whether it is a consistent pathological characteristic of the patients' voice would be difficult.

We anticipate that the results of this study will serve as a reference for selecting the type of voice data that should be used in both subjective and objective examinations of pathological voice disorders going forward. In particular, regarding thyroidectomy patients, voice analysis performed on the on-set fragment of the vowel /i/ may yield results that are more effective. Further, given the fact that vocal cord movement is best visualized in the high vowel /i/ when the glottal gesture is observed using the laryngeal endoscope, such an analysis will be able to provide a more efficient voice evaluation.

In conclusion, the results of the cepstral analysis performed on the voices of patients who underwent thyroidectomy procedures indicated pathological characteristics. Although significant

differences were not observed in the phonation of the entire vowel according to the examination periods, cepstral measurements of the voices were significantly lower in the on-set fragment of the vowel /i/ immediately after the operation compared to those before the operation, with an additional increase observed 6 weeks after the operation.

In addition, because measures varied according to the vowel fragments (on-set, mid, and off-set) used in voice analysis, appropriate voice data representing the characteristics of the voice disorder must be carefully selected for analysis.

## CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

This paper was supported by Research Institute of Clinical Medicine-Biomedical Research Institute of Chonbuk National University Hospital.

## REFERENCES

1. Timon CI, Hirani SP, Epstein R, Rafferty MA. Investigation of the impact of thyroid surgery on vocal tract steadiness. *J Voice*. 2010 Sep;24(5):610-3.
2. Hong KH, Kim YK. Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. *Otolaryngol Head Neck Surg*. 1997 Oct;117(4):399-404.
3. Grover G, Sadler GP, Mihai R. Morbidity after thyroid surgery: patient perspective. *Laryngoscope*. 2013 Sep;123(9):2319-23.
4. Hong KH, Ye M, Kim YM, Kevorkian KF, Berke GS. The role of strap muscles in phonation: in vivo canine laryngeal model. *J Voice*. 1997 Mar;11(1):23-32.
5. Hong KH, Yang YS, Lee HD, Yoon YS, Hong YT. The effect of total thyroidectomy on the speech production. *Clin Exp Otorhinolaryngol*. 2015 Jun;8(2):155-60.
6. Wolfe V, Martin D. Acoustic correlates of dysphonia: type and severity. *J Commun Disord*. 1997 Sep-Oct;30(5):403-15.
7. Wolfe V, Fitch J, Cornell R. Acoustic prediction of severity in commonly occurring voice problems. *J Speech Hear Res*. 1995 Apr;38(2):273-9.
8. Martin D, Fitch J, Wolfe V. Pathologic voice type and the acoustic prediction of severity. *J Speech Hear Res*. 1995 Aug;38(4):765-71.
9. Awan SN, Helou LB, Stojadinovic A, Solomon NP. Tracking voice change after thyroidectomy: application of spectral/cepstral analyses. *Clin Linguist Phon*. 2011 Apr;25(4):302-20.
10. Noll AM. Short-time spectrum and 'cepstrum' techniques for vocal-pitch detection. *J Acoust Soc Am*. 1964;36:296-302.
11. Noll AM. Cepstrum pitch determination. *J Acoust Soc Am*. 1967 Feb;41(2):293-309.
12. Debryne F, Ostyn F, Delaere P, Wellens W, Decoster W. Temporary voice changes after uncomplicated thyroidectomy. *Acta Otorhino-*

- laryngol Belg. 1997 Feb;51(3):137-40.
13. Kuhn MA, Bloom G, Myssiorek D. Patient perspectives on dysphonia after thyroidectomy for thyroid cancer. *J Voice*. 2013 Jan;27(1):111-4.
  14. Lee J, Na KY, Kim RM, Oh Y, Lee JH, Lee J, et al. Postoperative functional voice changes after conventional open or robotic thyroidectomy: a prospective trial. *Ann Surg Oncol*. 2012 Sep;19(9):2963-70.
  15. Li C, Tao Z, Qu J, Zhou T, Xia F. A voice acoustic analysis of thyroid adenoma patients after a unilateral thyroid lobectomy. *J Voice*. 2012 Jan;26(1):e23-6.
  16. Van Lierde K, D'Haeseleer E, Wuyts FL, Baudonck N, Bernaert L, Vermeersch H. Impact of thyroidectomy without laryngeal nerve injury on vocal quality characteristics: an objective multiparameter approach. *Laryngoscope*. 2010 Feb;120(2):338-45.
  17. Nam IC, Bae JS, Chae BJ, Shim MR, Hwang YS, Sun DI. Therapeutic approach to patients with a lower-pitched voice after thyroidectomy. *World J Surg*. 2013 Aug;37(8):1940-50.
  18. Hong KH, Kim HK, Kim YH. The role of the pars recta and pars oblique of cricothyroid muscle in speech production. *J Voice*. 2001 Dec;15(4):512-8.
  19. Stojadinovic A, Shaha AR, Orlikoff RF, Nissan A, Kornak MF, Singh B, et al. Prospective functional voice assessment in patients undergoing thyroid surgery. *Ann Surg*. 2002 Dec;236(6):823-32.
  20. Higgins MB, Netsell R, Schulte L. Vowel-related differences in laryngeal articulatory and phonatory function. *J Speech Lang Hear Res*. 1998 Aug;41(4):712-24.
  21. Emanuel FW, Sansone FE Jr. Some spectral features of "normal" and simulated "rough" vowels. *Folia Phoniatr (Basel)*. 1969;21(6):401-15.
  22. Solomon NP, Awan SN, Helou LB, Stojadinovic A. Acoustic analyses of thyroidectomy-related changes in vowel phonation. *J Voice*. 2012 Nov;26(6):711-20.
  23. Kilic MA, Ogut F, Dursun G, Okur E, Yildirim I, Midilli R. The effects of vowels on voice perturbation measures. *J Voice*. 2004 Sep;18(3):318-24.
  24. Maccallum JK, Zhang Y, Jiang JJ. Vowel selection and its effects on perturbation and nonlinear dynamic measures. *Folia Phoniatr Logop*. 2011;63(2):88-97.
  25. de Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *J Speech Hear Res*. 1995 Aug;38(4):794-811.
  26. Max L, Mueller PB. Speaking  $F_0$  and cepstral periodicity analysis of conversational speech in a 105-year-old woman: variability of aging effects. *J Voice*. 1996 Sep;10(3):245-51.
  27. Sorensen D, Horii Y. Frequency characteristics of male and female speakers in the pulse register. *J Commun Disord*. 1984 Feb;17(1):65-73.