ICIC Express Letters
Part B: Applications
Volume **10**, Number **12**, December **2019**

ICIC International ©2019 ISSN 2185-2766
pp. 1093–1098

EVALUATION OF LOW-SET TONGUE POSITION USING ZERO-CROSSING RATES AND MEL-FREQUENCY CEPSTRUM COEFFICIENTS

Yuka Ochiai[1], Kimiko Yamashita[1], Kaori Ishii[1], Masashi Nakayama[2]
Shunsuke Ishimitsu[2], Satoshi Horihata[3] and Kazutaka Kasai[1]

[1]Department of Orthodontics
[3]Department of Mathematical Science
Nihon University School of Dentistry at Matsudo
2-870-1, Sakae-Chou-Nishi, Matsudo-Shi, Chiba 271-8587, Japan
mayu16010@g.nihon-u.ac.jp; kasai.kazutaka@nihon-u.ac.jp

[2]Graduate School of Information Sciences
Hiroshima City University
3-4-1, Ozuka-Higashi, Asa-Minami-Ku, Hiroshima 731-3194, Japan

Received June 2019; accepted September 2019

Abstract. *Tongue-thrust swallowing habit from a low-set tongue position causes open bite when the tongue pushes the upper and lower front teeth. This study aimed to examine tongue position discrimination by speech zero-crossing rates and Mel-frequency cepstrum coefficients (MFCCs). Participants (N = 225) underwent lateral cephalography to measure TPD (tongue-palate-distance) (TPD5). In addition, a trans palatal arch (TPA) was also attached to the oral cavity in nine adult participants (Non group), and a low-set tongue was imitated. Participants' voices were recorded before and after wearing TPA, and analysis of zero-crossing rates and MFCCs was performed. TPD was measured to be 2.5(0∼16.5) mm from the central tongue, and a significant difference was found in zero-crossing rates and MFCCs between groups. As a result, we found that acoustic changes occurred due to the low-set tongue position. It was suggested that low-set tongue could be identified by speech.*

Keywords: Low-set tongue position, Tongue-thrust swallowing, Zero-crossing rate, Mel-frequency cepstrum coefficient

1. **Introduction.** Many orthodontic patients have oral habits such as tongue-thrust swallowing and a low-set tongue position [1,2]. Tongue-thrust swallowing habits cause an malocclusion called open bite when the tongue pushes the upper and lower front teeth. In the low-set tongue position, the tongue stays in the lower position, and the anterior teeth of the lower jaw are pushed, causing anterior crossbite. In addition, this tongue position is the cause of a narrow maxillary dental arch because the tongue force that should normally be applied to the maxilla is insufficient, and thus, maxillary lateral growth is weakened [3-8]. Therefore, to prevent malocclusion, it is necessary to evaluate the position of the tongue properly and eliminate abnormal tongue function. Conventionally, X-ray cinematography, palatograms, and pressure sensors have been used to assess tongue function. However, X-ray cinematography involves radiation exposure. Since the device must be attached to the dental arch using the palatogram and pressure sensor, it may affect physiological tongue movement. Therefore, the evaluation of tongue position by visual inspection is critical. Recently, authors have reported on quantitative evaluation of the tongue position during pronunciation by analyzing the formant frequency [9,10]. This acoustic analysis can be performed in a short time and the burden on the patient is less. Additionally, the physiological function of the tongue can be evaluated noninvasively and

DOI: 10.24507/icicelb.10.12.1093

objectively. We have previously performed speech analysis using the formant frequencies of speech as a method to evaluate tongue position objectively [11-14]. Ishii et al. focused on / $\int$ / that is characteristic of the tongue-thrust swallowing habits and performed speech analysis using formant frequencies F1 and F2 as a method to objectively evaluate the tongue position. However, speech analysis using formant frequencies is limited to evaluation of the position of the tip of the tongue, and it is difficult to estimate the tongue position in detail by formant characteristics and auditory evaluation of the pronunciation speech. Therefore, we have introduced the identification of tongue proboscis by analysis of the zero-crossing rates and MFCC as a newly acoustic analysis [15]. The oral habits causing malocclusion are not only prominent in tongue-thrust swallowing habits, but also in the low-set tongue position. However, acoustic evaluation of the low-set tongue position has not been conducted. Therefore, in this study, we examined the possibility of discrimination of the low-set tongue position through analysis of the zero-crossing rates and MFCC using speech that simulates this tongue position.

2. Materials and Methods.

2.1. Participants.

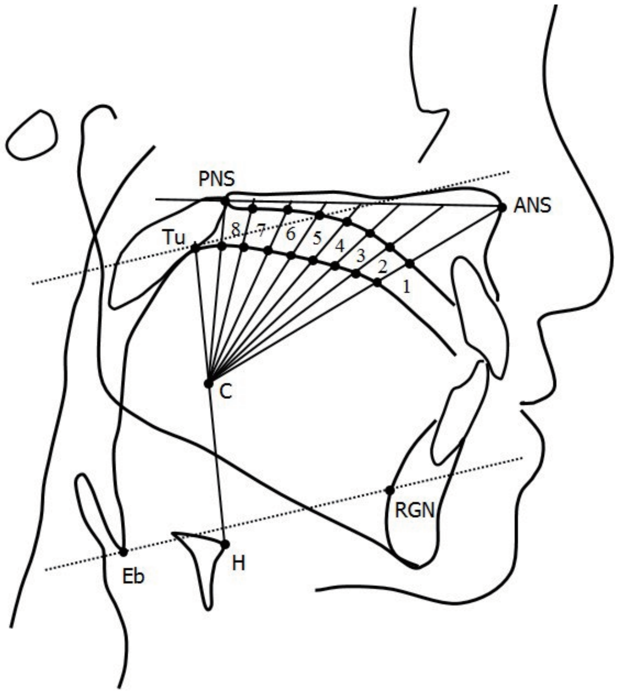
2.1.1. *Participants for evaluation of tongue position.* This study was approved by the ethics committee of Nihon University School of Dentistry at Matsudo (approval no.: EC19-18-17-16-15-007-4). Participants were 225 fourth-year students (144 males and 81 females) at the Nihon University School of Dentistry at Matsudo. No participant had malocclusion or an orthodontic appliance that could affect tongue movement.

2.1.2. *Participants for acoustic analysis of the simulated low-set tongue.* Participants for acoustic analysis of the simulated low-set tongue included 5 males and 4 females aged 28.3 ± 1.2 years. Inclusion criteria were as follows: an absence of tongue protrusion, a score on the tongue movement evaluation criteria [16] of 16 points (maximum score), and absence of irregularity in the skeleton and dentition.

2.2. **Protocol for evaluation of tongue position.** Lateral roentgenographic cephalography was performed in the 225 fourth-year students. Cephalography was performed with the head fixed according to the conventional method and the tongue resting in the intercalated position. The cephalogram was traced by the conventional method, and tracing and measurement were all performed by the same orthodontist to eliminate inter-rater bias. The measurement was performed according to the method of Ozbek [17] and Saitoh [18] (Figure 1). The upper tongue point (Tu) was defined as the uppermost point where the tongue and the straight line parallel to the straight line connecting retrognathion (RGN)-base of epiglottis (Eb) intersect. The middle point of the hyoid bone (H) and Tu was termed the “C” point. The distance between the tongue and palate on the line that equally divides the angle formed by C-anterior nasal spine (ANS) and C-posterior nasal spine (PNS) is set as tongue-palate-distance (TPD), and TPD1-TPD8 from the front. The center of the tongue (TPD5) was defined as the distance between the tongue and the palate. The median value of the tongue position in the 225 fourth-year students was calculated.

2.3. Acoustic analysis of the simulated low-set tongue position.

2.3.1. *Pronunciation condition.* In this study, we recorded uttered speeches using the trans palatal arch (TPA) to simulate the low-set tongue position (Figure 2). TPA is an auxiliary orthodontic treatment device that prevents the maxillary molar teeth from moving forward in orthodontic treatment. By placing it away from the maxillary palate mucosa, the central part of the tongue will be positioned low by the TPA, and the tongue will not rise above the TPA. /rakuda/ was pronounced nine times under the following two



1. TPD1, 2. TPD2, 3. TPD3, 4. TPD4, 5. TPD5, 6. TPD6, 7. TPD7, 8. TPD8
Eb: Deepest part of the epiglottis
RGN: Last point of symphysis
H: Anterior hyoid bone
Tu: Uppermost point where the tongue and the straight line parallel to the straight line connecting retrognathion (RGN)-base of epiglottis (Eb) intersect
C: Measured middle point of the hyoid bone (H) and Tu

Figure 1. Analysis of tongue position

Figure 2. Trans palatal arch (TPA)

conditions: normal pronunciation (Non group) and TPA positioned 10 mm away from the palate mucosa (TPA group).

2.3.2. *Acoustic analysis.* Stimuli were recorded with participants in the sitting position. A unidirectional condenser microphone (ATM 31a, Audio Technica, Inc., Tokyo) was placed about 20 cm from the lips in a position that would not be affected by respiratory exhalation. For the stimuli, unvoiced soft palate plosive /k/ and posterior tongue vowel /u/ in /ku/ were used. Participants pronounced /rakuda/ 9 times. The speech was quantized at a sample frequency of 22.05 kHz with an accuracy of 16 bits through an audio-technica interface (EDIROL UA-25EX, Roland Co., Shizuoka, Japan). The stimuli were extracted using voice waveform analysis and editing software (Acoustic Core, Arcadia, Inc., Osaka) stored on a personal computer (Vostro, Dell, Inc., Kanagawa, Japan). Analysis of zero-crossing rates and MFCC was performed according to Nakayama et al. [19]. Calculation of zero-crossing rates and MFCC were performed using speech analysis software, and a median value of each acoustic feature duration obtained from each speech sample in frame units.

The comparison between the Non and TPA groups was performed by the Wilcoxon test. All statistical analyses were performed using SPSS (SPSS Inc., Chicago, IL, USA).

3. **Results.** TPD5 in our study population was 2.5(0∼16.5) mm and 90%tile was 9.5 mm (Table 1). There were significant differences in zero-crossing rates and MFCC2 between Non and TPA groups ($p < 0.01$; Table 2). In addition, significant differences were also found in MFCC4 and MFCC8 between Non and TPA groups ($p < 0.05$; Table 2).

Table 1. Tongue-palate-distance ($N = 225$)

	Median	Range	90%tile
PTD1	0.0	0 ∼ 18.0	8.5
PTD2	1.5	0 ∼ 20.0	10.0
PTD3	1.8	0 ∼ 21.5	10.5
PTD4	2.0	0 ∼ 19.0	10.0
PTD5	2.5	0 ∼ 16.5	9.5
PTD6	4.0	0 ∼ 17.0	10.0
PTD7	5.5	0 ∼ 18.0	11.0
PTD8	8.5	0 ∼ 20.0	14.0

Table 2. Zero-crossing rates and MFCC1 to 13 values in Non and TPA groups

	Non group ($N = 9$)		TPA group ($N = 9$)		
	Median	Range	Median	Range	p-value
Zero-crossing	4.00	1.88 ∼ 5.43	3.48	2.00 ∼ 5.06	0.004**
MFCC1	−13.90	−5.80 ∼ 12.62	−14.22	−16.46 ∼ −12.22	0.212
MFCC2	0.91	0.33 ∼ 1.43	1.26	0.67 ∼ 2.62	0.001**
MFCC3	−0.69	−1.19 ∼ 0.44	−0.49	−1.23 ∼ 0.13	0.136
MFCC4	0.40	−0.04 ∼ 1.02	0.11	−0.41 ∼ 1.11	0.021*
MFCC5	0.33	−0.32 ∼ 0.81	0.20	−0.71 ∼ 0.73	0.130
MFCC6	0.16	−0.40 ∼ 0.96	0.25	−0.09 ∼ 0.82	0.302
MFCC7	−0.53	−0.99 ∼ 0.10	−0.51	−1.02 ∼ 0.16	0.118
MFCC8	−0.11	−0.54 ∼ 0.28	−0.17	−0.50 ∼ 0.16	0.021*
MFCC9	−0.37	−0.82 ∼ 0.01	−0.30	−0.89 ∼ 0.07	0.203
MFCC10	0.12	−0.31 ∼ 0.53	0.20	−0.36 ∼ 0.52	0.755
MFCC11	−0.53	−1.04 ∼ 0.24	−0.49	−1.07 ∼ 0.34	0.203
MFCC12	−0.15	−0.80 ∼ 0.60	−0.21	−0.67 ∼ 0.15	0.501
MFCC13	−0.05	−0.77 ∼ 0.39	−0.15	−0.62 ∼ 0.31	0.923

Notes: Wilcoxon signed rank test of signed-rank test; NS: not significant, *: $p < 0.05$, **: $p < 0.01$

4. **Discussion.**

4.1. Evaluation of tongue position. Since the TPD5 in the 225 fourth-year students was 2.5(0∼16.5) mm, it is considered that a certain distance exists between the tongue and palate at the center of the tongue. In addition, since 90%tile was 9.5 mm, we decided to diagnose a person with TPD5 of 10 mm or more as a low-set tongue in this study.

4.2. Speech analysis of the simulated low-set tongue. The zero-crossing rate is the number of zero crossing instances per unit time in the speech signal [20]. It is assumed that it detects the difference between normal and simulated low-set tongue pronunciation as frequency features, especially for consonants. Therefore, the zero-crossing rates may have utility in the evaluation of tongue function. In speech recognition, MFCC is widely used as a speech recognition parameter and is a kind of a cepstrum coefficient. The MFCC represents a value obtained by approximating and human perceptual response with a mel-scale which is a logarithmic linear frequency band [21]. Acoustic-feature vectors, MFCC, were set to 13 dimensions for acoustic evaluation. Cepstrum coefficient indicates roughness of spectral envelope. Low order means a gradual fluctuation of spectrum, while higher order means a fine fluctuation of spectrum. Since TPD5 was 2.5(0∼16.5) mm and 90%tile was 9.5 mm, the tongue position 10 mm from the palate mucosa was defined as the low-set tongue position.

Significant differences between the Non and TPA groups in the zero-crossing rates and MFCC2, MFCC4, and MFCC8 were observed. Therefore, the low-set tongue position could be identified with the envelope from MFCC2, 4, 8. It was suggested that acoustic changes in speech during pronunciation may correlate with tongue position. Yamashita et al. [15] found that MFCC8 is related to tongue protrusion. It is conceivable that the tongue was pushed forward by the TPA. Therefore, zero-crossing rates, MFCC2, and MFCC4 are believed to be related to tongue height.

5. Conclusions. It was suggested that the low-set tongue position can be identified by using the zero-crossing rates, MFCC2 and MFCC4, which quantify the acoustic features of phonetic speech. Therefore, a more objective tongue evaluation can be performed by evaluating speech using these methods.

REFERENCES

[1] T. Ohno, Y. Yoshida, S. Takami and M. Irie, Clinical practice of myofunctional therapy, *Tokyo: Nihon Shika Shuppan*, pp.18-39, 1986 (in Japanese).

[2] K. Enoki, K. Motohashi and Y. Nakamura, Form, size and functional anomalies of the tongue, *Orthod Waves Jpn. Ed.*, vol.14, pp.13-20, 1955.

[3] M. L. Hanson and M. V. Andrianopoulos, Tongue thrust and malocclusion: A longitudinal study, *Int. J. Orthodontics*, vol.20, no.2, pp.9-18, 1982.

[4] W. L. Kydd, J. S. Akamine, R. A. Mendel and B. S. Kraus, Tongue and lip forces exerted during deglutition in subjects with and without an anterior open bite, *J. Dent. Res.*, vol.42, pp.858-866, 1963.

[5] M. G. Sahad, A. C. Nahás, H. Scavone-Junior, L. B. Jabur and E. Guedes-Pinto, Vertical interincisal trespass assessment in children with speech disorders, *Braz. Oral Res.*, vol.22, no.3, pp.247-251, 2008.

[6] B. Melsen, K. Stengaard and J. Pedersen, Sucking habits and their influence on swallowing pattern and prevalence of malocclusion, *Eur. J. Orthod.*, vol.1, no.4, pp.271-280, 1979.

[7] A. S. Cayley, A. P. Tindall, W. J. Sampson and A. R. Butcher, Electropalatographic and cephalometric assessment of tongue function in open bite and non-open bite subjects, *Eur. J. Orthod.*, vol.22, no.5, pp.463-474, 2000.

[8] M. Ishikawa, K. Okamura, S. Shoji, N. Isogawa and Y. Takagi, Influence of the tongue rest position on the anterior occlusal relationship, *Jpn. J. Ped. Dent.*, vol.44, pp.534-540, 2006.

[9] Y. Hama, K. Inoue, T. Tamura, M. Hayashi and S. Matsunaga, Studies on an application of acoustic analysis to the diagnosis of open-bite tendency and its usefulness, *Nihon Univ. Dent. J.*, vol.64, pp.412-420, 1990.

[10] M. Kudo, S. Takahashi, N. Isogawa, T. Sato, N. Yano, M. Ishikawa and Y. Takagi, Articulation characteristics of children with tongue thrust swallowing: First report on an acoustic analysis of alveolar /s/, *Jpn. J. Ped. Dent.*, vol.43, pp.79-84, 2005.
[11] K. Ishii, K. Saitoh and K. Kasai, Maxillofacial and pharyngeal airway morphology and acoustic characteristics in adult tongue thrust swallowers, *Orthod. Waves Jpn. Ed.*, vol.71, pp.33-41, 2012.
[12] K. Ishii, K. Saitoh and K. Kasai, Clinical application of acoustic analysis in evaluation of tongue function, *Orthod. Waves Jpn. Ed.*, vol.71, no.3, pp.170-177, 2012.
[13] I. Nishio, K. Ishii, K. Saitoh et al., The long-term study about changes in hard and soft tissues and the position of the tongue at speech after postorthognathic surgery, *Orthod. Waves Jpn. Ed.*, vol.74, no.1, 2015.
[14] I. Nishio, K. Ishii, K. Saitoh et al., Adaptation of the speech analysis to mandibular prognathism, *JSMFT*, vol.4, no.1, pp.9-15, 2015.
[15] K. Yamashita, M. Nakayama, K. Ishii et al., Evaluation of tongue function after MFT using zero-crossing and MFCC, *J. Oral Med. Sci.*, vol.17, no.1, pp.1-8, 2018.
[16] Y. Ishino, Y. Yamashita, K. Nemoto et al., Severity of ankyloglossia and functional disorders: Quantitative evaluations of voluntary tongue movements, articulatory and swallowing functions, *J. the Japanese Stomatological Society*, vol.50, pp.26-34, 2001.
[17] M. M. Ozbek, U. T. Memikoglu, A. T. Altug-Atac and A. A. Lowe, Stability of maxillary expansion and tongue posture, *Angle Orthod.*, vol.79, no.2, pp.214-220, 2009.
[18] K. Saitoh, Long-term changes in pharyngeal airway morphology after mandibular setback surgery, *Am. J. Orthod. & Dentofacial Orthop.*, vol.125, no.5, pp.556-561, 2004.
[19] M. Nakayama, S. Ishimitsu, K. Yamashita, K. Ishii, K. Kasai and S. Horihata, Tongue habit discrimination system using acoustical feature for oral habits improvement, *Electronics and Communications in Japan*, vol.138, no.3, pp.242-248, 2018.
[20] F. Gouyon, F. Pachet and O. Delerue, On the use of zero-crossing rate for an application of classification of percussive sounds, *Proc. of the COST G-6 Conference on Digital Audio Effects*, DAFX-1-6, 2000.
[21] S. Furui, Comparison of speaker recognition methods using statistical features and dynamic features, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.29, pp.342-350, 1981.