ICIC Express Letters
Volume **14**, Number **6**, June **2020**

 ISSN 1881-803X
pp. 563–570

DETECTION SYSTEM OF PHARYNGEAL TONSIL HYPERTROPHY: OPTIMAL CEPSTRAL COEFFICIENT FOR DETECTING ANTI-FORMANT

Kohei Kabashima[1], Shunsuke Ishimitsu[1], Masashi Nakayama[1]
Shohei Komatsu[2], Kazutaka Kasai[2], Kaori Ishii[2] and Satoshi Horihata[2]

[1]Graduate School of Information Science
Hiroshima City University
3-4-1, Ozuka-Higashi, Asaminami, Hiroshima 731-3194, Japan
k-kabashima@hfce.info.hiroshima-cu.ac.jp; ishimitsu@hiroshima-cu.ac.jp

[2]Department of Orthodontics
Nihon University School of Dentistry at Matsudo
2-870-1, Sakae-Chou-Nishi, Matsudo, Chiba 271-8587, Japan

Received December 2019; accepted March 2020

Abstract. *Pharyngeal tonsil lymphoid tissue hypertrophies as acquired immunity improve during childhood. Identifying pharyngeal tonsil hypertrophy at an early stage is important for treatment, as hypertrophied tissues may lead to sleeping and breathing disorders. Therefore, we propose a non-invasive method for detecting pharyngeal tonsil hypertrophy using only speech sounds. This paper focuses on an anti-formant caused by the obstruction of the pharyngeal cavity. Anti-resonance frequency was recognized in the spectral envelope and was calculated using the cepstral coefficient. Therefore, anti-formant differences are expected to contribute to the identification of pharyngeal tonsil hypertrophy. However, the optimal cepstral coefficient has not yet been clarified. In this study, the optimal cepstral coefficient was determined using statistical analysis and experiments. The efficacy of the proposed method was demonstrated using acoustic parameters for speech sounds formed with and without pharyngeal tonsil hypertrophy.*
Keywords: Pharyngeal tonsil hypertrophy, Anti-formant, Cepstrum analysis, Cepstral coefficient

1. **Introduction.** Pharyngeal tonsil lymphoid tissue hypertrophies as acquired immunity improve throughout childhood. This tissue mass is referred to as the adenoid [1], and hypertrophy culminates between 4 to 6 years of age. Then, the pharyngeal tonsils atrophy to their adult size [2,3]. However, in some cases, hypertrophy has been observed after the atrophy term has passed. Pharyngeal tonsil hypertrophy (PTH) leads to stenosis between the nasal cavity and epipharynx. Stenosis inhibits nasal breathing and leads to oral breathing. Sleep apnea syndrome (SAS) may occur as a result of PTH [4-6]. Therefore, early detection of PTH is important to avoid disordered sleeping and breathing. Cone-beam computed tomography (CBCT) and roentgenographic cephalometry are the conventional methods used in examining pharyngeal tonsil hypertrophy. These methods provide an accurate measurement but involve radiation exposure [7-9]. To address this issue, we propose a method for detecting PTH that evaluates the anti-formants of speech sounds [10]. A non-invasive detection method of the adenoids has not been reported, so far. Therefore, our study is highly novel and useful, if applied practical application. Previous research reported a fundamental experiment in which polyvinyl chloride (PVC) pipes of different obstruction rates were used as a vocal tract model [11]. It confirmed that acoustic changes due to obstruction could be detected by evaluating the transfer function of each pipe. To apply these findings into a sound-assisted hypertrophy detection, one

DOI: 10.24507/icicel.14.06.563

must first detect a suitable anti-formant. Resonance and anti-resonance are detected from the spectral envelope. The cepstral coefficient determines the shape of the spectral envelope. Therefore, it is possible to identify PTH by determining the optimal number of cepstral coefficients.

In this report, anti-formants were detected in 28 participants with and without adenoids by conducting cepstrum analysis. The optimal number of cepstral coefficients was obtained using statistical analysis for anti-formants of each participant. The proposed method was then evaluated using these coefficients.

2. **Fundamental Principle for Observing Pharyngeal Tonsil Hypertrophy.** Speech sounds are produced in the vocal cords and transferred through sound organs. They are perceived as radiated sound from the lips and nose via the oral and nasal cavities, which function as articulators. Formants and anti-formants are characterized in these cavities and pathways, and they are referred to as resonance and anti-resonance frequency, respectively. PTH leads to stenosis between the nasal cavity and nasopharynx. As a result, changes in acoustic characteristics lead to anti-formants accompanied by the closure or non-closure of the nasal cavity [12,13]. The anti-formant appears as an anti-resonance between the first and second formants. Therefore, the spectral envelope differs between the closure and non-closure caused by PTH.

3. **Using the Cepstral Coefficient to Track Anti-Formants.** In our study, it was necessary to detect resonance and anti-resonance using cepstrum analysis. This section describes the importance of the cepstral coefficient in evaluating anti-formants and tracking resonance and anti-resonance.

3.1. **The spectral envelope in cepstrum analysis.** Cepstrum analysis is a method to separate sound source from the vocal cords and transfer characteristics of articulators. The cepstrum is defined as the inverse Fourier transform of the logarithm of a signal spectrum. Figure 1 shows the cepstrum of speech sounds. The low quefrency portion provides information on the transfer characteristics of articulators, and the high quefrency portion provides information on the vocal cord. Both types of information can be separated from

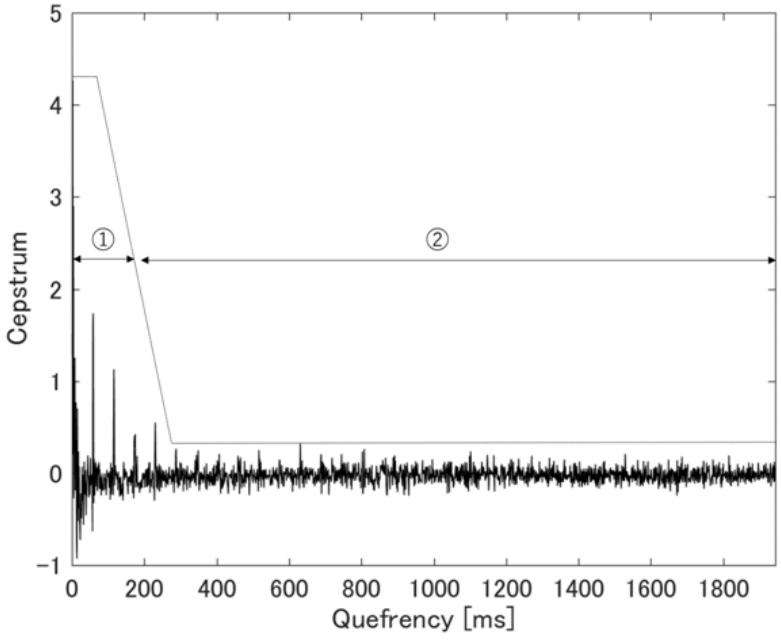


FIGURE 1. Cepstrum analysis: ① low quefrency; ② high quefrency

the sound source by filtering each quefrency portion using an arbitrary cepstral coefficient. This cepstral coefficient determines the approximate shape of the spectral envelope.

For example, Figure 2 shows the spectral envelope with the number of cepstral coefficients 32 and 64. These data confirmed that the resonance and anti-resonance differed at each cepstral coefficient. Therefore, it is necessary to determine the optimal number of cepstral coefficients first, to better determine the PTH using anti-formants.

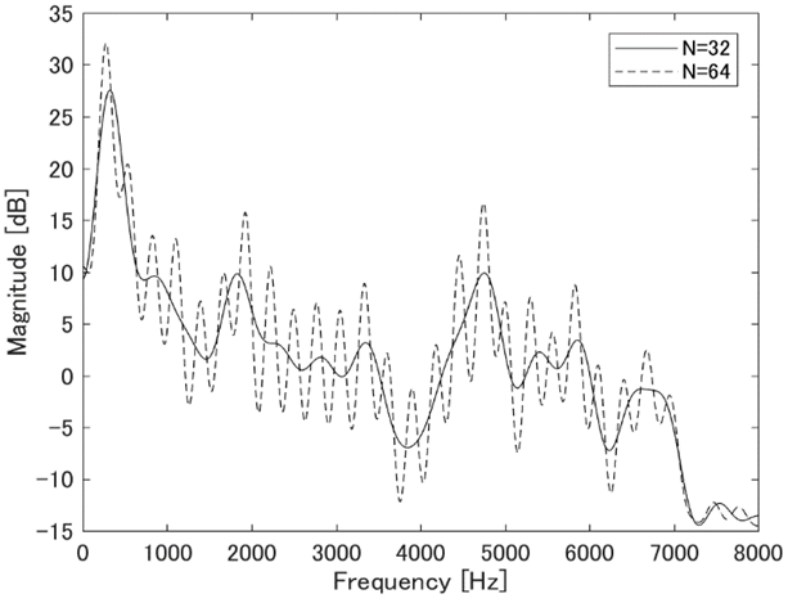
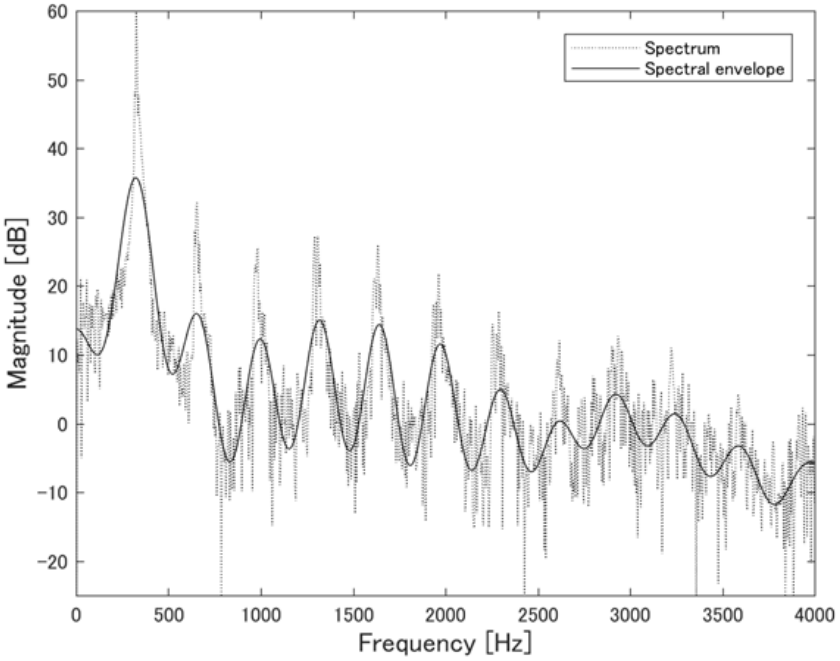


FIGURE 2. Comparison of the spectral envelope at different cepstral coefficients (N: cepstral coefficient)

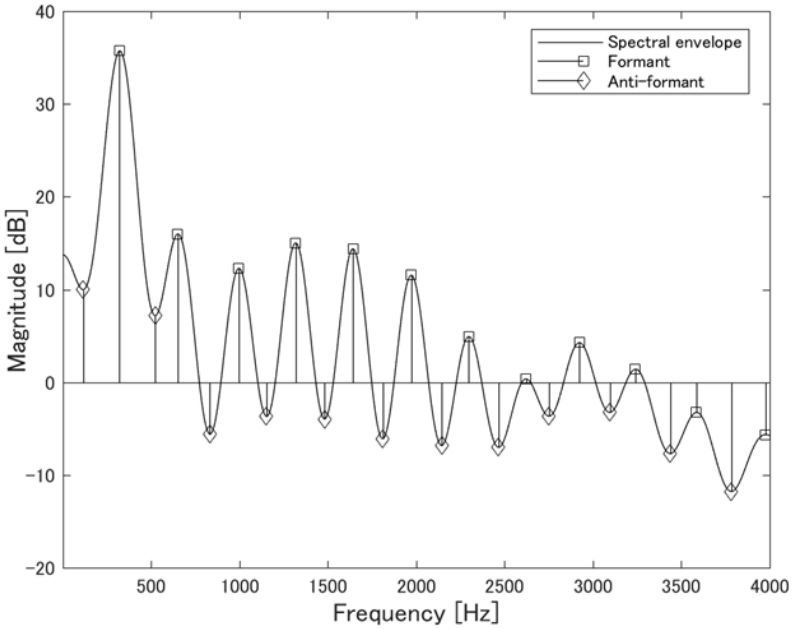
3.2. **Resonance and anti-resonance tracking using cepstrum analysis.** The resonance and anti-resonance frequencies were obtained from the spectral envelope using cepstrum analysis. Figure 3 shows the spectral envelope of speech sounds. Figure 3(a) shows the spectral envelope obtained by cepstrum analysis, and Figure 3(b) shows the resonance and anti-resonance frequencies obtained from the spectral envelope. It was confirmed that formants and anti-formants are detected from the spectral envelope.

4. **Experiment 1: Statistical Analysis to Identify Cepstral Coefficient.** This section considers the optimal number of cepstral coefficients for detecting PTH, by evaluating significant differences between the anti-formants of healthy and diseased groups.

4.1. **Dataset and method.** This subsection describes the details of the clinical dataset. It was necessary to obtain clinical data of speech sounds produced by children with PTH. Speech sounds were recorded from clinical participants in the consultation room of Nihon University Dental Hospital at Matsudo. The participants were between 6 and 11 years of age; the participant group comprised 14 children without PTH and 14 children with PTH. We considered that the number of samples is enough. They were asked to pronounce nine trials of a Japanese nasal sound /N/ repeatedly, since this sound articulates as a nasal sound from the palate. We compiled a dataset of 252 samples. The nasal resistance of all participants was measured using rhinomanometry to analyze the breathing ability of the nasal cavities, and no participant was found to have a breathing problem. The participants were asked to sit on a chair, and a unidirectional condenser microphone was placed approximately 20 cm from their mouth to record the sounds. Recording was done with 16 kHz sampling and 24 bits.



(a) Spectrum and its envelope of speech sound



(b) Resonance and anti-resonance frequencies obtained from the spectral envelope

FIGURE 3. Tracking resonance and anti-resonance frequency

We calculated the median of the anti-formant wavelengths from each participant to determine discrimination parameters. It was necessary to calibrate the difference in physique because the gender and age of participants varied. Therefore, vocal tract length was measured using CBCT images, and the median value was divided by the vocal tract length to calculate the wavelength to vocal tract length ratio (WVR) for normalization. The p-value of WVR for each cepstral coefficient was calculated using the Mann-Whitney U test by comparing the healthy and disease groups' WVR. A p-value < 0.05 was considered statistically significant.

4.2. **Result and discussion.** We analyzed this dataset using cepstrum analysis with the number of cepstral coefficients between 4 and 80. The selection of these cepstral coefficients has been described in Section 5.2. No significant difference was expected when the number of cepstral coefficients exceeded 56. Therefore, with some margin, the results obtained when using the number of cepstral coefficients between 32 and 76 are shown in Table 1 and Figure 4 as highlights. Extreme outliers were excluded in Figure 4. As shown in Table 1, a p-value < 0.01 was considered statistically significant when the numbers of cepstral coefficients were 48 and 50. A cepstral coefficient of 52 to 56 was also as close as possible to the significance level 0.01. As a result, 48 to 56 were considered the optimal number of cepstral coefficients for determining PTH as the p-values of 48 to 56 were quite small among the number of cepstral coefficients between 4 and 80.

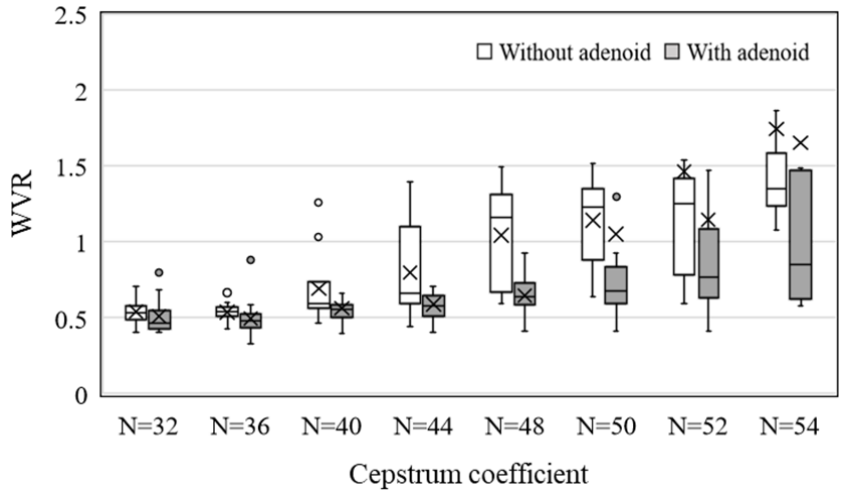
Table 1. Significant differences between anti-formants of groups with and without adenoid (Level of significance *: p-value < 0.05, **: p-value < 0.01)

Cepstral coefficient	Without adenoid ($N = 14$)			With adenoid ($N = 14$)			p-value
	Minimum	Median	Max	Minimum	Median	Max	
32	0.406	0.535	0.707	0.407	0.463	0.799	0.098
36	0.428	0.543	0.665	0.329	0.480	0.883	0.048*
40	0.466	0.591	1.262	0.399	0.555	0.920	0.035*
44	0.443	0.662	1.396	0.404	0.580	0.928	0.017*
48	0.589	1.159	1.491	0.410	0.635	0.928	0.002**
50	0.641	1.231	1.516	0.410	0.674	5.455	0.004**
52	0.590	1.253	5.766	0.413	0.766	5.455	0.054*
54	0.682	1.348	5.766	0.422	0.805	5.661	0.035*
56	1.130	1.507	6.107	0.476	0.794	5.661	0.012*
60	1.130	1.642	6.983	0.606	1.387	6.379	0.108
64	1.303	5.416	7.376	0.603	3.706	6.483	0.358
68	1.035	5.380	8.030	0.926	5.645	7.770	0.818
72	1.044	5.341	7.891	1.619	6.035	7.993	0.232
76	1.053	4.111	8.031	1.619	6.101	9.653	0.118

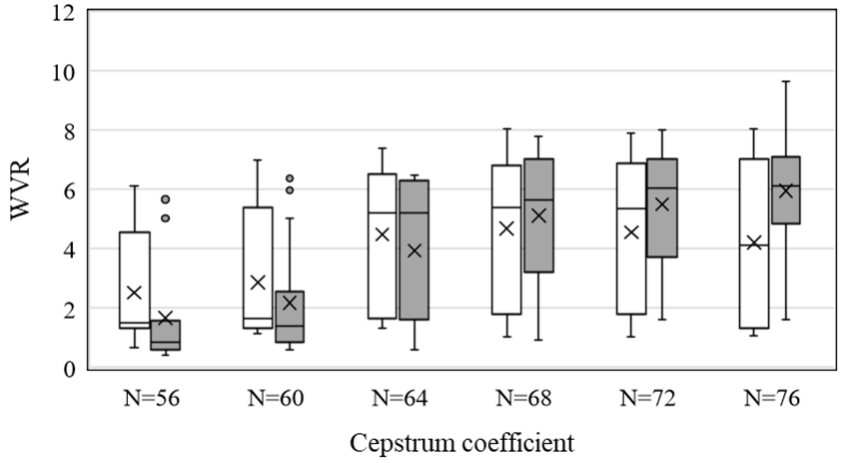
5. **Experiment 2: Discrimination Experiment for Clinical Dataset.** In Section 4, we confirmed significant differences for each cepstral coefficient by statistically analyzing the clinical data. This section verified the reliability of the optimal number of cepstral coefficients defined in Section 4 through experimentation. The efficacy of the proposed method was proven by demonstrating the discrimination performance using clinical data.

5.1. **Dataset and method.** The dataset and parameter WVR were the same as those outlined in Section 4. We arbitrarily determined a threshold value based on the WVR. The discrimination rate of PTH was calculated by comparing the threshold value and WVR. If the WVR exceeded the threshold value, it was judged to be healthy. Otherwise, it was judged to be diseased.

5.2. **Result and discussion.** The discrimination rate gradually increased as the number of cepstral coefficients increased, and the maximum discrimination rate was achieved with the number of cepstral coefficients of 48 and 56. Beyond 56, the discrimination rate was diminished. There was almost no difference in the discrimination rate, when the number of cepstral coefficients exceeded 68. Therefore, the selected number of cepstral coefficients was between 4 and 80, with some margin, because the identification rate was not expected to increase if the coefficient exceeded 68. The results obtained with the number of cepstral



(a) Number of cepstral coefficients from 32 to 54



(b) Number of cepstral coefficients from 56 to 76

FIGURE 4. Statistical distribution with and without adenoid

TABLE 2. Discrimination rate for each cepstral coefficient length

Cepstral coefficient	Thr.	Discrimination rate [%]					
		0.50	0.65	0.80	0.95	1.10	1.25
32		67.86	50.00	50.00	50.00	50.00	50.00
36		71.43	50.00	46.43	50.00	50.00	50.00
40		57.14	53.57	57.14	60.71	53.57	53.57
44		57.14	67.86	64.29	67.86	57.14	53.57
48		57.14	67.86	75.00	82.14	25.00	67.86
50		53.57	64.29	75.00	78.57	32.14	67.86
52		53.57	57.14	64.29	71.43	35.71	64.29
54		53.57	64.29	67.86	78.57	42.86	67.86
56		53.57	64.29	78.57	82.14	50.00	75.00
60		50.00	53.57	57.14	67.86	50.00	64.29
64		50.00	53.57	53.57	53.57	50.00	53.57
68		50.00	50.00	50.00	53.57	46.43	46.43
72		50.00	50.00	50.00	50.00	46.43	39.29
76		50.00	50.00	50.00	50.00	46.43	42.86

coefficients between 32 and 76 were described as highlights. Table 2 shows the discrimination rates for the clinical dataset with each cepstral coefficient. As shown in Table 2, about 82% discrimination rate was confirmed, when the number of cepstral coefficients was 48 or 56, and the threshold value was 0.95. These coefficients are consistent with the considerations outlined in Section 4. These results indicated that 48 and 56 were the optimal number of cepstral coefficients for detecting PTH. We confirmed the possibility of developing a non-invasive method for detecting PTH using only patient-produced speech sounds. However, nearly 20% of the participants were misdiagnosed. Thus, the proposed method is considered more suitable as a screening tool than a diagnostic tool.

6. **Conclusion and Future Work.** In this study, we evaluated the efficacy of a method for detecting PTH using anti-formants. This method utilizes information derived from closure and non-closure conditions in PTH. The optimal number of cepstral coefficients for identifying PTH was determined using statistical analysis. The efficacy of the proposed method was confirmed through clinical experiments. The findings of this study are as follows.

- Significant differences were calculated for each cepstral coefficient and significance was confirmed at a p-value < 0.01 when the number of cepstral coefficients was 48 and 50. And the cepstral coefficients of 52 to 56 were also as close as possible to the significance level 0.01.
- This method achieved a discrimination rate of 82% with the best combination of parameters when the number of cepstral coefficients was 48 or 56, by comparing WVR and threshold values.

These findings indicated that the optimal numbers of cepstral coefficients for identifying PTH are 48 to 56. This proposed method is expected to discriminate between healthy and hypertrophied tonsils, by evaluating patients' anti-formants. In future studies, we plan to improve the accuracy of the proposed method by adjusting the parameters.

REFERENCES

[1] T. O. Adedeji, Y. B. Amusa and A. A. Aremu, Correlation between adenoidal nasopharyngeal ratio and symptoms of enlarged adenoids in children with adenoidal hypertrophy, *African Journal of Paediatric Surgery*, vol.13, no.1, pp.14-19, 2016.

[2] S. Fujii, H. Kato, M. Fujii, F. Kato, T. Saito, H. Yamada and M. Akasaka, A study on growth changes in the pharyngeal airway, *The Japanese Journal of Pediatric Dentistry*, vol.29, no.4, pp.777-783, 1991.

[3] T. Kitamura, A. Sakabe and N. Ueda, Usefulness of cephalometry and pharyngeal findings in the primary diagnosis of obstructive sleep apnea syndrome, *Nippon Jibiinkoka Gakkai Kaiho*, vol.111, no.1, pp.695-700, 2008.

[4] J. M. Farber, Clinical practice guideline: Diagnosis and management of childhood obstructive sleeping apnea syndrome, *Pediatrics*, vol.110, pp.1255-1257, 2002.

[5] J. E. Park, S. Gray, H. Bennani, J. S. Antoun and M. Farella, Morphometric growth changes of the nasopharyngeal space in subjects with different vertical craniofacial features, *American Journal of Orthodontics and Dentofacial Orthopedics*, vol.150, no.3, pp.451-458, 2016.

[6] M. Kimura, H. Wada and T. Tanigawa, Sleep disordered breathing in children, *Japanese Society of Behavioral Medicine*, vol.23, no.2, pp.70-75, 2008.

[7] A. Manabe, T. Ishida, H. S. Yoon, S. S. Yang, E. Kanada and T. Ono, Differential changes in the adenoids and tonsils in Japanese children and teenagers: A cross-sectional study, *Scientific Reports 7*, Article No. 9734, 2017.

[8] T. Iwasaki, I. Saitoh, Y. Takemoto, E. Inada, R. Kanomi, H. Hayasaki and Y. Yamasaki, Evaluation of upper airway obstruction in Class II children with fluid-mechanical simulation, *American Journal of Orthodontics and Dentofacial Orthopedics*, vol.139, no.2, pp.135-145, 2011.

[9] S. Y. Chang, Effect of adenoid on the facial skeleton by means of roentogenocephalometric approach, *Japanese Society of Oral and Maxillofacial Surgeons*, vol.19, no.4, pp.330-347, 1973.

[10] S. Imaizumi, Multilateral study of nasal sound, *Japanese Journal of Rhinology*, vol.39, no.1, pp.51-52, 2001.

[11] K. Kabashima, S. Ishimitsu, S. Iijima, M. Nakayama, Y. Fukuda, S. Komatsu, K. Kasai, K. Ishii and S. Horihata, Experimental study of pharyngeal tonsil: A resonance model, *ICIC Express Letters, Part B: Applications*, vol.9, no.12, pp.1257-1263, 2018.

[12] K. Kabashima, M. Nakayama, S. Ishimitsu, S. Iijima, Y. Fukuda, K. Komatsu, S. Kasai and S. Horihata, Performance of pharyngeal tonsil hypertrophy detection system with phonetic sound, *Spring Meeting Acoustical Society of Japan*, pp.325-326, 2018.

[13] R. D. Kent and C. Reed, *Acoustic Analysis of Speech*, Kaibundo, 1996.