

# 후두미세수술 전후 /아/의 음향적 특성 비교

황연신(연세대), 성철재(충남대)

## <차 례>

- |                |                     |
|----------------|---------------------|
| 1. 머리말         | 3.3.2. 주파수 대역별 HNR  |
| 2. 선행 연구       | 3.3.3. F0           |
| 2.1. 국내 연구     | 3.3.4. 포먼트의 주파수 대역폭 |
| 2.2. 국외 연구     | 4. 결과               |
| 3. 실험          | 4.1. HNR            |
| 3.1. 녹음 및 디지털화 | 4.2. 주파수 대역별 HNR    |
| 3.2. 피험자       | 4.3. F0             |
| 3.3 분석 방법      | 4.4. 포먼트의 주파수 대역폭   |
| 3.3.1. HNR     | 5. 마무리              |

## <Abstract>

### Comparative Study on the Acoustic Characteristics of the Korean Vowel /a/ before and after LMS

Yeon-Sin Hwang, Cheol-Jae Seong

The aim of this study is to show the differences in acoustic parameters between a pathological voice /a/ caused by vocal polyp and a normal voice /a/ produced after LMS (Laryngeal Microscopic Surgery). It was expected that voices of two kinds could be analyzed effectively in terms of HNR in specific frequency bands than in all frequency bands. For this study, 10 patients' voice were recorded before and after LMS and then were manipulated in terms of four acoustic parameter. It was found out that (a) frequency bands of 500 Hz in the range of 1,000 Hz to 4,000 Hz were very useful to obtain HNR values; (b) frequency bands in the range of 1,248 Hz to 5,500 Hz on a log scale were very useful to obtain HNR values; (c) F0 dropped after LMS but not significantly; (d) the bandwidth of the second formant (B2) decreased significantly after LMS, while that of the first formant (B1) decreased after LMS but not significantly.

\* Keywords: HNR divided by frequency band, LMS, Vocal fold polyp, F0.

## 1. 머리말

본 논문의 목적은 후두미세수술(laryngeal microscopic surgery)을 받은 후 청지각적으로 음질 향상을 보인 성대 폴립(vocal fold polyp) 환자의 수술 후 음성 개선의 여부를 음향학적 매개변수(acoustic parameter)를 통하여 수술 전과 비교하고 그 결과가 어떠한지 연구하는 것이다<sup>1)</sup>.

지금까지 후두미세수술 전후 음성을 음향학적으로 비교·분석한 연구들에서는 대부분 MDVP(Multi-Dimensional Voice Program)<sup>2)</sup>의 매개변수 중 jitter(주파수 변동률)<sup>3)</sup>, shimmer(진폭 변동률)<sup>4)</sup>, NHR(Noise to Harmonic Ratio; 조화음 대 소음 비율)<sup>5)</sup>을 이용하여 수술 전후 차이를 보여주었는데, 이렇게 많은 연구들<sup>6)</sup>에서 동일한 매개변수들을 통해 분석하였으므로 연구 간 변별성을 찾기 힘든 결과를 가져왔다. 또한 MDVP의 NHR 값은 jitter와 shimmer 값과는 달리 수술 전후 차이를 변별력 있게 제시하지 못했는데 왜 그러한지 그 이유에 대한 명확한 설명을 그간 연구들에서 찾기 힘들었다. 따라서 본 논문은 [4], [5], [6] 등의 연구에서 수술 전후 음성의 차이를 변별력 있게 제시하지 못했던 매개변수 HNR(Harmonic to Noise Ratio)<sup>7)</sup> 값을 보완하여 HNR이란 매개변수에 변별력을 주고자 한다. 이를 위해 주

- 
- 1) 본 논문은 제 1 저자의 석사 학위 논문 내용 중 일부를 발전시킨 것이다[19].
  - 2) MDVP(Multi-Dimensional Voice Program)는 가장 많이 활용되는 음성 분석 기구 중 하나이다. 그 이름에서 알 수 있듯이 음성을 다양한 측면에서 분석할 수 있다는 것이 가장 큰 장점이다. ‘음성’의 특성이 매우 가변적이고 다변적임을 고려할 때, 이를 분석함에 있어 다차원적인 접근 방법을 사용한다는 것은 매우 유효적절한 일이다. 그렇게 때문에 MDVP의 정확하고 적절한 활용은 음성 장애 환자의 병적 음성 뿐 아니라 정상 음성의 특질에 대해 파악하는 데에도 큰 도움을 줄 수 있다[1].
  - 3) 피치 주기와 주기 간 음도 변이의 규칙성·불규칙성을 상대적으로 평가하는 변인이다. 상대적인 평가이므로 Percentage(%) 단위로 계산되며 정상 역치는 1.04 %이다[1].
  - 4) 피치 주기와 주기 간 음성 강도 변이의 규칙성·불규칙성을 상대적으로 평가하는 변인이다. 단위는 Percentage(%)이며 MDVP에서 제시하는 정상 역치는 3.81 %이다[1].
  - 5) 70~4,500 Hz 사이에 존재하는 조화음과 1,500~4,500 Hz 사이에 존재하는 비정상 조화음간 비율의 평균치를 보여준다. MDVP에서 제시하는 정상 역치는 0.19이다[1].
  - 6) 여기에 해당하는 선행 연구로 참고 문헌 [4], [5], [6], [7], [8], [9]를 들 수 있다.
  - 7) NHR이 조화음에 대한 소음의 비율을 나타내는 것이라면, HNR은 소음에 대한 조화음의 비율을 나타낸다. 즉 NHR은 조화음의 수치를 분모로, 소음의 수치를 분자로 놓아 소음 비율의 평균치를 구하고, HNR은 소음의 수치를 분모로, 조화음의 수치를 분자로 놓아 조화음 비율의 평균치를 구한다. 결론적으로 NHR과 HNR은 구하려는 대상이 다를 뿐 궁극적으로 뜻하는 바는 같다. 다시 말해 NHR은 소음과 조화음의 비율을 소음을 통해 제시하는 것이고, HNR은 소음과 조화음의 비율을 조화음을 통해 제시하는 것이다. 다만 NHR은 1을 전체로 보고 그 비율을 소수점으로 나타내는데 반해, HNR은 상용로그

파수 대역별로 HNR 값을 살펴봄으로써 수술 전후 통계적으로 유의미한 차이를 나타내는 주파수 대역(frequency band)을 구하려 한다. 이렇게 하는 이유는 조화음과 소음의 값이 목소리 특성(voice quality)에 따라 민감하게 변화하는 주파수 대역이 있다고 생각하기 때문이다. 따라서 기존에 사용했던 모든 주파수 대역에서의 HNR 값이 아니라 특정 주파수 대역에서의 HNR 값을 구하여 HNR를 이용해 좀 더 면밀한 분석을 하고자 한다.

본 논문에서 주파수 대역별 HNR 포함하여 분석을 위해 사용한 음향학적 매개변수는 총 4개가 있는데 그 목록은 아래와 같다.

- ① HNR(소음 대 조화음 비율, Harmonic to Noise Ratio)
- ② 주파수 대역별 HNR
- ③ FO(기본 주파수)
- ④ 포먼트(formant)의 주파수 대역폭(bandwidth)

먼저 조화음의 특성과 관련된 매개변수들인 HNR과 주파수 대역별 HNR을 이용하여 수술 전후 음성의 차이를 알아 볼 것이다. 조화음의 비율은 정상 음성에서 보다 병리적 음성(pathologic voice)에서 감소하게 되므로 본 매개변수들을 이용하여 수술 전후 음성의 차이를 살펴보는 것은 적절한 일이라 판단한다.

둘째, 공명의 특성과 관련된 매개변수인 포먼트의 주파수 대역폭을 사용하여 음향학적 분석을 할 것이다. 향상된 발성 즉 양질의 음원은 공명의 향상을 가져온다. 따라서 주파수 대역폭을 통해 공명의 향상 정도를 살펴보는 것도 수술 전후 음성의 차이를 밝히는 중요한 작업이 될 것이다.

마지막으로 음높이를 나타내는 FO를 통하여 정상적 음높이와 병리적 상태에서의 음높이가 어떠한 차이가 있는지 살펴볼 것이다.

## 2. 선행 연구

### 2.1. 국내 연구

후두미세수술은 주로 성대 폴립(vocal fold polyp), 성대 결절(vocal fold nodules), 라인케 부종(Reinke's edema) 등에 행하여진다. 수술의 일반적인 원칙은 정확하고 섬세한 수기(手技)로 병변을 완전히 제거하고 정상 조직을 최대한 보존하는 것이다[2]. 이를 위해 겸자(鉗子)<sup>8)</sup>로 종괴(腫塊)<sup>9)</sup>를 잡고 가위나 칼로 병변 부위를 제

(log10)를 씌우고 10을 곱해서 단위 dB로 결과 값을 나타내므로 표면적으로 나타나는 수치는 NHR과 차이가 난다.

거하는 방법을 쓰기도 하며, 레이저로 결절을 기화(氣化)하는 방법이 사용되기도 한다.

성대 폴립은 성대에 있는 분비액으로 가득 찬 덩어리를 말한다. 성대를 지나치게 사용하면, 성대 최대 접촉점의 점막에 출혈이 다소 있을 수 있으며 이 출혈과 연관된 자극으로부터 결국 분비액으로 가득 찬 덩어리가 생기는데 이 덩어리를 성대 폴립이라 한다[3]. 성대 폴립은 성대 마비(vocal fold paralysis)처럼 성대가 전혀 움직이는 못하는 상태가 아니라, 성대 표면에 생긴 분비액 덩어리로 인해 성대의 내전과 외전이 완벽하게 이루어지지 않는, 즉 정상적인 성대 접촉이 방해 받는 상태라고 볼 수 있다. 따라서 이 덩어리, 즉 폴립만을 제거하는 수술을 통해 효과적인 음질 향상이나 정상 음성으로의 회복을 기대할 수 있다고 보고되어 왔다.

본 논문은 후두미세수술을 받은 성대 폴립 환자의 수술 전과 후의 음성을 연구 대상으로 삼았다. 왜냐하면 후두미세수술로 폴립이 제거되면 완전한 내전과 외전이 가능하여 정상 음성으로 발성할 수 있다고 여겨지기 때문이다.

성대 폴립 환자의 후두미세수술 전후 음성의 차이를 알아본 국내 선행 연구들의 연구 방향을 살펴보면 대체적으로 다음 네 가지로 나누어 볼 수 있다. 첫째, 청지각적 분석(auditory-perceptual analysis), 음향학적 분석(acoustic analysis), 공기역학적 분석(aerodynamic analysis) 방법 중 어떠한 분석 방법이 수술 전후 차이를 가장 잘 보여주는지 살펴본 연구, 둘째, 수술 전후 음성과 정상인의 음성을 비교하여 수술 후 음성이 정상 범주로 되돌아 왔는지 정상군과 비교를 통하여 알아본 연구, 셋째, 음향학적 분석 시 어떠한 음향학적 매개변수가 수술 전후 차이를 가장 명확하게 보여주는지 밝힌 연구, 마지막으로 음성 장애별로 어떠한 질환이 수술 결과가 가장 좋았는지 비교하고 밝힌 연구가 그것이다.

이러한 연구 방향을 참고삼아 최근의 주요 국내 연구를 살펴보면 다음과 같다.

[4]는 후두미세수술을 받은 성대 폴립, 성대 결절, 라인케 부종의 27명 환자들을 대상으로 수술 전후 음성 /아/를 녹음하고 음향학적 분석을 하였다. 연구 결과, FO 값은 수술 후 대부분 환자군에서 약간의 증가를 보였으나 통계적으로 유의미한 변화는 없었다. jitter 값은 성대 폴립 환자군에서 수술 후 유의미하게 감소하였으나 성대 결절, 라인케 부종 환자군에서는 통계적으로 유의미한 감소는 없었다. shimmer 값의 경우 성대 폴립 환자군과 성대 결절 환자군에서 수술 후 유의미하게 감소하였으나, 라인케 부종 환자들에서는 통계적으로 유의미한 감소가 없었다. NHR은 모든 환자군에서 약간의 감소가 관찰되었으나 통계적으로 유의미한 변화는 없었다. 결론적으로, 성대 폴립에서는 수술 후 jitter와 shimmer 값이 유의미하게 감소하였으며, 성대 결절에서는 수술 후 shimmer 값만이 유의미하게 감소하였고,

8) 날이 서지 않은 가위 모양으로 생긴 외과 수술 기구. 조직이나 기관을 집어서 누르거나 고정하는 데에 쓴다.

9) 조직이나 장기의 일부에 생긴 경계가 분명한 종기.

라인케 부종에서는 수술 후 음향학적 매개변수들의 유의미한 변화가 없었다. 따라서 후두미세수술 후, 병변에 따른 결과를 비교하기 위한 음향학적 매개변수는 jitter와 shimmer가 사용될 수 있을 것이라고 하였다. 또한 shimmer는 성대 폴립과 성대 결절 두 환자군에서 유의미하게 감소되었으므로 수술 전후 성대 기능을 가장 잘 반영할 수 있는 음향학적 매개변수라고 하였으며, 다른 매개변수들에 비해 가장 뚜렷한 변화를 보인 것으로 보아 음성의 질과 민감하게 관련이 있다고 하였다.

[4]는 비교적 많은 숫자의 피험자를 확보하여 분석함으로써 연구의 유의성을 확보했으나 다음과 같은 몇 가지 미비한 점이 관찰된다. 우선 수술 전후 F0나 NHR이 왜 통계적으로 유의미한 차이가 나지 않았는지에 대한 구체적인 설명이 없다. 그리고 ‘수술 전후 shimmer의 수치 변화가 가장 크다. 따라서 성대 기능을 가장 잘 반영하는 매개변수다’라고 하였는데 이는 무리한 결론이라 여겨진다. 수술 후 정상 범주로의 회복 정도를 고려하지 않고 어떤 매개변수의 수치 변화가 가장 크다고 해서 그것이 성대 기능을 가장 잘 반영하는 매개변수라고 결론 내리는 것은 무리가 있기 때문이다. 혹은 [4]의 논의를 그대로 따라간다면, ‘성대 폴립은 jitter와 shimmer 값이 통계적으로 유의미한 차이가 나므로 성대 점막의 진폭과 파동에서 모두 회복을 보였으며, 성대 결절은 shimmer 값만이 통계적으로 유의미한 차이가 나므로 성대 점막의 진폭에서만 큰 회복을 보였다’라고 결론을 내리는 것이 더 타당하다고 여겨진다.

두 번째로 언급할 선행 연구는, 양성 후두 질환군(성대 결절 3예, 성대 폴립 27예, 성대 구증(sulcus vocalis)<sup>10)</sup> 1예, 후두 낭종(laryngeal cyst)<sup>11)</sup> 1예, 후두염(laryngitis)<sup>12)</sup> 1예) 33명과 정상군 35명을 대상으로 /우/, /오/, /아/, /에/, /이/를 녹음하고 음향학적으로 분석한 연구이다[5]. 분석 결과, 정상군과 양성 후두 질환군 사이에 통계적으로 유의미한 매개변수는 shimmer, sh dB(shimmer in dB)<sup>13)</sup>, vAm(peak-amplitude variation)<sup>14)</sup>, SPI(Soft Phonation Index)<sup>15)</sup>로 나왔다.

[5]의 경우 정상군과 양성 후두 질환군을 감별하는 통계적으로 유의미한 매개

- 10) 성대의 유리연을 따라 평행하게 흡이 나 있는 것을 성대구라 하고 이로 인해 음성 장애가 발생하는 것을 성대 구증이라 한다. 발생 원인으로 인후두 역류, 성대낭의 파열, 선천성·노인성 후두증, 그 외 염증성 질환, 후두 손상 등이 거론되고 있다.
- 11) 성대 낭종은 점액선의 분비관이 막혀서 발생한다. 이는 후두 양성 종양의 5%를 차지하며 후두 어느 부위에서도 발생할 수 있으나, 점액을 분비하는 선조직이 많은 후두 상부, 그 중에서도 후두개 설측면에 자주 발생한다.
- 12) 성대가 박테리아 또는 바이러스에 감염이 되어 발생하며 성대가 붓고 빨간색으로 변하는 질환이다.
- 13) 피치의 주기와 주기 간 음성 강도 변이의 규칙성·불규칙성을 상대적으로 평가하는 변인이다. shimmer와 같은 개념이나 단위를 dB로 쓰는 것이 shimmer와 다르다. MDVP에서 제시하는 정상 역치는 0.35 dB이다[1].

변수를 명시적으로 제시한 점이 우수하다고 여겨지나, 연구 결과를 기술만 하였을 뿐 그 결과에 대한 타당한 설명이 전반적으로 부족하다. 즉 제시한 매개변수 대부분이 ‘진폭의 동요’와 관련된 것인데 왜 그러한지, 또한 ‘주파수 동요’와 관련된 매개변수는 통계적으로 유의미한 차이가 없는데 이 역시 왜 그러한지 타당한 설명을 하고 있지 않다.

세 번째로 언급할 선행 연구는 성대 양성 병변(성대 폴립, 성대 결절, 라인케 부종, 후두 각화증(hyperkeratosis)<sup>16</sup>, 성대 낭종 등)으로 후두미세수술을 받은 환자를 대상으로 공기역학적 검사와 음향학적 검사를 실시한 논문이다[6]. 결과를 살펴보면, 성대 폴립의 경우 jitter와 평균 호기율(mean air flow)<sup>17</sup>이 수술 후 통계적으로 유의미하게 감소하였고, 성대 결절의 경우 jitter와 shimmer가 수술 후 통계적으로 유의미하게 감소하였으며, 성대 낭종의 경우 최대발성지속시간(maximum phonation time)<sup>18</sup>이 수술 후 통계적으로 유의미하게 증가하였다. 라인케 부종의 경우 통계적으로 유의미한 차이를 보이는 음성 지표는 없었으나 전반적인 음성 매개변수 값의 향상은 있었고, 후두 각화증의 경우 수술 후에 뚜렷한 음성 매개변수 값의 향상은 없었다고 하였다. 결론적으로, 음향학적 검사와 공기역학적 검사는 음성 장애 환자의 수술 전후 음성 평가에 있어 객관적이고 용이한 검사로 여겨진다고 하였다. 그리고 성대 폴립과 결절의 경우 jitter가 성대 기능의 상태를 반영하는 예민한 매개변수이며, 이는 후두미세수술 후 음성 평가에 있어 중요한 지표로 이용될 수 있을 것이라고 하였다. 그리고 NHR은 모든 질환군에서 통계적으로 유의미하지는 않지만 수술 후 증가를 보인다고 하였다.

[6]은 두 검사—공기 역학적 검사와 음향학적 검사—를 함께 실시하여 다른 연구들과 구별되는 연구의 다양성을 획득하였으나, 두 검사를 실시하였음에도 불구하고 두 검사의 결과를 개별적으로 제시하였을 뿐 두 검사 간 상관관계를 밝히지 않아 두 가지 검사를 실시한 의의가 상쇄되는 결과를 가져왔다. 또한 수치가 낮은 수록 양질의 음성임을 나타내는 매개변수 NHR이 수술 후 오히려 그 값이 증가한

14) 접해 있는 주기(Cycle) 간의 음성 강도 변이에 대한 표준편차를 상대적으로 보여주고 있다. MDVP에서 제시하는 정상 역치는 8.20 %이다[1].

15) 70~1,600 Hz 사이의 저주파에 존재하는 조화음 에너지(harmonic energy)와 1,600~4,500 Hz 사이에 고주파에 존재하는 조화음 에너지(harmonic energy) 간의 비율을 보여준다(고도홍 외, 2001).

16) 성대 안쪽 가장 자리에 케라틴(단백질이 경화된 것)이 축적된 질환이다.

17) 일반적으로 성인은 90~140ml/sec가 정상이다. 이때 90ml/sec 이하는 성대의 과기능자, 200ml/sec 이상은 성대의 저기능자라고 판단할 수 있다[1].

18) 10~18초면 정상으로 판단한다. 전문적인 가수의 경우에는 30초 이상 연장하는 것도 가능하다. 성대 결절이나 폴립, 마비가 존재하는 경우 최대연장발성 길이가 짧다. 대부분 10초를 넘기지 못한다[1].

이유 역시 설명하고 있지 않다. 마지막으로 [6]의 shimmer에 대한 연구 결과는 'shimmer는 성대 폴립과 성대 결절 음성 평가에 유용한 매개변수다'라고 한 [4]의 연구와 상반된 것이기도 하다.

네 번째로 언급할 선행 연구는 성대 결절 혹은 성대 폴립으로 진단 받고 후두 미세수술을 받은 환자와 정상군을 대상으로 /에/를 녹음하고 음향학적 분석을 실시한 논문이다[7]. [7]의 연구 결과는 크게 여섯 개로 나뉘는데 결과를 간단히 살펴보면 다음과 같다. 첫째, 성대 결절과 성대 폴립의 경우, 수술 전후 음성을 객관적으로 평가할 수 있는 음향학적 매개변수는 F0의 변동폭(F0 range)<sup>19)</sup>과 jitter, shimmer이었다. 둘째, 수술 후 1 일째 성대 결절과 성대 폴립의 음향학적 매개변수 값은 수술 전에 비해 통계적으로 유의미하게 감소하였으며, 수술 후 28 일째 음향학적 매개변수 값은 대부분 정상 범주 속하게 되었다. 셋째, 남녀 간 통계적으로 유의미한 없었다. 넷째, 성대 폴립의 경우, F0의 변동의 폭은 수술 후 기간이 경과할수록 통계적으로 유의미한 감소를 보였으며, 이 중 남성은 정상 범위에 가깝게 회복되었고, 여성은 정상 범주 내로 완전히 속하게 되었다. 다섯째, jitter 값은 수술 후 기간이 경과할수록 통계적으로 유의미한 감소를 보였고, 그 중 남성은 정상 범위로 회복되었으며 여성은 정상군보다도 더 향상된 음성으로 회복되었다. 여섯째, shimmer 값은 수술 후 통계적으로 유의미하게 감소하였고 정상 범위로 회복되었다. 일곱째, 전반적으로 성대 폴립의 음향학적 변화폭이 성대 결절보다 컸다.

[5]의 연구는 남녀별 차이와 시간 경과에 따라 차이를 자세히 다루어 다른 연구들에서 다루지 않은 부분을 다루어 연구의 변별성을 획득하였으나, 다른 연구들과 마찬가지로 매개변수 값을 제시하였을 뿐 왜 그런 결과 값이 나왔는지 구체적인 설명이 부족하다고 여겨진다.

다섯 번째로 언급할 연구로, 새로 개발된 다채널 음성 분석기를 이용하여 정상군 50명(남자 22명, 여자 28명)과 후두 내시경 검사 상 성대 점막 병변이나 기능적인 이상이 확인된 환자 57명(남자 17명, 여자 40명)을 대상으로 jitter, shimmer, SPI, HNR, F0, 음성 강도, 평균 호기율을 살펴본 논문을 들 수 있다[8]. [8]의 결과를 살펴보면, 정상군과 환자군에서 F0의 차이는 없었지만 음성 강도는 여자 환자군에서 통계적으로 유의미하게 높은 값을 보였다고 하였다. 이러한 결과는 환자군에서 정상 음성을 형성하기 위해 무의식적으로 과도한 성문하압을 형성하려 하는 경향이 있음을 의미하고 이는 음성 강도 조절이 매우 어려웠음을 시사한다고 하였다. jitter와 shimmer, 평균 호기율은 환자군에서 증가하였고, SPI는 통계적으로 유의미한 차이는 없었으나 남성 환자군이 정상군에 비해 큰 변이(variance)를 보였다고 하였다. HNR은 남성 환자군만 통계적으로 유의미한 차이를 보였고 여성 환자군은

19) 기본주파수의 상대적인 표준편차로서 기본주파수의 변이에 대해 보여준다. MDVP에서 제시하는 정상 역치는 1.10%이다(고도홍 외, 2001).

통계적으로 유의미한 차이가 없었다고 하였다. 그 이유는 여성 환자군의 음성 강도가 정상군에 비해 통계적으로 유의미하게 높았기 때문이라고 하였다. HNR은 음성 강도와 통계적으로 유의미한 상관관계를 보이는데, 이는 음성 강도가 증가함에 따라 소음 부분보다는 조화음에 더 높은 에너지가 집중되기 때문이라고 했다.

마지막으로 언급할 선행 연구는 성대 폴립 또는 성대 결절로 진단 받고, 후두 미세수술을 받은 환자 94명을 대상으로 실험한 논문이다[9]. 음성 녹음은 수술 전 날과 수술 후 2주일째 시행하였으며, 사용한 문장은 ‘높은 산에 올라가 맑은 공기를 마시며 소리를 지르면 가슴이 활짝 열리는 듯하다’이다. 분석 결과는 다음과 같다. 후두미세수술을 통하여 성대 병변 자체를 제거해줌으로써 수술 후 jitter와 shimmer의 수치가 호전됨을 알 수 있었으나 FO 평균값의 경우 남녀에서 서로 반대 현상을 보여, 수술 후에 남자는 감소하고 여자는 증가하였다고 하였다. 녹취 시점이 수술 후 2주째 이루어졌다는 점에서 이와 같은 현상을 최종적인 변화라고 단정 지을 수는 없으나, 병변 자체가 제거되어 편안한 발성이 가능해 짐으로써 나타난 현상으로 생각된다고 하였다. 또한 FO의 변화를 수술 후 후두 미세 구조의 회복 과정에 따른 변화를 간접적으로 예측할 수 있는 부분으로 생각한다고 하였다. 따라서 평균 FO의 변화 여부에 대하여서는 수술 후 조금 더 긴 회복 기간을 통한 추가적인 연구가 필요하다고 하였다.

이와 같이 대부분 연구에서 거의 동일한 음향학적 매개변수들로 분석하고 결과를 내어 제시하였으나 많은 연구들에서 왜 이러한 결과가 나왔는지에 대한 구체적이고 세세한 설명을 하고 있지 않다. 특히 특정 음향학적 매개변수가 수술 전후 차이를 가장 잘 나타내주는 것이라고 제시하고 있으나 왜 그러한지 그 이유를 설득력 있게 설명하지 못한 점이 위에서 제시한 선행 연구들의 가장 큰 아쉬운 점이라 여겨진다.

## 2.2. 국외 연구

본 논문과 관련된 최근의 국외 연구로 구조적 병변(17명은 성대 결절, 8명은 성대 낭종)을 지닌 25명 소년들의 음성(이 중 22명은 발성 장애(dysphonia)를 보이고 3명은 발성 장애를 보이지 않음)과 구조적 병변 없이 발성 장애를 보인 11명 소년들의 음성, 그리고 14명의 정상군을 대상으로 NHR을 조사한 논문을 들 수 있다[11]. 결과를 살펴보면, NHR은 구조적 병변을 가진 음성과 구조적 병변 없이 발성 장애만을 가진 음성들 모두에서 유의미하게 높았다. 다만 로지스틱 회귀분석을 통해 살펴본 결과 NHR은 발성 장애 발생만을 설명하는데 쓰일 수 있는 매개변수라고 하였다. 즉 NHR이 0.01 증가하면 발성 장애가 발생할 확률이 두 배로 늘어난다고 하였다. 또한 NHR은 청지각적으로 병리적 발성 장애 음성을 진단한 후 이를 입증하는 지표로 쓰일 수 있으며 시간이 흐름에 따라 병리적 발성 장애 음성



이 어떻게 변화하는지 평가하는데 유용하게 쓰일 수 있다고 하였다. 하지만 NHR은 청각적 평가의 보조적 도구로서 쓰일 수 있을 뿐이지 대체 도구로 쓰일 수는 없다고 하였다.

또 하나의 국외 연구로 88명의 정상군과 98명의 성대 점막에 병변이 있는 환자군(성대 결절 25명, 성대 폴립 32명, 성대 유두종(laryngeal papilloma)<sup>20)</sup> 24명, 성대암(glottic carcinoma) 17명)과 50명의 편측 성대 마비 환자군(unilateral vocal fold paralysis)을 대상으로 음향학적 분석을 하고, 성대 폴립과 성대 결절 환자의 음성을 후두미세수술 전후 음향학적 검사와 청지각적 검사를 통해 비교한 논문을 들 수 있다[16]. 결과는 다음과 같다. 첫째, NNE<sup>21)</sup>가 두 환자 집단군에서 모두 유의미하게 높아 정상군과 비정상군을 구별하는 최적의 매개변수라고 하였다. 둘째, GRBAS Scale<sup>22)</sup>로 분석한 청지각적 검사와 jitter, shimmer, NNE로 분석한 음향학적 검사 모두에서 수술 전후 통계적으로 유의미한 차이를 나타냈다. 따라서 이 논문에서 사용한 모든 분석 방법이 정상군과 환자군의 음성을 비교하는데 도움을 줄 수 있다고 하였으며 수술 전후 음성 상태를 평가하는 도구로도 쓰일 수 있다고 하였다.

### 3. 실험

#### 3.1. 녹음 및 디지털화

음성 자료는 서울대학교 병원 음성 검사실에서 녹음되었다. 구체적인 녹음 방법은 다음과 같다. 피험자를 편안한 자세로 앉게 한 후 마이크를 피험자 입과 45도 정도 옆에 위치시키고 거리는 10~15 cm 떨어뜨려 놓은 후, 평상시와 같은 높이와 크기로 /아/를 연장 발성하라고 지시하였다. 동일한 피험자의 수술 전 음성과

20) 유두종은 어둡고 축축한 기도 공동(空洞)에서 발생하는 바이러스로 인해 발생한 물사마귀처럼 생긴 조직이다.

21) 발성시 산출되는 소음 에너지를 측정하는 것으로서 조화음보다는 소음 에너지 그 자체에 초점을 맞추고 있다. 실제로 병리적 음성에서 NNE(Normal Noise Energy)를 많이 사용하고 있는데 그 이유는 NNE가 성문 폐쇄(Glottal Closure)의 정도와 관련이 있기 때문이다[12].

22) GRBAS Scale이란, 조조성(粗糙性(Rough: R)), 기식성(氣息性(Breathy: B)), 노력성((努力性(Strained: S)) 및 무력성(無力性(Asthenic: A))으로 나누어 그 정도를 0, 1, 2, 3의 4단계로 평가하는 방법이다. 이때 G(Grade)는 총괄적인 인상을 나타내는 것으로서 0은 전혀 쉼 목소리가 없는 정상적인 상태, 3은 쉼 목소리 정도가 가장 심한 상태, 1과 2는 그 중간 상태이다. G의 경우에서와 같이 R, B, A, S에 있어서도 0은 각각 청각적 인상이 전혀 없는 상태이고 3은 가장 심한 정도를 나타낸다.

수술 후 음성을 녹음하였는데, 수술 전 음성 자료는 서울대학교병원 이비인후과에서 성대 폴립으로 진단을 받은 후 녹음한 것이고, 수술 후 음성 자료는 수술 후 한 달간 음성 휴식을 갖게 한 후 최소 한 달 후에 녹음한 것이다.

실험에 사용된 음성자료를 처음 녹음한 시기는 2006년 1월 26일이었고 마지막으로 녹음한 시기는 2007년 5월 4일이었다. 이 중 2006년 1월부터 11월까지의 음성 자료는 KAY사의 CSL Model 4300B를 이용하여 녹음 및 디지털화하였고(표본 추출률 50,000 Hz, 양자화 16 bit), 2006년 12월부터 2007년 5월까지의 자료는 KAY사의 CSL Model 4500을 매개로 녹음 및 디지털화하였다(표본 추출률 44,100 Hz, 양자화 16 bit). 녹음 시 사용한 마이크는 SHURE 사의 단일 지향성 다이내믹 마이크인 SM48이었다.

본 논문에서 사용한 음성 파형은 원 자료(raw data)를 표본 추출률 10,000 Hz로 다운샘플링하고 16 bit로 양자화 하여 만든 것이다. 원 자료를 10,000 Hz로 다운샘플링한 이유는 모음(vowel)의 경우 10,000 Hz의 나이퀴스트 주파수(nyquist frequency)인 5,000 Hz 이하 대역에서 모음의 음가를 결정하는 주요 포먼트는 물론 스펙트럼의 주요 구성성분이 관찰되기 때문이다. 다운 샘플링에 사용한 음성 분석 프로그램은 praat 4.5.17이었다.

수치 정리, 표와 그림 작업을 위하여 microsoft 사의 excel 2003을 이용하였고, 통계 분석을 위하여 SPSS 12.0 KO를 이용하였다.

### 3.2. 피험자

피험자는 서울대학교병원 이비인후과에서 성대 폴립으로 진단 받고 후두미세 수술을 받은 환자이다. 수술 전후 음성 파형이 확보된 16명 중 수술 후 음질 향상을 보인 14명에서 다시 남성 10명을 선택하여 분석 대상으로 삼았다. 본 연구에서 여성을 제외한 이유는 남성보다 훨씬 적은 4명의 여성 피험자를 포함할 경우 여성이라는 성별이 또 하나의 변인(variable)으로 작용하여 결과에 영향을 미칠 수 있기 때문이었다.

수술 후 음질 향상은 청지각적 척도인 'GRBAS Scale'을 이용하여 판가를 하였는데, 실험자가 청지각적으로 판단하여 수술 후 GRBAS에서 1점 이상 향상을 보인 경우 실험 자료에 포함하였다. 피험자의 배경 정보는 아래 표와 같다.

<표 1> 피험자 배경 정보

| 이름  | 나이  | 성별 | 진단명       | 수술 전 녹음일    | 수술 후 녹음일    |
|-----|-----|----|-----------|-------------|-------------|
| HJS | 58세 | 남  | Vf. polyp | 2006.01.26. | 2006.03.14. |
| KJH | 32세 | 남  | Vf. polyp | 2006.06.20. | 2006.08.29. |
| LGH | 55세 | 남  | Vf. polyp | 2006.08.24. | 2006.12.22. |
| LJH | 40세 | 남  | Vf. polyp | 2006.03.23. | 2006.05.30. |
| LYD | 54세 | 남  | Vf. polyp | 2006.02.23. | 2006.06.27. |
| MSW | 41세 | 남  | Vf. polyp | 2007.01.05. | 2007.03.23. |
| PJH | 27세 | 남  | Vf. polyp | 2006.12.15. | 2007.02.02. |
| SBS | 32세 | 남  | Vf. polyp | 2006.12.22. | 2007.02.23. |
| WYG | 38세 | 남  | Vf. polyp | 2006.08.23. | 2006.10.23. |
| YYJ | 32세 | 남  | Vf. polyp | 2007.02.23. | 2007.05.04. |

### 3.3. 분석 방법

약 5~10초 정도 연장 발생된 /아/에서 안정 구간 2~3초를 선택하여 분석 대상으로 삼았다. 총 4개의 매개변수를 이용하여 분석하였는데 매개변수 목록을 <표 2>에 제시하였다. 본 논문에서 선택한 4개의 매개변수는 음성 파형에서 값을 얻을 수 있는 것이므로 분석 대상은 음성 파형이 된다.

<표 2> 매개변수 목록과 매개변수가 적용된 구체적 분석 대상

| 매개변수                                      | 분석 대상 |
|---|-------|
| HNR(Harmonic to Noise Ratio, 소음 대 조화음 비율) | 음성 파형 |
| 주파수 대역별 HNR                               | 음성 파형 |
| F0(기본 주파수)                                | 음성 파형 |
| 포먼트의 주파수 대역폭 (formant's bandwidth)        | 음성 파형 |

매개변수 중 HNR과 주파수 대역별 HNR은 성대의 진동과 관련된 발생부의 특징이므로 10,000 Hz로 다운샘플링된 분석객체의 성도전달 특성을 없애기 위해서 deemphasis를 하고 이를 대상으로 필터링 작업을 진행하였다.

#### 3.3.1. HNR

HNR은 소음 대 조화음 비율을 의미한다. praat에서 사용된 HNR을 구하는 과정

은 다음과 같다[13]. 우선 자기상관법을 통해 F0와 음성의 주기성을 구한다. 자기상관법을 적용하는 수식은 (1)과 같다.

$$r_x(\tau) \equiv \int x(t)x(t+\tau)dt \quad (1)$$

수식 (1)에서  $x(t)$ 는 시간 신호를 의미하는 상수이며  $\tau$ 는 상관관계를 알아보기 위해 비교하는 두 시간 신호 사이의 시간차(time lag)를 의미한다. 이 함수는  $\tau$ 가 0일 때 최댓값을 갖는다. 만약  $\tau$ 가 0이 아니더라도  $r_x(\tau)$ 가 최댓값이 나온다면 이는 시간 신호  $x(t)$ 가 주기적이라는 것을 의미한다. 결국  $r_x(0)$ 일 때 주기적 한 신호의 시간  $T_0$ 가 나오게 되므로 수식 (2)와 같은 등식을 얻을 수 있다. 수식 (2)에서  $n$ 은 정수(integer)를 말한다.

$$r_x(nT_0) = r_x(0) \quad (2)$$

결국  $\tau$ 가 0일 때 조화음의 구성 성분(H)인 주기적 신호  $x$ 의  $T_0$  값을 얻게 되므로 수식 (3)과 같은 등식 관계를 얻을 수 있다.

$$r_x(\tau_{\max}) = r_H(T_0) = r_H(0) \quad (3)$$

수식 (4)는 자기상관법을 구하는 수식을 표준화시켜(normalize) 0부터 1까지 값만 나오게 한 수식이다.

$$r'_x(\tau) \equiv \frac{r_x(\tau)}{r_x(0)} \quad (4)$$

수식 (4)를 통해 수식 (5)를 도출할 수 있으며 조화음 이외의 구성 성분은 소음(N)이라 간주할 수 있으므로 수식 (6)을 도출할 수 있다.

$$r'_x(\tau_{\max}) = \frac{r_H(0)}{r_x(0)} \quad (5)$$

$$1 - r'_x(\tau_{\max}) = \frac{r_N(0)}{r_x(0)} \quad (6)$$

HNR은 소음 대 조화음 비율을 dB로 나타낸 것이므로 결론적으로 다음 수식 (7)을 유도할 수 있다.

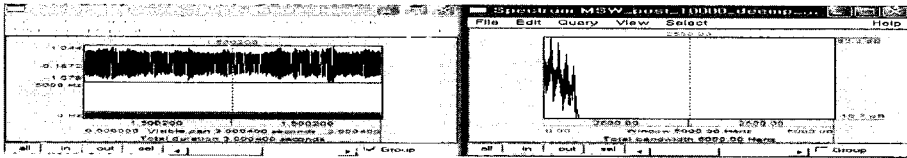
$$HNR = 10 \log_{10} \frac{r'_x(\tau_{\max})}{1 - r'_x(\tau_{\max})} \quad (7)$$

HNR 값을 구하기 위한 절차는 다음과 같다. praat 편집 창에서 음성 파형을 불러 전체를 드래그하여 선택하고, 메뉴 ‘pulses’에서 ‘show pulses’를 선택한 후 다시 그 하부 메뉴인 ‘voice report’를 선택하여 거기서 제공하는 HNR을 이용하여 값을 얻었다.

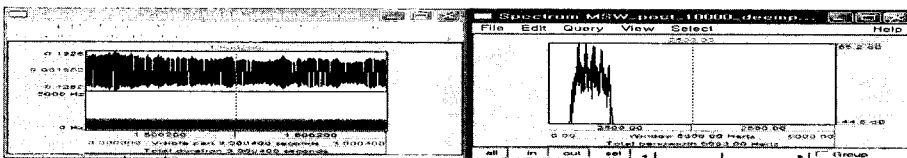
### 3.3.2. 주파수 대역별 HNR

#### 3.3.2.1. 500Hz 대역폭으로 나눈 HNR

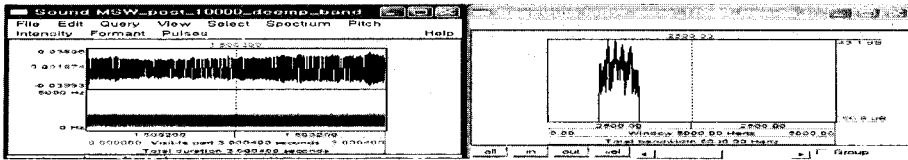
주파수 대역별 HNR 값을 구하기 위해 전체 표본 추출률 10,000 Hz의 나이퀴스트 주파수 5,000 Hz를 500 Hz 단위로 나누어 총 10개 대역으로 나누어 살펴 보았다. 500 Hz 단위로 나눈 이유는 1,000 Hz 간격으로 나누어 분석할 경우 분석 폭이 너무 넓어 세밀하지 못하고 무딘 결과 값을 유도할 수 있기 때문이며, 반대로 500 Hz 이하의 주파수 대역으로 나누어 분석하면 분석하는 절차나 시간이 지나치게 오래 걸려 분석 도구로서 가치가 상실될 수 있기 때문이다. 따라서 본 논문에서는 이 두 가지 문제점을 다 피해갈 수 있는 가장 적절한 주파수 대역을 500 Hz로 간주하고 그렇게 시행하였다. 500 Hz 단위로 나누기 위하여 praat 객체 창에서 메뉴 ‘Synthesize’의 ‘Filter(pass Han band)’를 사용하여 필터링(Filtering)하였다. <그림 1~10>은 피험자 MSW의 주파수 대역별 음성 파형과 스펙트로그램(spectrogram), 그리고 스펙트럼(spectrum) 나타낸 것이다.



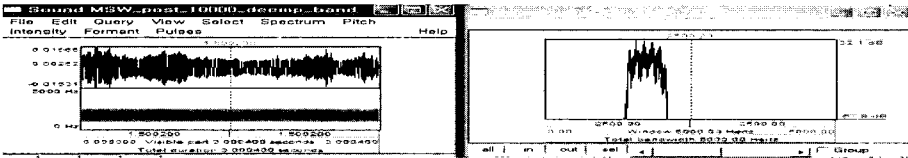
<그림 1> 0~500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



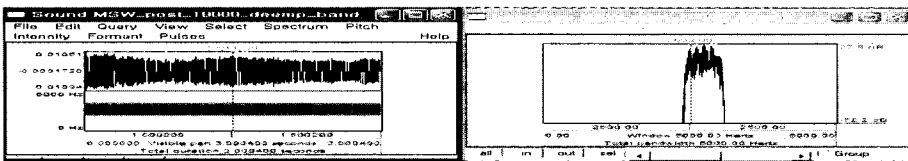
<그림 2> 500~1,000 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



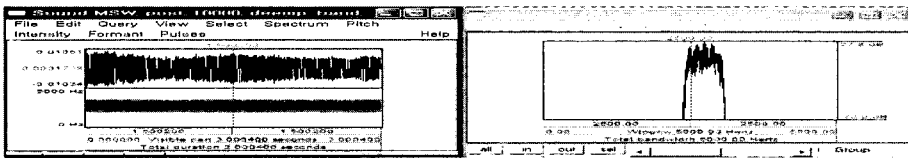
<그림 3> 1,000~1,500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



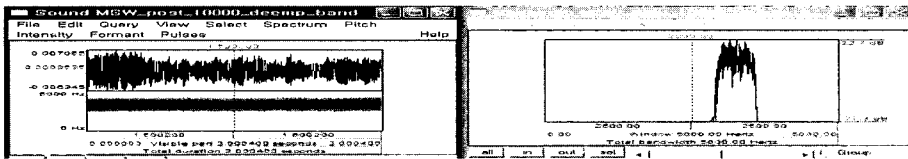
<그림 4> 1,500~2,000 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



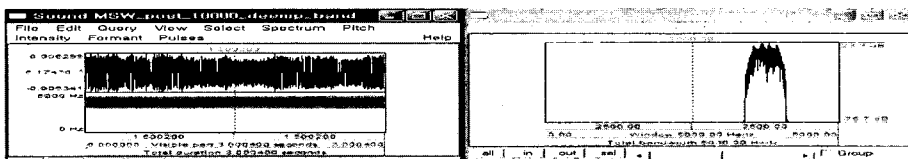
<그림 5> 2,000~2,500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



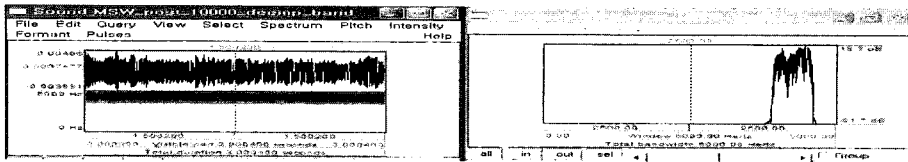
<그림 6> 2,500~3,000 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



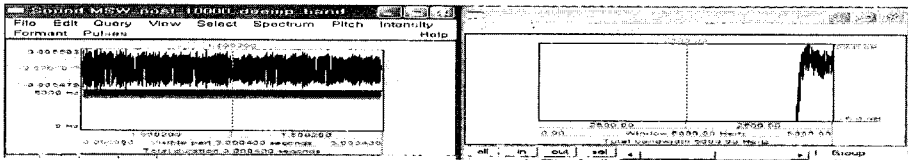
<그림 7> 3,000~3,500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



<그림 8> 3,500~4,000 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



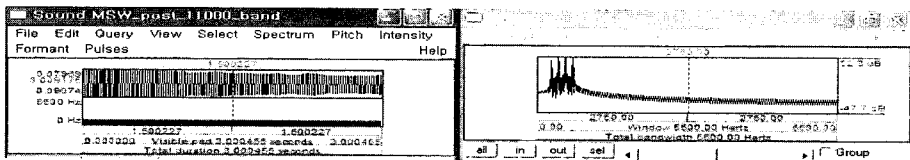
<그림 9> 4,000~4,500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



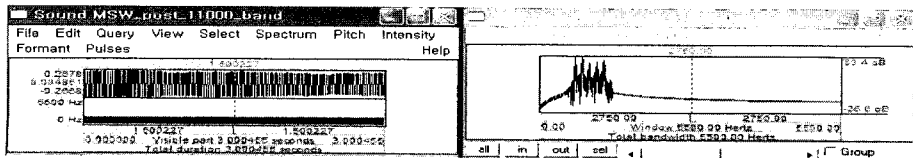
<그림 10> 4,500~5,000 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼

### 3.3.2.2. 로그스케일에 따라 나눈 HNR

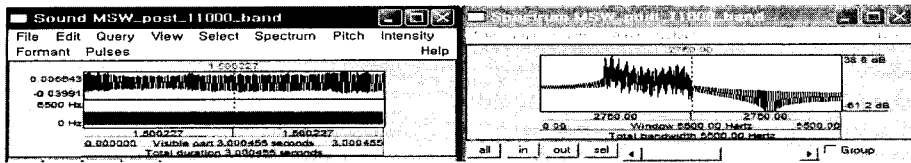
[14]에 의하면 주파수 대역을 나누는 방법 중 가장 합리적인 것은 사람의 청각 기관을 모방해 로그스케일(log scale)에 따라 주파수 대역을 설정하는 것이라고 하였다. 왜냐하면 사람의 청각기관은 저주파수 대역은 정밀하게 구별하여 듣고, 고주파수 대역은 정밀하게 구별하여 듣지 못하기 때문이라고 했다. 따라서 로그스케일로 주파수 대역을 나누었을 때 정상 음성과 병리적 음성을 구별하는 과제에서 더 이상의 성능 향상이 없었다고 즉 가장 최상의 차이를 나타내는 결과를 내었다고 하였다. 본 논문에서도 [14]에 제시된 대로 로그스케일에 따라 4개의 대역폭 ([300~620.8 Hz], [620.8~1248.5 Hz], [1248.5~2658 Hz], [2658~5500 Hz])으로 나누어 HNR을 구하는 실험을 진행하였다. <그림 11~14>에 로그스케일에 따라 주파수 대역을 나눈 피험자 MSW의 음성 파형과 스펙트로그램과 스펙트럼을 제시하였다.



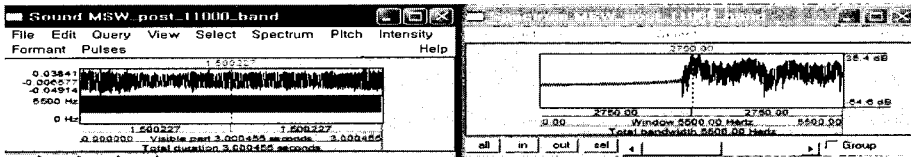
<그림 11> 300~620.8 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



<그림 12> 620.8~1,248.5 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



<그림 13> 1,248.5~2,658 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼



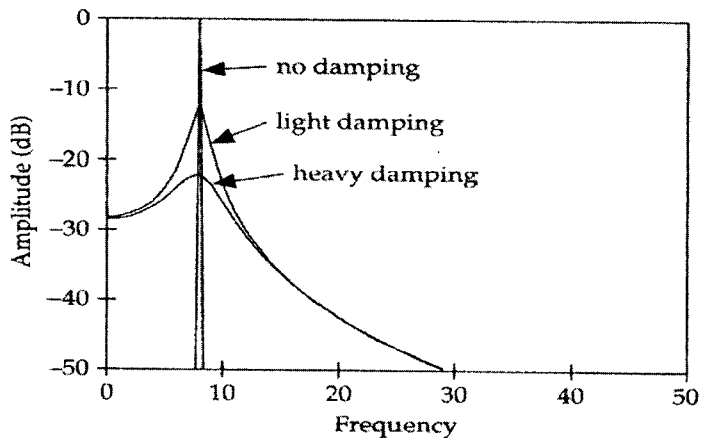
<그림 14> 2,658~5,500 Hz 대역 음성 파형, 스펙트로그램, 스펙트럼

### 3.3.3. F0와 포먼트의 주파수 대역폭

Praat에서는 자기 상관법으로 pitch를 추출하였으며 본 논문에서는 그 추출된 pitch 값을 F0 값으로 삼았다.

### 3.3.4. 포먼트의 주파수 대역폭

Praat를 이용하여 F1과 F2의 공명 주파수 대역폭을 살펴보았다. 감폭률이 높을 수록 넓은 대역폭으로 나타난다(<그림 15> 참고).



<그림 15> 감폭에 따른 주파수 대역폭[12]



## 4. 결과 및 논의

### 4.1. HNR

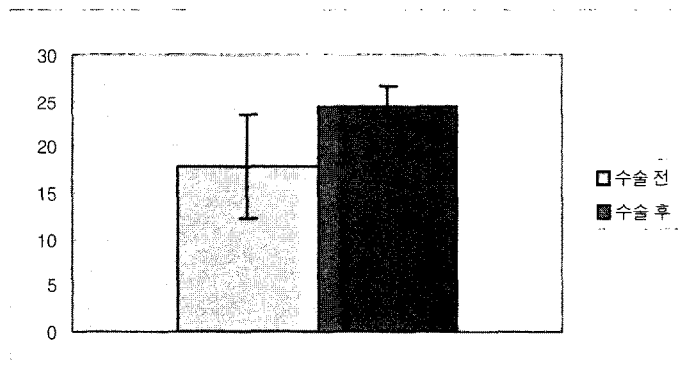
HNR을 측정한 결과 수술 전 평균은 17.81 dB(s.d. 5.63 dB)이었고 수술 후에는 24.42 dB(s.d. 2.1 dB)로 6.61 dB가 증가하였으며 증가율은 137.11%이었다. 대응표본 t-검정(paired t-test)을 한 결과 두 평균값은 통계적으로 유의미한 차이를 나타냈다[t(9) = 3.541\*\*, p < 0.01].

조화음의 비율이 증가하였다는 것은 수술로 성대 폴립이 제거되어 성대의 내전(adduction)과 외전(abduction)이 수술 전보다 완벽하게 이루어졌다는 것을 의미한다. 이로 인해 양질의 음원이 성대로부터 나왔고, 결과적으로 조화음의 에너지는 늘고 소음은 줄어들었기 때문에 나온 결과이다.

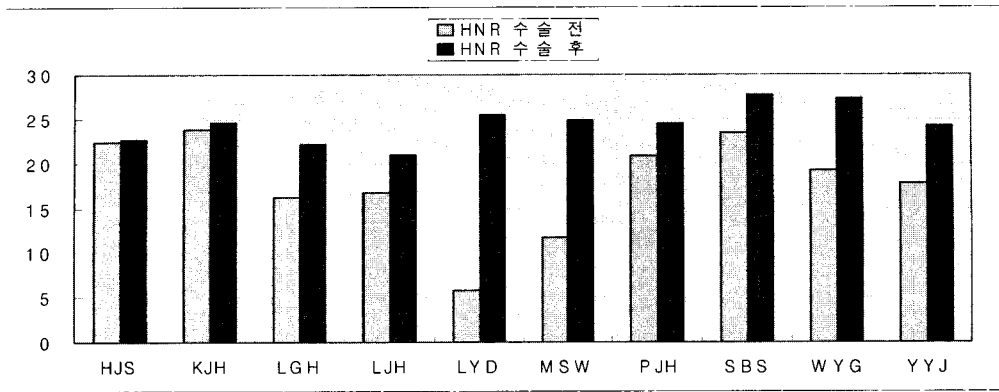
피험자별로 살펴보면 LYD가 가장 많이 증가하였고(19.75 dB), HJS가 가장 적게(0.22 dB) 증가하였으며, 수술 후 HNR 수치가 감소한 피험자는 없었다. 아래 <표 3>과 <그림 16>에 수술 전후 HNR 값을 제시하였고 <그림 17>에 피험자별 수술 전후 HNR 값을 제시하였다.

<표 3> 수술 전후 HNR 평균값과 통계값(단위 dB)

|        | 수술 전   | 수술 후   |
|--------|--------|--------|
| 평균     | 17.81  | 24.42  |
| (표준편차) | (5.63) | (2.10) |



<그림 16> 수술 전후 HNR 평균값 (단위 dB)



<그림 17> 피험자별 수술 전후 HNR (단위 dB)

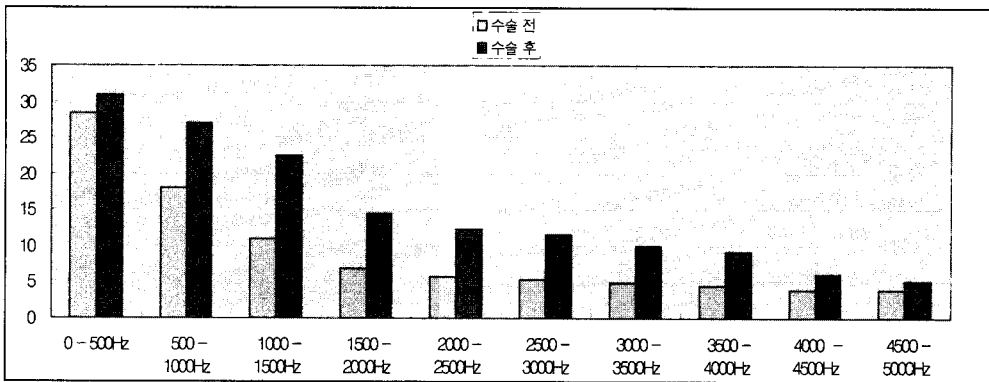
## 4.2. 주파수 대역별 HNR

### 4.2.1. 500Hz 대역폭으로 나눈 HNR

주파수 대역별 HNR의 평균값, 표준 편차, 통계값을 <표 4>에 나타내었다. 그리고 주파수 대역별 HNR의 수술 전후 평균값의 그림을 <그림 18>에 나타내었다.

<표 4> 수술 전후 주파수 대역별 HNR 평균값(단위 dB), 통계값

| 주파수 대역         | 수술 전<br>평균<br>(표준 편차) | 수술 후<br>평균<br>(표준 편차) | t 값        | 유의 확률     |
|----------------|-----------------------|-----------------------|------------|-----------|
| 0~500 Hz       | 28.33 (6.64)          | 30.83 (4.05)          | -1.054     | p = 0.319 |
| 500~1,000 Hz   | 17.96 (5.58)          | 26.96 (2.11)          | -4.109**   | p < 0.01  |
| 1,000~1,500 Hz | 10.93 (4.05)          | 22.54 (2.82)          | -6.219***  | p < 0.001 |
| 1,500~2,000 Hz | 6.91 (3.11)           | 14.45 (4.23)          | -4.437**   | p < 0.01  |
| 2,000~2,500 Hz | 5.71 (2.93)           | 12.40 (1.83)          | -5.171**   | p < 0.01  |
| 2,500~3,000 Hz | 5.18 (1.68)           | 11.66 (1.68)          | -10.931*** | p < 0.001 |
| 3,000~3,500 Hz | 4.79 (1.97)           | 9.96 (1.94)           | -5.653***  | p < 0.001 |
| 3,500~4,000 Hz | 4.48 (1.50)           | 9.10 (2.03)           | -6.646***  | p < 0.001 |
| 4,000~4,500 Hz | 4.00 (1.34)           | 6.13 (2.12)           | -2.472*    | p < 0.05  |
| 4,500~5,000 Hz | 3.91 (1.23)           | 5.17 (1.45)           | -1.650     | p = 0.133 |



<그림 18> 주파수 대역별 수술 전후 HNR(단위 dB)

표와 그림을 통해 제시된 결과를 살펴보면 다음과 같다. HNR 값은 전체 0~5,000 Hz 대역 중 500~4,500 Hz 대역에서만 수술 후 통계적으로 유의미한 차이를 나타냈다. 0~500 Hz 대역과 4,500~5,000 Hz 대역에서는 수술 후 값이 증가하였으나 통계적으로 유의미하지는 않았다.

이와 같은 결과가 나온 것은 다음과 같은 이유에서이다. 성대에 난 폴립은 성대의 내전 시 성문의 완전한 차단을 어렵게 만든다. 이런 상태는 기식음 내지 기식 섞인 목소리를 발생시킨다. 기식음 내지 기식 섞인 목소리의 가장 두드러진 특징은 스펙트럼의 기울기가 현저히 급격해지고 에너지가 높은 소음이 F3, 심지어는 F2 영역에까지 발생한다는 것이다. 이는 1,000 Hz에서 4,000 Hz까지의 영역에서 소음의 발생 비율이 높아지는 것을 의미하며 HNR이 낮아진다는 것을 의미한다 [17]. 따라서 1,000 Hz에서 4,000 Hz 대역이 수술 전후 HNR 값을 비교하기에 가장 적당한 주파수 대역이라 여겨진다.

또한 다음과 같은 설명도 가능하다. 조화음이 지배적인 500 Hz 이하 저주파수 대역에서 조화음과 소음의 비율을 구할 경우 HNR 값은 음질에 상관없이 일관되게 커지기 마련이다. 다시 말해 병리적 음성도 HNR 값이 커지고, 정상 음성의 경우도 마찬가지이다. 따라서 병리적 음성과 정상 음성을 구별할 변별력이 사라진다. 이것은 4,500 Hz 이상 고주파수 대역에서도 마찬가지이다. 즉 병리적 음성이나 정상 음성 모두 소음의 수치가 지나치게 높아 HNR 값이 두 음성 모두에서 작게 나오게 되고, 따라서 둘의 음성을 구별할 만한 변별력은 사라지게 된다. 결국 이 두 대역을 포함하여 조화음과 소음의 비율을 계산한다면 변별력 있고 정밀한 HNR을 구하기가 어렵게 된다.

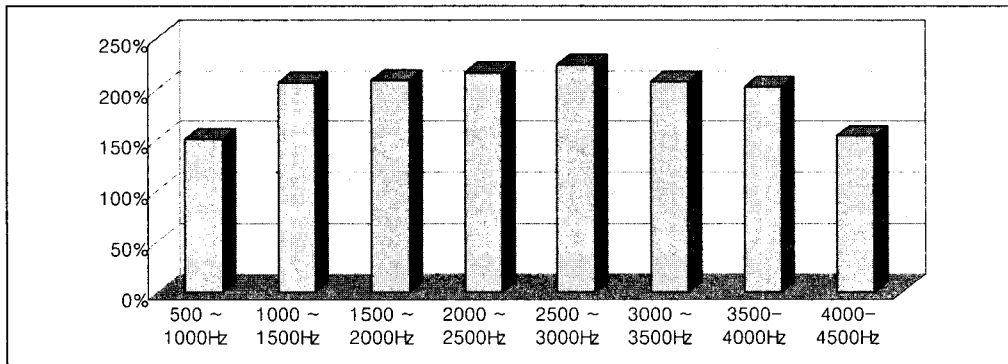
따라서 조화음을 많이 포함하고 있는 저주파수 대역과 소음을 많이 포함하고 있는 고주파수 대역은 정상 음성과 병리적 음성의 차이를 보기에 적절치 않은 대역이다. 결국 먼밀한 HNR 값을 구하기 위해서는 이 두 대역을 제외하고, 조화음과 소음의 수치가 음질의 상태에 따라 민감하게 변화하는 주파수 대역에서 값을

구해야 한다.

그렇다면 수술 전후 차이가 통계적으로 유의미한 500 ~ 4,500 Hz 중 수술 후 HNR 값의 증가율이 큰 대역은 어느 대역일까. 이를 알아보기 위해 수술 후 HNR 증가율을 <표 5>와 <그림 19>에 제시하였다.

<표 5> 수술 후 HNR 증가율 (단위 %)

| 주파수 대역         | 증가율      |
|----------------|----------|
| 500~1,000 Hz   | 150.13 % |
| 1,000~1,500 Hz | 206.12 % |
| 1,500~2,000 Hz | 209.01 % |
| 2,000~2,500 Hz | 217.01 % |
| 2,500~3,000 Hz | 224.14 % |
| 3,000~3,500 Hz | 207.78 % |
| 3,500~4,000 Hz | 203.08 % |
| 4,000~4,500 Hz | 153.16 % |



<그림 19> 수술 후 HNR 증가율 (단위 %)

<표 5>를 보면 1,000~4,000 Hz는 수술 후 HNR 증가율이 200%를 넘는 것에 반해, 500~1,000 Hz 대역과 4,000~4,500 Hz 대역은 수술 후 증가율이 200 %를 넘지 않는다(500~1,000 Hz 증가율: 150.13 %, 4,000~4,500 Hz 증가율: 153.16 %). 또한 이 두 대역은 전체 주파수 대역의 평균값 증가율 137.11 %와 비교하였을 때에도 큰 차이를 보이지 않으며, 두 경우 모두 200 %를 넘지 않는 공통점이 발견된다.

결론적으로 본 논문에서는 다음과 같은 과정을 통해 HNR을 구할 대역을 선택하려 한다. 첫째, 통계적으로 유의미하지 않은 대역은 제외한다. 따라서 0~500 Hz와 4,500~5,000 Hz를 제외한다. 둘째, 통계적으로 유의미한 대역 중 수술 후 증가

율이 200 %를 넘지 않는 대역은 제외한다. 따라서 500~1,000 Hz와 4,000~4,500 Hz를 제외한다. 정리하면 HNR 값을 구하는 가장 정밀하고 변별력 있는 주파수 대역은 1,000~4,000 Hz 대역이라고 결론 내린다.

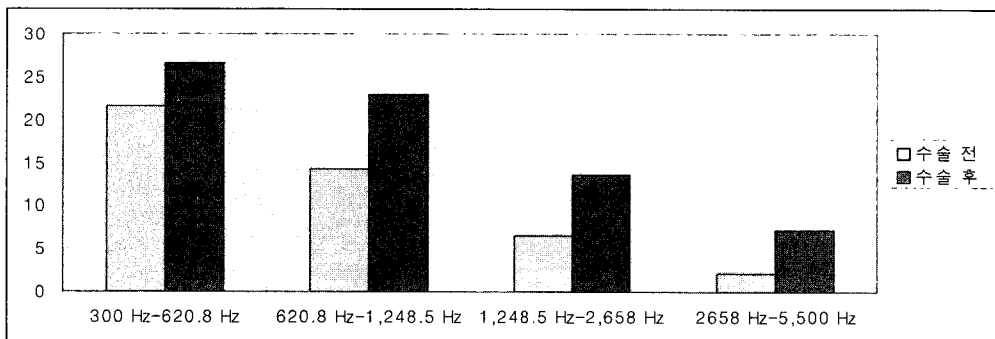
MDVP에서 NHR은 70~4,500 Hz 대역에서 값을 구하였다. 이 경우 1,000 Hz 이하와 4,000Hz 이상 대역에서 NHR 값을 계산하였기에 정밀한 결과를 내지 못한 것으로 판단된다. 이것이 기존 연구에서 NHR이 통계적으로 유의미한 차이를 보이지 못했던 이유라고 여겨진다.

4.2.4. 로그스케일에 따른 HNR

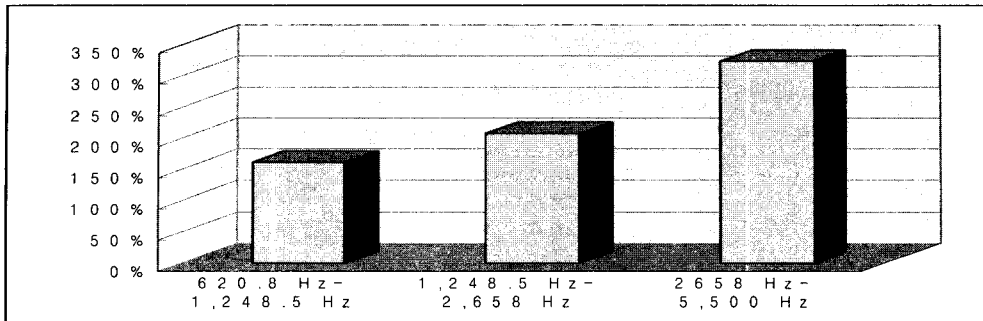
로그스케일에 따른 주파수 대역별 HNR의 평균값, 표준편차, 통계값을 <표 6>에 나타내었고 평균값의 그림을 <그림 20>에 나타내었다. 그리고 수술 전후 차이를 통계적으로 유의미한 대역에서의 증가율을 <그림 21>에 나타내었다.

<표 6> 로그스케일에 따라 나는 주파수 대역별 HNR의 평균값, 증가율, 통계값

| 주파수 대역           | 수술 전<br>평균<br>(표준 편차) | 수술 후<br>평균<br>(표준 편차) | 증가율      | t 값     | 유의 확률    |
|------------------|-----------------------|-----------------------|----------|---------|----------|
| 300~620.8 Hz     | 21.61(5.89)           | 26.54(2.84)           | 122.81 % | -2.261  | p = 0.05 |
| 620.8~1,248.5 Hz | 14.26(5.68)           | 23.00(2.27)           | 161.29 % | -3.976* | p < 0.05 |
| 1,248.5~2,658 Hz | 6.58(3.95)            | 13.62(3.04)           | 206.99 % | -4.173* | p < 0.05 |
| 2,658~5,500 Hz   | 2.21(1.46)            | 7.13(2.42)            | 322.62 % | -5.093* | p < 0.05 |



<그림 20> 로그스케일에 따라 나는 수술 전후 HNR (단위 dB)



<그림 21> 로그스케일에 따라 나눈 HNR의 수술 후 증가율 (단위 %)

표와 그림을 통해 알아본 결과에 의하면 300~620.8 Hz 대역에서는 수술 전후 차이가 통계적으로 유의미하지 않은 것으로 나타났고, 이 대역을 제외한 나머지 대역에서는 통계적으로 유의미한 결과를 보였다. 통계적으로 유의미한 대역 중 620.8~1,248.5 Hz에서의 수술 후 증가율 161.29 %는 전체 대역에서의 증가율 137.11 %와 큰 차이가 나지 않을 뿐더러 수술 후 증가율이 200 %를 넘지 않으므로 HNR을 구하기에 적절치 않은 대역이라 여겨진다.

결론적으로 로그스케일을 활용할 때 HNR을 구하기에 가장 적절한 대역은 1,248.5 Hz~5,500 Hz라 여겨진다.

로그스케일에 따라 나눈 결과와 500 Hz 대역폭으로 나눈 결과를 비교하면, 약 1,000 Hz 이하의 대역은 변별력 있는 HNR 값을 구하기에 적절치 않은 대역이라는 공통점이 발견된다. 차이점으로는 4,000 Hz 이하까지 선택한 500 Hz 대역폭으로 나눈 실험에 비해 로그스케일에 따라 나눈 실험은 4,000~5,500Hz 대역을 포함한다는 것이다.

이러한 차이점이 있기는 하지만 두 실험 모두 저주파수 대역과 고주파수 대역을 제외한 나머지 대역, 즉 소음과 조화음이 음질의 상태에 따라 민감하게 변화하는 중간 주파수 대역에서 HNR 값을 구한다는 결론에 두 실험 결과 모두가 들어 맞는다는 것을 알 수 있다. 결국 실험 방법에 따라 주파수 대역이 조금씩 차이가 날 수는 있겠으나 조화음이 지배적인 저주파수 대역과 소음이 지배적인 고주파수 대역을 제외한 중간 대역에서 HNR을 구해야 한다는 일반적인 결론을 두 실험 결과에서 모두 얻을 수 있다.

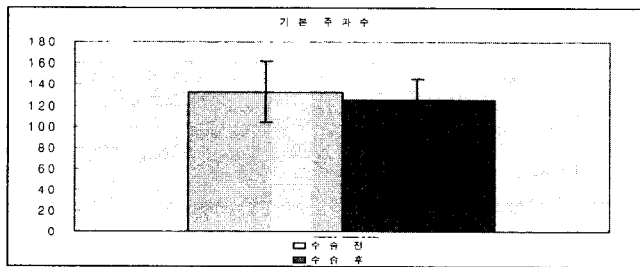
#### 4.3. F0

수술 전후 F0의 평균값을 <표 7>에 제시하였다. <그림 17>은 평균값을 그림으로 나타낸 것이고 <그림 18>은 피험자별 값을 그림으로 나타낸 것이다. 결과적으로 F0는 수술 후에 14.46 Hz 감소하였으나 통계적으로 유의미하지 않았다( $t(9) =$

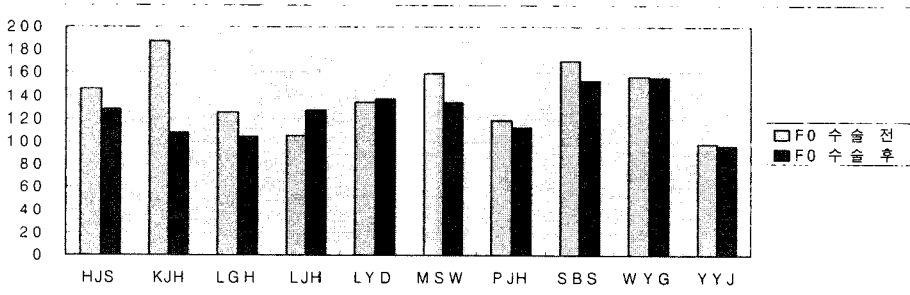
1.721,  $p = 0.119$ ]. 다만 수술 후 F0가 성인 남자의 평균값인 125 Hz대로 내려갔다는 것은 성대 진동이 안정되고 있음을 보여주는 결과라 할 수 있다. 하지만 본 결과에서 두 명의 피험자는 전체 평균과는 반대로 증가하여서 본 결과를 일반화하기엔 무리가 따른다. 다만 수술 후 기간의 경과에 따라 추적 조사한다면 목소리의 안정화와 함께 F0가 어떻게 변화되는지에 대한 좀 더 상세한 연구가 가능할 것이라 여겨진다.

<표 7> 수술 전후 F0 평균값과 표준편차(단위 Hz)

|       | 수술 전   | 수술 후   |
|-------|--------|--------|
| 평균    | 132.85 | 125.62 |
| 표준 편차 | 28.88  | 20.19  |



<그림 22> 수술 전후 F0 평균값(단위 Hz)



<그림 23> 피험자별 수술 전후 F0 값(단위 Hz)

#### 4.4. 포먼트의 주파수 대역폭

<표 8>과 <그림 24>은 수술 전후 F1의 주파수 대역폭의 평균값을 나타낸 것이고, <표 9>과 <그림 25>는 수술 전후 F2의 주파수 대역폭의 평균값을 나타낸 것이다. <그림 26>과 <그림 27>은 피험자별 주파수 대역폭을 나타낸 것이다.

F1의 주파수 대역폭(B1)과 F2의 주파수 대역폭(B2)을 구해본 결과, B1은 수술

전후 통계적으로 유의미한 차이가 없었고[t(9) = 2.135, p = 0.062], B2는 수술 후 약 100 Hz가 감소하였는데 이는 통계적으로 유의미한 차이가 있었다[t(9) = 2.379\*, p < 0.05].

성대에 난 폴립은 성문의 완전한 차단을 방해하여 성문 아래와 성문 위 공간의 음향적 결합(acoustic coupling)을 초래한다. 이것은 공명 특성의 변화를 가져오는데 그 중 가장 두드러진 변화가 B1의 변화이다. B1은 성문의 넓이와 직접적인 관계가 있다. 수술 후에 포먼트 대역폭이 감소한다는 것은 폴립으로 인한 성문의 불완전한 개폐가 사라져 정상적인 개폐로 바뀌었다는 것을 보여준다[18].

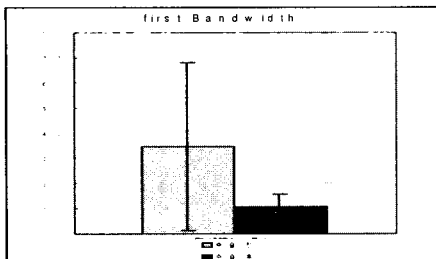
또한 본 결과는 성대로부터 양질의 에너지원이 제공됨에 따라 조화음의 주파수 대역폭이 좁아졌고, 결국 조화음으로 구성된 포먼트의 주파수 대역폭도 좁아졌기 때문에 나온 결과라고 판단된다. 다만 이 결과는 피험자 간 수술 전후 변화율의 차이가 크고 KJH와 LJH는 오히려 수술 후에 B2의 값이 늘어났으므로 일반화하기엔 무리가 따른다고 보인다. 따라서 공명 개선의 여부에 대해서는 더 많은 피험자를 확보하여 분석할 필요가 있다고 생각된다.

<표 8> 수술 전후 F1의 주파수 대역폭 값(단위 Hz)

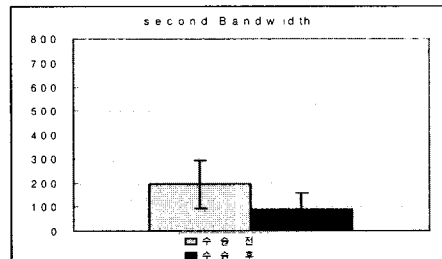
|       | B1     |        |
|-------|--------|--------|
|       | 수술 전   | 수술 후   |
| 평균    | 346.93 | 108.07 |
| 표준 편차 | 334.10 | 49.43  |

<표 9> 수술 전후 F2의 주파수 대역폭 값(단위 Hz)

|       | B2     |       |
|-------|--------|-------|
|       | 수술 전   | 수술 후  |
| 평균    | 193.93 | 90.81 |
| 표준 편차 | 99.33  | 64.34 |

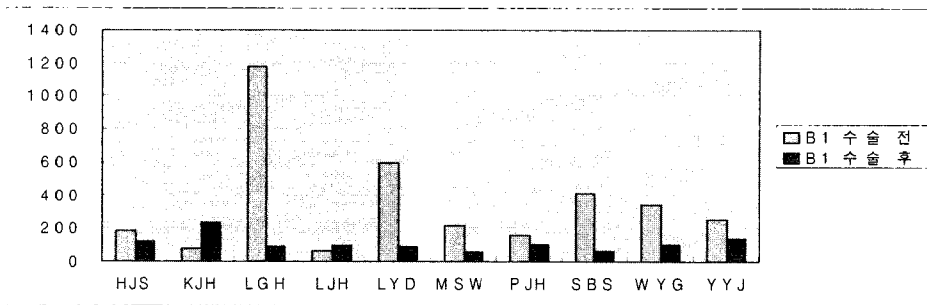


<그림 24> 수술 전후 F1의 주파수 대역폭 평균값(단위 Hz)

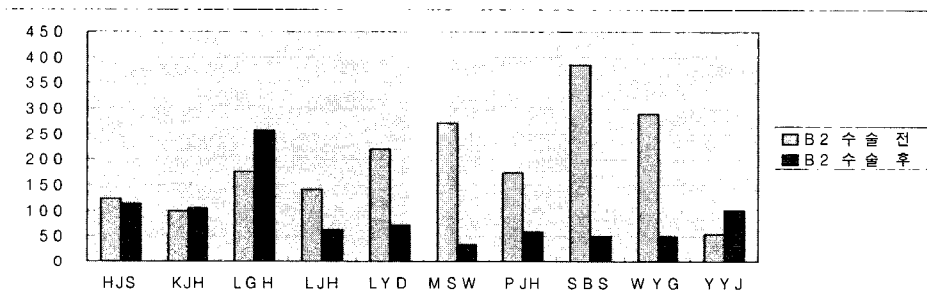


<그림 25> 수술 전후 F2의 주파수 대역폭 평균값(단위 Hz)





<그림 26> 피험자별 수술 전후 F1의 주파수 대역폭 값(단위 Hz)



<그림 27> 피험자별 수술 전후 F2의 주파수 대역폭 값(단위 Hz)

### 5. 맺음말

지금까지 후두미세수술 후 모음 /아/의 음질 향상의 정도를 주파수 대역별 HNR을 포함한 4개의 음향학적 매개변수를 통해 살펴보았다. 그 결과와 논의 내용을 요약하여 기술하면 다음과 같다.

첫째, HNR 값은 수술 후가 수술 전보다 통계적으로 유의미하게 높았다. 이것은 성대로부터 양질의 음원이 제공되어 조화음의 에너지가 증가하여 나온 결과이며 이는 음성의 음질이 향상되었음을 의미하는 것이다.

둘째, 500Hz 대역별로 나누어 HNR을 살펴본 결과, 변별력 있고 정밀한 HNR 값을 도출하기에 가장 적합한 대역은 1,000~4,000 Hz 대역이라는 결론을 얻었다. 이것은 HNR을 측정할 때 조화음이 지배적인 1,000 Hz 이하의 저주파수 대역과 소음이 지배적인 4,000 Hz 이상의 고주파수 대역을 제외한 나머지 대역에서 HNR 값을 구해야 보다 정밀하고 변별력 있는 HNR 값을 도출할 수 있다는 것을 의미한다. 기존에 MDVP를 사용하여 HNR 값을 구한 많은 연구들에서 HNR 값이 변별력 있고 구체적인 결과를 제시하지 못한 것은 1,000 Hz 이하의 조화음을 포함하여 식을 계산하였기 때문이라고 판단된다.

셋째, 로그스케일에 따라 주파수 대역을 나누어 HNR 값을 구해본 결과, 1,248.5~5,500 Hz 대역이 HNR을 구하기에 가장 적합한 대역이라는 결과가 나왔다. 이것은 HNR을 측정할 때 조화음이 지배적인 1,248.5 Hz 이하 대역과, 소음이 지배적인 5,500 Hz 이상의 대역을 제외한 나머지 대역에서 HNR 값을 구해야 변별력 있는 HNR 값을 구할 수 있다는 것을 의미한다.

넷째, HNR을 구할 가장 적합한 대역을 구하기 위한 두 실험, 즉 500 Hz 대역 별로 나누는 실험과 로그스케일에 따라 대역을 나누는 실험의 결과를 비교하면, 조화음이 지배적인 저주파수 대역과 소음이 지배적인 고주파수 대역을 제외한 나머지 대역에서 HNR 값을 구해야 한다는 공통점을 얻을 수 있다. 다만 조화음이 지배적인 저주파수 대역은 약 1,000 Hz 이하로서 두 실험에서 모두 비슷한 결과를 내었지만, 소음이 지배적인 고주파수 대역은 각각 4,000 Hz와 5,000 Hz 이상으로 두 실험에서 차이를 보인다. 그러나 이는 실험 설계에서 충분히 발생할 수 있는 차이로 여겨지며, 두 실험 모두에서 조화음과 소음이 민감하게 변화하는 중간 대역에서 HNR 값을 구해야 한다는 공통된 결론을 이끌어 낼 수 있다.

다섯째, F0는 수술 후에 낮아졌지만 통계적으로 유의미하지 않았다. 다만 수술 후 F0가 성인 남성의 평균 주파수인 125 Hz대로 낮아졌다는 것은 수술 후 음성이 안정화되었음을 의미한다.

여섯째, F2의 주파수 대역폭은 수술 후에 유의미하게 감소하였고, F1의 주파수 대역폭은 수술 후에 감소하였으나 통계적으로 유의미하지 않았다. 이는 수술 후에 F2 부근의 공명이 수술 전보다 효과적으로 이루어졌다는 것을 의미하는 것이다. 즉 성대로부터 양질의 음원이 제공됨에 따라 조화음들의 주파수 대역폭이 좁아졌고 이로 인해 조화음들로 구성된 포먼트 역시 주파수 대역폭이 좁아져서 나온 결과라고 판단한다.

본 논문의 의의는 다음 두 가지로 정리할 수 있다.

첫째, 본 논문은 지금까지 후두미세수술 전후 음향학적 평가에서 많이 쓰인 jitter, shimmer, HNR 혹은 NHR 등의 매개변수 이외 주파수 대역별 HNR과 포먼트의 주파수 대역폭과 같은 매개변수를 통해 수술 전후 통계적으로 유의미한 음향학적 차이를 규명함으로써 기존 연구들과 차이점을 두었다.

둘째, 본 논문에서는 변별력 있는 HNR 값을 얻기에 최적의 주파수 대역(1,000~4,000 Hz 대역)을 제시하여 기존 연구에서 NHR이나 HNR이 정밀한 결과 값을 유도해내지 못 한 이유를 설명하였다.

본 논문에서 사용한 음향학적 매개변수들과 청지각적 매개변수 그리고 공기역학적 매개변수간의 상관관계를 밝히는 것은 본 논문의 매개변수들의 신뢰도를 높이는 매우 중요한 작업이므로 후행 연구가 필요하다고 여겨진다. 또한 본 논문에서 사용한 매개변수들 간에도 상관관계를 조사하여 어떠한 매개변수들끼리 서로 관련이 깊은지 살펴보는 연구도 진행할 필요가 있다. 마지막으로 남성의 음성뿐만

아니라 여성의 음성도 실험 자료에 포함하여 병리적 음성에서 성별에 따른 차이가 있는지 밝혀내는 후행 연구도 필요하므로 진행할 필요가 있다고 생각된다.

### 참고문헌

- [1] 고도홍, 정옥란, 신효근, 최홍식, 김현기, 왕수건, 이정학, 양병곤, 김진숙, 김연희, 배소영, 박병규, 신지영, 표화영, 안종복, 박상희, 배재연, 정용호, *음성 및 언어 분석기기 활용법*, 서울: 한국 문화사, 2001.
- [2] 대한이비인후과학회, *이비인후과학*, 서울: 일조각, 2005.
- [3] D. R. Boone, S. C. MacFarlane, *The Voice and Voice Therapy*, 6th ed., Boston: Allyn and Bacon, 2005.
- [4] 남순열, 박정혁, 전현경, 김상윤, “후두미세 수술 전후 성대 양성 병변 환자의 음성 분석”, *대한이비인후과학회지*, 41권, pp. 925-928, 1998.
- [5] 왕수건, 백무진, 양병곤, 조철우, 박현민, 권순복, 김태섭, 홍현석, “병적음성의 감별진단을 위한 음향학적 변수”, *대한이비인후과학회지*, 42권, pp. 1561-1567, 1999.
- [6] 김영모, 조정일, 김철호, 김영진, 하현령, “후두 미세 수술 전후의 음성기능분석”, *대한이비인후과학회지*, 42권, pp. 1174-1178, 1999.
- [7] 박수현, 이승균, 조남순, 이민재, 이병돈, 장혁순, 강주원, “후두 결절 및 후두 폴립의 술 전·후 음향학적 지표의 변화에 관한 연구”, *대한이비인후과학회지*, 42권, pp. 634-638, 1999.
- [8] 권택균, *다채널 음성분석과 다변량 로짓 분석을 통한 포괄적 음성지표의 개발*, 서울대학교 의학과 석사학위 논문, 2003.
- [9] 진성민, 송윤경, 이경철, 권기환, 반재호, 김태형, 김상호, 강일규, “성대 용종 및 결절에 대한 후두미세수술 전후의 음성변화”, *대한이비인후과학회지*, 46권, pp. 147-150, 2003.
- [10] 안종복, 유재연, 권도하, 정옥란, “일반학생과 성악도를 대상으로 Dr. Speech의 음향학적 측정치와 EGG 측정치의 상관관계 비교 연구”, *대한음성언어의학회지*, 13권, 1호, pp. 28-32, 2002.
- [11] G. P. Jotz, O. Cervantes, M. Abrahão, F. A. P. Settami, E. C. Angelis, “Noise-to-Harmonics Ratio as an Acoustic Measure of Voice Disorders in Boys”, *Journal of Voice*, Vol. 16, No. 1, pp. 28-31, 2002.
- [12] K. Johnson, *Acoustic and Auditory Phonetics*, 2nd ed., United Kingdom: Blackwell, 2003.
- [13] K. Shamam, A. Krishna, N. U. Cholayya, “Study of Harmonics-to-Noise Ratio and Critical-Band Energy Spectrum of Speech as Acoustic Indicators of Laryngeal and Voice Pathology”, *EURASIP Journal no Advances in Signal Proceeding*, 2007.
- [14] P. Boersma, “Accurate Short-term Analysis of the Fundamental Frequency and the Harmonics-to-Noise Ratio of a Sampled Sound”, *Proceedings of IFA*, Vol. 17, 1993.
- [15] P. Ladefoged, *Elements of Acoustic Phonetics*, 2nd ed., Chicago: The University of Chicago Press, 1996.

- [16] V. Uloza, V. Saferis, I. Ulziene, "Perceptual and Acoustic Assessment of Voice Pathology and the Efficacy of Endolaryngeal Phonomicrosurgery", *Journal of Voice*, Vol. 19, No. 1, pp. 138-145, 2005.
- [17] Klatt, D. H. and L. C. Klatt (1990) Analysis, Synthesis, and Perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, Vol. 87, No. 2, pp. 820-857.
- [18] Stevens, K. (1998), *Acoustic Phonetics*, Cambridge: The MIT Press.
- [19] 황연신, 후두미세수술 전후 모음 /아/의 발성 및 공명 특성 비교 연구, 충남대학교 석사 학위 논문, 2007.

접수일자: 2008년 7월 25일

게재결정: 2008년 9월 6일

▶ 황연신(Yeon-Sin Hwang)

주소: 서울시 강남구 도곡동 146-92

소속: 연세대학교 의과대학 음성언어의학연구소

전화: 02) 2019-2587

E-mail: sofihwang@naver.com

▶ 성철재(Cheol-Jae Seong): 교신저자

주소: 305-764 대전 유성구 궁동 220

소속: 충남대학교 인문대학 언어학과

전화: 042) 821-6395

E-mail: cjseong@cnu.ac.kr