

모음연장 음성 샘플의 분석 구간에 따른 음향학적 파라미터 비교

신유정
호원대학교 언어치료학과

Comparison of Acoustic Parameters According to the Section of Analysis in Sustained Vowel Phonation

Yu-Jeong Shin

Department of Language Therapy, Howon University

요약 본 논문은 임상에서 음성장애 환자의 객관적 음성 분석 대상으로 주로 쓰이는 모음연장 발성이 분석하는 구간에 따라 어떠한 음향학적 차이를 보이는지 밝히고자 하였다. 본 연구에서는 성대결절 환자 17명의 /a/ 모음연장 발성을 시작, 중간, 끝 구간으로 편집하여 MDVP를 통해 각 구간의 jitter, shimmer, NHR을 산출하였고, 비교를 위하여 정상 음성 집단 12명의 음성도 분석하였다. 산출 결과는 R 통계프로그램을 활용하여 Fridman test와 사후 검정을 실시하였다. 음성장애 환자 집단은 모음연장 발성의 끝 구간이 중간 구간에 비해 jitter, shimmer, NHR 값이 모두 유의하게 높은 것으로 나타났다. 또한, 발성의 시작 구간은 중간 구간에 비해 세 파라미터 모두에서 높게 산출됐지만 유의한 차이는 없었다. 반면, 정상 집단은 발성의 시작, 중간, 끝 모든 구간에서 유의한 차이가 없었다. 모음연장 발성은 구간에 따라 음향학적 파라미터의 분석 결과가 다르고 발성 끝 구간에서 중간 구간보다 유의하게 음성이 불안정해지는 것으로 나타났다. 이러한 결론은 임상 현장에서 모음연장 발성의 분석 구간 선택과 결과 해석에 유용하게 활용될 수 있을 것이다.

Abstract This study aimed to investigate the acoustic differences that occur in diverse sections of sustained vowel phonation, which is often used in an objective speech analysis of voice disorder patients. The subjects included 17 voice disorder patients (vocal nodules) and 12 normal individuals without any voice disorder. The participants' sustained vowel phonation of /a/ was divided into onset, middle, and offset, and the jitter, shimmer, and NHR in each section were analyzed using the MDVP(Multi-Dimensional Voice Program). The Friedman test and post hoc analysis were used. In the vocal nodules group, the jitter, shimmer and NHR were significantly higher in the off section of sustained vowel phonation than in the middle section, and there were no significant differences between the beginning and middle sections. In contrast, in the group of normal individuals, there were no significant differences between any of these sections. The values of the acoustic parameters according to the section of analysis in the sustained vowel phonation are different and the vocal in the end section is significantly more unstable than that in the middle section. The results of this study will be useful for selecting the sections to be analyzed in sustained vowel phonation and interpreting the results of the analysis.

Keywords : MDVP parameters, sampling section, sustained vowel phonation, voice analysis, voice disorder

1. 서론

음성장애 환자의 음성은 평가자의 청지각에 의한 주

관적 평가와 음향학적 분석기기를 활용하는 객관적 평가
결과 모두를 포함하여 다차원적으로 평가해야한다[1].
이 중, 기기를 사용한 분석은 적절한 음성 진단과 치료를

본 논문은 호원대학교 연구과제로 수행되었음.

*Corresponding Author : Yu-Jeong Shin(Howon Univ.)

Tel: +82-63-450-7487 email: uzhang@naver.com

Received June 14, 2017

Revised (1st June 29, 2017, 2nd July 6, 2017)

Accepted July 7, 2017

Published July 31, 2017

위한 음성의 양적(quantitative) 수치를 제공한다[2].

국내 음성언어재활사가 근무하고 있는 이비인후과 임상에서 가장 보편적으로 사용하는 객관적 분석 프로그램은 Coputerized Speech Lab(CSL)의 Multi-Dimentional Voice Program(MDVP)이며, 음향학적 분석을 위한 음성 자료로는 모음 /아/의 연장 발성이 많이 사용된다[3].

모음 연장 발성은 주기적이거나 준주기성을 갖는 신호에서만 정확한 분석이 가능한 대부분의 음향학적 분석 [4]에서 억양, 강세와 같은 변동 요인이 없어 비교적 안정적으로 녹음이 가능[5][6]하기 때문에 분석 자료로 주로 쓰이고 있다.

MDVP 분석을 위한 음성 녹음은 모음 연장 발성을 분석 대상으로 할 때, 보통 모음 /아/를 평상 시 목소리와 크기로 약 3초 혹은 4초간 지속하도록 하여 녹음한다 [7][8]. 녹음된 음성은 분석하는 임상가나 연구자에 따라 모음 전체 구간을 다 분석하기도 하고 안정된 구간을 편집하여 분석하기도 한다. jitter(주파수변동률), shimmer(진폭변동률)와 같은 발성의 변동률 값을 음성 자료의 안정성에 크게 영향을 받기 때문에 안정된 구간을 찾기 위한 연구들[9][10]도 활발히 이루어지고 있다. 하지만 ‘안정적인’ 구간에 대한 객관적이고 정확한 기준이 없기 때문에 현재 임상과 연구 현장에서는 모음 전체 녹음 길이의 중간 부분 중 육안으로 비교적 흔들림이 없어 보이는 구간을 설정하여 분석하고 있는 경우가 대부분이다.

그러나 장애 음성의 특성은 음성장애의 진단명과 중증도에 따라 다양한 양상으로 나타날 수 있으므로 단순히 모음의 안정 구간만을 분석한다면 환자 음성의 특징적 병리 현상을 간과할 수 있는 가능성이 있다. 모음 연장 발성의 분석 구간에 따른 청지각적 평가와 음향학적 평가를 분석한 한 선행 연구[11][12]에서는 조조성(roughness) 항목이 음성 신호의 비주기성 및 불안정성과 관련되어 있고, 모음 시작 구간에서의 평가 결과가 가장 안정적인 구간보다 전체 음질에 대한 중요한 정보를 줄 수 있다고 밝혔다.

음향학적 분석 대상으로서의 모음 연장 발성에 대한 연구는 연속된 발화(continuous speech)와의 비교, 모음의 분석 길이와 안정된 구간의 추출 방법에 따른 결과 차이, 모음 별 음성 평가의 결과 차이 등의 주제로 많은 연구 결과가 발표되었으나 모음 /아/의 연장 발성 자체를 대상으로 각 구간 별 음향학적 특성에 대한 연구는 없었으므로 본 연구에서는 모음 /아/의 연장 발성 자료를 시

작과 중간, 끝 구간으로 나눈 후, 구간에 따라 음향학적 파라미터가 어떠한 양상으로 나타나는지 확인하고자 하였다. 이를 통해 현재 사용하고 있는 음향학적 분석 자료의 타당성을 점검, 고찰하고자 하였다.

따라서 본 연구는 음성언어재활사가 가장 많이 접하는 성대결절 환자[2]음성을 대상으로 모음 /아/ 연장 발성의 시작, 중간, 끝 구간에 대한 jitter, shimmer, NHR(Noise to Harmonic Ratio; 소음 대 배음 비율)을 산출하고, 이와 함께 정상 음성 집단에 대한 분석도 실시하였다.

2. 본론

2.1 연구방법

2.1.1 연구대상

본 연구는 전북대학병원 이비인후과에 내원하여 성대결절로 진단받은 환자 17명(남 2명, 여 15명, 평균연령만 44.2세)의 음성을 분석 대상으로 하였으며, 대상 환자는 성대결절 외에 음성에 영향을 미칠 수 있는 다른 질환이 없는 환자로 제한하였다.

또한, 정상 음성의 모음 연장 발성에서도 분석 구간 별 파라미터의 차이가 있는지 확인하기 위해 언어치료과에 재학 중인 대학생 12명(남 5명, 여 7명, 평균연령만 20.8세)의 음성을 분석하였다. 음성에 영향을 줄 수 있는 폐질환이나 호흡기 질환이 있는 학생은 제외하였고, 녹음 당시 음성에 문제가 없는 학생들을 대상으로 하였다.

Table 1. Information of subjects

	Number	Mean age
Patients	17 (M 2, F 15)	44.2
Control group	12 (M 5, F 7)	20.8

2.1.2 자료수집 및 분석

음성 자료는 방음 시설이 된 음성검사실에서 AKG사의 C420 헤드셋 콘텐서 마이크를 사용하여 녹음되었으며 음성표본추출률은 44100Hz였다. 대상자는 마이크에서 약 10cm 떨어진 곳에서 /아/ 모음을 편안한 음도와 강도로 4초간 연장 발성하였다. 녹음된 자료는 CSL의 MDVP를 통해 음성 파라미터를 산출하였다. 모음 연장 발성의 구간별 분석을 위해 선행연구들[13][14]의 분석

시간(1초)를 참고하여 모음의 발생 시작 1초, 중간 지점의 안정 구간 1초, 발생 끝 부분 1초를 편집하였고 각 구간에 대한 jitter, shimmer, NHR을 분석하였다.

2.1.3 통계분석

MDVP를 통해 얻어진 결과는 통계 프로그램 R을 활용하여 분석하였다. 각 집단 별 모음연장 발생의 구간에 따른 음성 파라미터 차이의 유무를 보기 위해 Kruskal-Wallis test를 사용하였고 95% 신뢰수준($p < .05$)을 기준으로 하였다. 유의한 차이를 보인 결과에 대해서는 Tukey HSD test로 사후검정을 실시하였다.

2.2 연구 결과

2.2.1 성대결절 환자의 모음 연장 발생 구간 별 음성 파라미터

성대결절 환자의 모음 연장 발생을 시작, 중간, 끝 구간으로 나누어 분석한 결과, 표 1과 같이 jitter, shimmer, NHR 값 모두 구간에 따라 유의한 차이($p < .05$)가 있는 것으로 나타났다. 이에 사후 검정을 실시한 결과는 그림 1, 2, 3에서 보는 바와 같이 jitter, shimmer, NHR 값 모두 발생 끝 구간이 중간 구간과 비교하여 통계적으로 유의하게 높은 수치를 보였다. 또한, shimmer 값에서는 첫 발생 시작 구간과 끝 구간에서도 유의한 차이가 나타났다.

2.2.2 정상 음성의 모음 연장 발생 구간 별 음성 파라미터

정상 음성 집단의 모음연장 발생 구간 별 파라미터를 분석한 결과는 <표 2>와 같다. jitter, shimmer, NHR 모두 모음연장 발생의 시작, 중간, 끝 구간 별 유의한 차이를 보이지 않았다.

Table 2. Comparison of acoustic parameters according to the section of sustained vowel in subjects with vocal nodules

	onset	middle	offset	<i>p</i>
jitter (%)	2.2 ± 1.9	1.8 ± 1.3	3.9 ± 5.0	0.001**
shimmer (%)	5.1 ± 3.3	3.4 ± 1.6	9.4 ± 13.5	0.000**
NHR	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.2	0.013*

* $p < .05$, ** $p < .01$

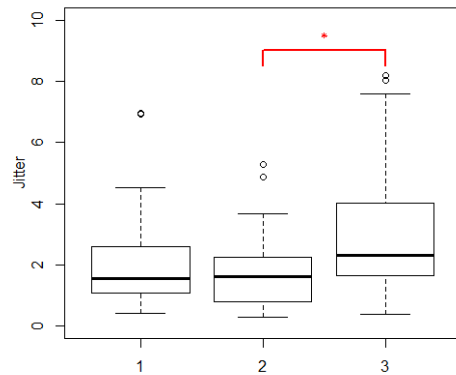


Fig. 1. Box plot of jitter(%) according to the section (1-onset, 2-middle, 3-offset) of sustained vowel in subjects with vocal nodules

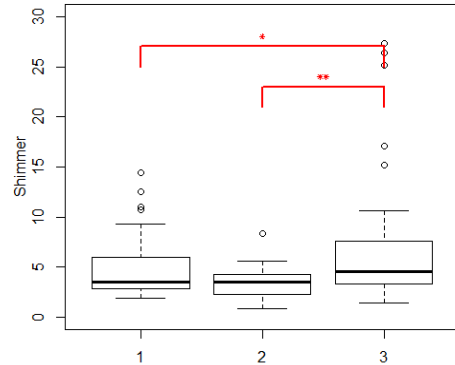


Fig. 2. Box plot of shimmer(%) according to the section (1-onset, 2-middle, 3-offset) of sustained vowel in subjects with vocal nodules

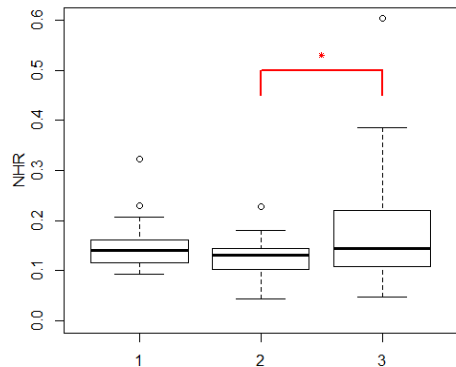


Fig. 3. Box plot of NHR according to the section (1-onset, 2-middle, 3-offset) of sustained vowel in subjects with vocal nodules

Table 3. Comparison of acoustic parameters according to the section of sustained vowel in normal subjects

	onset	middle	offset	p
jitter (%)	1.0 ± 0.4	1.0 ± 0.5	1.2 ± 0.5	0.479
shimmer (%)	4.4 ± 1.7	3.8 ± 1.2	4.6 ± 1.8	0.506
NHR	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.706

3. 고찰 및 결론

3.1 고찰

본 연구는 음성재활 임상 현장에서 가장 흔히 쓰이고 있는 음향학적 분석 대상인 ‘모음 /아/의 연장 발성’이 분석 구간에 따라 음성 파라미터에 어떠한 차이가 나타나는지 살펴보았다. 그 결과, 성대결절 환자 집단은 중간 구간과 끝 구간의 jitter와 shimmer 값에서 유의한 차이 ($p < 0.05$)가 나타났다. jitter와 shimmer 모두 끝 구간 값이 중간 구간보다 더 높았는데 이는 끝 구간에서 음성이 유의한 수준으로 더 불안정해짐을 뜻한다. 모음 발성의 끝 구간은 발성의 끝마침으로 인해 성대 진동이 느려지며 성대 프라이(glottal fry)가 산출되기 쉽고 이는 진동 운동의 주기성과 기본주파수에 영향을 미치기 때문에 [15][16] 변동률 수치인 jitter와 shimmer 값도 높아졌을 것으로 판단된다. 발성 시작과 끝의 PTF(Phonation Threshold Flow; 발성역치기류)를 분석한 연구[17] 결과 역시 발성 끝 부분에서의 PTF 값이 시작에 비해서 항상 낮게 나타났고, 이를 성대 진동의 이력현상(hysteresis phenomenon)으로 설명하였다. 이러한 선행 연구 결과들은 발성 시작부터 끝까지의 음질이 일정하지 않다는 본 연구의 결과를 뒷받침할 수 있으며 끝 구간을 제외하고 안정 구간을 분석하는 방법이 타당하다고 해석할 수 있을 것이다.

반면, 모음 연장 발성의 시작 구간과 중간 구간, 시작 구간과 끝 구간은 각각 음성 파라미터 값에 유의한 차이가 없었다. 음성장애 환자의 발성 시작은 정상 음성에 비해 부적절한 성대 접촉을 초래할 가능성이 높기 때문에 [18] 모음 발성의 시작 구간에서 더 불안정한 음향학적 특성을 가질 것으로 예측해볼 수 있다. 그러나 본 연구의 결과에서는 시작 구간에서 jitter와 shimmer 값이 중간

구간에 비해 약간 높게 산출되긴 하였으나 통계적으로 유의하지는 않았다. 발성의 양상은 후두 풀립이나 반회 신경마비 등에서는 끝이 긴 경우가 많고 후두암 등에서는 시작이 긴 경우가 많으며 이 구간에서 큰 흔들림을 수반[19]하는 등 음성장애에 따라 음성의 병리적 특성이 다르게 나타나는데 본 연구에서는 성대결절 환자 음성만을 대상으로 하였기 때문에 제한된 결과가 산출됐을 가능성이 있다. 갑상선 절제술 환자의 수술 전후 음성을 분석한 한 연구[20]에서는 모음 연장 발성을 구간 별로 분석한 결과, 모음의 시작 구간이 수술 전후 유의한 차이가 나타났다고 밝힌 바 있다. 이러한 선행 연구 결과는 음성장애의 진단명에 따라 발성 시작의 음향학적 특성이 각각 다르게 나타날 수 있기 때문에 짧은 음성 샘플을 대상으로 전반적인 환자 음성을 판단하는 임상에서 샘플 구간을 결정할 때 보다 신중해야 한다는 것을 뜻한다. 추후 다양한 음성장애 집단을 대상으로 음성장애 진단명에 따른 구간 별 음향학적 차이를 분석하는 후속 연구가 필요할 것으로 보인다.

또한, 정상 음성 집단의 모음 연장 발성은 구간에 따른 유의한 차이가 나타나지 않았다. 즉, 정상 음성은 발성의 시작과 중간, 끝 구간이 비교적 일정하게 대상자의 음성을 반영한다고 볼 수 있다. 이는 음성장애 환자의 음성은 정상 음성에 비해 짧은 발성 시간 동안에도 불안정한 특성을 보인다는 것을 밝혀주는 결과이다. 정상 집단을 대상으로 음성의 시작 구간과 끝 구간의 음성 파라미터를 분석한 연구[10]에서도 두 구간이 중간 구간에 비해 높은 수치를 보였으나 통계적으로 유의한 결과가 없었다.

3.2 결론

MDVP를 활용한 음성 분석의 결과는 녹음 시 주변 환경과 마이크 위치 등에 따라 민감하게 달라질 수 있고 [6], 음성의 주기성이 감소하고 음성장애가 심해질수록 기본주파수 변화를 추적하기 어려워 그 신뢰도가 떨어지는[21] 치명적 단점을 지닌다. 그럼에도 불구하고 MDVP는 환자의 음향학적 파라미터를 쉽고 간편하게 산출할 수 있기 때문에 임상 현장에서 많이 쓰이고 있다. 따라서 음성의 분석 대상 자료와 분석 방법이 결과에 미치는 영향에 대한 정확한 이해가 필요하다. 본 연구 결과에 따르면, 장애 음성의 모음 /아/의 연장 발성은 구간에 따라 유의한 차이가 있으므로 분석 구간의 적절한 선택

이 필요하며 산출된 결과에 대한 타당한 해석이 요구된다고 할 수 있다. 또한, 모음 /아/는 한국어 모음 중에서도 높은 변동률을 보이는 불안정한 모음[22]이라는 최근 연구 결과는 음향학적 분석 자료에 대한 근본적인 고찰이 필요하다는 것을 보여준다.

본 연구는 대상자 수가 적고, 기능적 음성장애로 분류되는 성대결절 환자의 음성만을 대상으로 하였기 때문에 일반화가 어려울 수 있다는 제한점을 가진다. 그러나 실제 임상에서 쓰이는 음성 샘플의 우려점을 객관적으로 확인하였고, 샘플 선정의 중요성을 다시금 확인했다는 데에 연구의 의의가 있겠다. 추후 연구에서는 다른 음성 장애 집단의 음성 및 중증도별 음성 분석을 세분화하여 모음 연장 발생 자료에 대한 보다 구체적인 분석이 이루어져야 할 것이다.

References

- [1] Piccirillo JF, Painter C, Haiduk A, Fuller D, Fredrickson J., *Assessment of two objective voice function indices*. Annals of Otolaryngology, Rhinology & Laryngology, vol. 107, pp. 396 - 400, 1998.
DOI: <https://doi.org/10.1177/000348949810700506>
- [2] Bhuta, T., Patrick, L., Garnett, J. D., *Perceptual evaluation of voice quality and its correlation with acoustic measurements*. Journal of voice, vol. 18, no. 3, pp. 299-304, 2004.
DOI: <https://doi.org/10.1016/j.jvoice.2003.12.004>
- [3] Choi Seong Hee, *Speech-Language Pathologists' Voice Assessment and Voice Therapy Practices: A Survey for Standard Clinical Guideline and Evidence-Based Practice*. Communication Sciences & Disorders, vol. 18, no. 4, pp. 473-85, 2013.
DOI: <https://doi.org/10.12963/csd.13082>
- [4] Titze, I. R., *Workshop on Acoustic Voice Analysis: Summary Statement*. National Center for Voice and Speech: Iowa City, IA., 1995.
- [5] Parsa, V., Jamieson, D. G., *Acoustic discrimination of pathological voice: sustained vowels versus continuous speech*. Journal of Speech, Language & Hearing Research, vol. 44, no. 2, pp. 327-339, 2001.
DOI: [https://doi.org/10.1044/1092-4388\(2001\)027](https://doi.org/10.1044/1092-4388(2001)027)
- [6] Zhang, Y., Jiang, J. J., *Acoustic Analyses of Sustained and Running Voices From Patients With Laryngeal Pathologies*, Journal of Voice, vol. 22 no. 1, pp. 1-9, 2008.
DOI: <https://doi.org/10.1016/j.jvoice.2006.08.003>
- [7] KayPentex, *Multi-Dementional Voice Program(MDVP) Software Introduction Manual*, Lincoln Park. NJ. USA. 2008.
- [8] Go Do Heung et al., *Experimental Phonetics for Speech-Language Pathologists*. Seoul: Hakjisa. 2015.
- [9] Olszewski, A. E., Shen, L., Jiang, J. J., *Objective Methods of Sample Selection in Acoustic Analysis of Voice*. Annals of Otolaryngology, Rhinology & Laryngology, vol. 120, no. 3, 155-161, 2011.
DOI: <https://doi.org/10.1177/000348941112000303>
- [10] Choi S. H., Lee J., Sprecher A. J., Jiang J. J., *The effect of segment selection on acoustic analysis*, Journal of Voice, vol. 26, no. 1, pp. 1-7, 2012.
DOI: <https://doi.org/10.1016/j.jvoice.2010.10.009>
- [11] de Krom, *Consistency and Reliability of Voice Quality Ratings for Different Types of Speech Fragments*. Journal of Speech, Language, and Hearing Research, vol. 37, no. 5, pp. 985, 1994.
DOI: <https://doi.org/10.1044/jshr.3705.985>
- [12] de Krom, *Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments*. Journal of Speech, Language, and Hearing Research, vol. 38, no. 4, pp. 794, 1995.
DOI: <https://doi.org/10.1044/jshr.3804.794>
- [13] Martin, D., Fitch, J., & Wolfe, V., *Pathologic voice type and the acoustic prediction of severity*. Journal of speech and hearing research, vol. 38, no. 4, pp. 765-771, 1995.
DOI: <https://doi.org/10.1044/jshr.3804.765>
- [14] Wolfe, V., Fitch, J., & Cornell, R., *Acoustic prediction of severity in commonly occurring voice problems*. Journal of speech and hearing research, vol. 38, no. 2, pp. 273-279, 1995.
DOI: <https://doi.org/10.1044/jshr.3802.273>
- [15] Sorensen, D., Horii, Y., *Frequency characteristics of male and female speakers in the pulse register*. Journal of Communication Disorders, vol. 17, no. 1, pp. 65-73, 1984.
DOI: [https://doi.org/10.1016/0021-9924\(84\)90026-1](https://doi.org/10.1016/0021-9924(84)90026-1)
- [16] Max, L., Mueller, P. B., *Speaking fo and cepstral periodicity analysis of conversational speech in a 105-year-old woman: Variability of aging effects*. Journal of Voice, vol. 10, no. 3, pp. 245-251, 1996.
DOI: [https://doi.org/10.1016/S0892-1997\(96\)80005-1](https://doi.org/10.1016/S0892-1997(96)80005-1)
- [17] Regner M. F., Tao C., Jiang J. J., Zhuang P. *Onset and Offset Phonation Threshold Flow in Excised Canine Larynges*. Laryngoscope, vol. 118, no. 7, pp. 1313-1317, 2008.
DOI: <https://doi.org/10.1097/MLG.0b013e31816e2ec7>
- [18] Seong-Tae Kim, Chul min Ahn, Soon Yuhl Nam, *The Study for Voice Onset Types in Benign Vocal Fold Lesions*. The Journal of the Korean Society of Logopedics and Phoniatics, vol. 20, no. 2, pp. 131-135, 2009.
- [19] Ahn Hwoe Young. *Analysis of voice*, Seoul: Jinsu. 1992.
- [20] Yu Jeong Shin, Ki Hwan Hong, *Cepstral Analysis of Voice in Patients With Thyroidectomy*. Clinical and Experimental Otorhinolaryngology, vol. 9, No. 2, 2016.
DOI: <https://doi.org/10.21053/ceo.2015.00199>
- [21] Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody, M. M., Sataloff, R. T. *Cepstral peak prominence: a more reliable measure of dysphonia*. Annals of Otolaryngology, Rhinology & Laryngology, vol. 112, no. 4, pp. 324-333, 2003.
DOI: <https://doi.org/10.1177/000348940311200406>
- [22] Choi Seong Hee, Choi Chul Hee. *The Stability and Variability based on Vowels in Voice Quality Analysis*. Journal of Korean Society of Speech Sciences, vol. 7, no. 1, pp. 79-86, 2015.
DOI: <https://doi.org/10.13064/KSSS.2015.7.1.079>

신 유 정(Yu-Jeong Shin)

[정회원]



- 2010년 8월 : 이화여자대학교 언어 병리학과 (언어병리학석사)
- 2014년 2월 : 전북대학교 임상언어 병리학과 (임상언어병리학박사)
- 2014년 4월 ~ 현재 : 호원대학교 언어치료학과 조교수

<관심분야>

음성장애, 의사소통장애, 언어치료