

# 음성 연령에 대한 음향학적 분석

## - 동음을 중심으로

김지채 · 한지연 · 정옥란  
대구대학교 언어치료학과

acoustic analysis of the aging voice - Baby voice

Jichae Kim, Jiyeon Han, Okran Jeong

Department of Speech Language Pathology, Daegu University School of Electronics

E-mail : veronika0626@daum.net

### Abstract

The purpose of this study is to examine the difference in acoustic features between Young Voices and Aged Voices, which are actually come from the same age group. The 12 female subjects in their thirties were participated and recorded their sustained vowel /a/, connected speech, and reading. Their voices were divided into Younger Voices and Aged Voices, which means voices sound like younger person and sound like in their age or more aged ones. Praat 4.4.22 was used to record and analyze their acoustic features like Fo, SFF, Jitter, Shimmer, HNR, Pitch-range. And the six female listeners guessed the subjects' age and judged whether they sound younger or as like their actual age.

We used the Independent t-Test to find the significant difference between those two groups' acoustic features. The result shows a significant difference in Fo, SFF. The above and the previous studies tell us the group who sounds like younger or baby like voice has the similar acoustic features of actually young people.

### I. 서론

사람들 간의 관계에서 첫인상은 어떤 사람의 특징을 결정짓는데 아주 큰 영향을 미친다. 첫인상이 좋으면

어떤 사람의 행동에 대해 전반적으로 좋게 평가하게 된다는 것으로 이를 초두효과라 한다. 이런 첫인상에 영향을 미치는 요인으로는 얼굴의 모양이나 눈의 형태 등이 있으며 그 중에는 음성도 포함된다. 우리는 처음으로 사람을 만나서 인사를 나눌 때 동시에 그 사람의 음성에 대해서 지각적으로 분석을 하게 된다. 듣기에 좋은 음성인지, 나쁜 음성인지부터 성량이 풍부한지, 음도가 높은지, 크기가 적당한지를 지각한다. 또한 음성에서 화자의 나이에 대해서도 판단하게 된다. 최근에는 이와 더불어 실제 나이보다 젊게 보이는 동안(童顏)이 큰 이슈가 되고 있다. 첫인상을 얼굴의 생김새와 더불어 음성으로도 지각한다는 것을 생각해 볼 때, 현재 이슈가 되고 있는 동안의 특징 역시 단순히 얼굴의 형태만을 보고 판단하는 것이 아니라, 음성의 특징도 판단하는 기준에 포함된다고 보는 것이 옳을 것이다.

Shipp과 Hollien(1969)는 청자는 화자를 보거나 화자에 대한 사전 지식이 없이도 음성을 듣고 생활연령을 비교적 정확하게 지각할 수 있다고 하였다. Pracek과 Sanders가 화자의 음성 연령에 대한 연구를 시작한 이래로 화자의 음성 연령에 대한 지각적 판단을 연구한 논문이 다수 발표되었다. [1], [2], [3]

음성 연령을 지각적으로 판단하는 것에 대한 연구는 논문은 청자가 얼마나 정확하게 화자의 음성 연령을 지각하는지에 대한 연구와 여러 연령 집단으로 지각된 음성의 음향학적 특징을 분석하는 것으로 나눌 수 있다. Huntley 등과 Jacques 등은 젊은 청자가 나이든 청자들 보다 음성을 더욱 정확하게 지각한다고 하였고,

[4].[5] 여성 청자가 남성 청자보다 더 정확하게 음성 연령을 지각하는 경향이 있다고 하였다.[6] 또한 나이든 음성보다 젊은 음성을 더 정확하게 지각한다고 하였다.[7]

화자의 연령을 지각하는 요인으로 중 음향학적 파라미터로는 SFF(speaking fundamental frequency)와 MPFR(maximum phonation frequency), jitter와 shimmer 등의 음질 관련 요인 등이 있다.[8] Debruyne와 Decoster(1999)는 20대와 노인의 정상 성인의 음성을 청자에게 들려준 후 젊은 음성으로 나누어 평가하도록 하고 두 집단의 음성에 대해 음향학적 분석을 실시하였다. 그 결과 젊은 음성으로 지각된 20대와 나이든 음성으로 지각된 노인의 음성은 SFF와 HNR, H1-H2, F1와 F2의 위치에서 통계적으로 유의한 차이가 있다고 하였다. [3]

이에 따라 본 연구자는 청자에게 동일 연령 집단인 화자의 연령에 대한 지각적 평가를 실시한 후, 청자가 실제 나이보다 젊게 지각한 집단과 실제 나이로 지각하거나 더 나이가 들었다고 지각한 집단으로 나누어 두 집단이 차이가 있는지를 음향학적 분석으로 알아보고자 한다.

## II. 연구 방법

### 1. 연구대상

#### (1) 화자

이 연구에 참여한 화자는 모두 대구 및 경상도 지역에 거주하고 있는 30대 여성으로 총 12명이 참여하였다. 모두 음성장애의 징후가 없는 정상 성인으로, 대학 졸업 이상의 학력을 소지하고 있으며, 연구자가 평가하기에 심하지 않는 경상도 방언을 구사하는 화자로 하였다. 화자의 평균 연령은 34:07세(표준편차 2:04)로 32세부터 38세 사이의 여성으로 하였다.

#### (2) 청자

이 연구에 참여한 청자는 모두 대구 및 경상도 지역에 거주하고 있는 10대 후반의 여성, 총 6명이 참여하였다. 청자의 평균 연령은 20:4세(표준편차 0:07)로 모두 청각장애 및 언어장애가 없는 정상 성인으로 대구대학교 언어치료학과 1, 2학년에 재학 중인 청능 훈련을 받지 않는 여성을 청자로 선정하였다.

### 2. 연구절차

#### (1) 화자의 음성 샘플 수집

화자의 음성을 녹음하기 위하여 음성 녹음 및 분석 도구인 praat(version 4.4.22)을 사용하여 녹음하였다.

샘플링 모드는 44,100Hz, mono sound로 녹음하였고, 편안한 상태에서 마이크와의 거리를 15cm 두고 발성하게 하여 녹음하였다.

수집한 샘플의 형식은 연장발성과 읽기, 문장발화를 사용하였다. 수집된 샘플의 형식은 다음과 같다.[8]

표 1. 음성 샘플의 종류

| 샘플 형식        | 내용  |
|--------------|---|
| 연장발성         | 편안한 상태에서 /a/를 연장 발성한다.                          |
| 읽기 :<br>산책문단 | “높은 산에 올라가 맑은 공기를 마시며 소리를 지르면 가슴이 활짝 열리는 듯 하다.” |
| 문장발화         | “안녕하십니까? 제 이름은 ○○○입니다.”<br>“여보세요? 누구십니까?”       |

#### (2) 청자의 지각적 판단

6명의 청자에게 12명의 화자에게서 수집한 음성 샘플을 무작위로 들려주고 화자의 연령을 청각적으로 지각하게 하였다. 지각 판단은 7점 척도(10-19세, 20-29세, 30-39세, 40-49세, 50-59세, 60-69세, 70-79세)를 사용하여 화자의 음성이 어느 연령대로 지각되는지 체크하게 한 후, 자세한 나이를 직접 적게 하였다.

### 3. 샘플 분석 및 통계 처리

동일 연령 집단에 대한 청자의 음성 연령의 지각 정도를 알아보기 위하여 음성 샘플을 동음 집단과 실제 연령 집단으로 나누었다. 연구의 신뢰도를 위하여 청자의 검사자간 신뢰도가 70%이상인 데이터만을 사용하였다.

두 집단이 실제로 어떤 요인에서 차이가 있는지를 알아보기 위하여 연장 발성 /a/에서 F0와 jitter, shimmer, HNR을, 읽기와 문장발화에서 SFF와 pitch range, HNR을 측정하였고 독립표본 t-검정(SPSS 10.0 for windows)을 사용하여 집단간의 차이가 통계적으로 유의한지 알아보았다. 음성 분석도구는 praat(version 4.4.22)를 사용하였고, 각 파라미터의 분석 방법은 양병곤(2003)을 기초로 하였다.[9]

## III. 연구 결과

### 1. 지각 판단에 대한 평균 및 표준편차

5명의 여성청자에게 12명의 여성화자의 음성을 들려주고 연령을 지각적으로 평가하게 한 결과는 아래와 같

다(표 2).

표 4. 음성연령에 대한 지각적 평가 결과

|         | N | 평균연령 (세:개월) |       |       |       |
|---------|---|-------------|-------|-------|-------|
|         |   | 실제 연령       | 표준 편차 | 지각 연령 | 표준 편차 |
| 동음 집단   | 6 | 33:06       | 1:06  | 26:11 | 5:02  |
| 실제지각 집단 | 6 | 35:08       | 2:05  | 34:06 | 7:07  |

동음집단과 실제지각집단이 동일한 집단임을 알기 위해 독립 t-검정을 실시하였다. 그 결과 실제지각집단에서는 지각 연령과 실제 연령의 차이가 통계적으로 유의미하지 않았다. 이는 두 집단이 같은 모집단에서 추출되었음을 의미한다.

## 2. 동음집단과 실제지각집단의 음향학적 분석결과

### (1) 연장발성 /a/에 대한 음향학적 분석결과

동음집단과 실제지각집단의 음성이 음향학적으로 차이가 있는지 알아보기 위하여 연장발성 /a/를 녹음한 뒤 안정된 1초간의 F0와 jitter, shimmer, HNR을 측정하고, 두 집단의 측정값을 독립 t-검정을 사용하여 그 차이가 유의미한지 통계적으로 살펴보았다. 결과는 다음과 같다(표 3).

표 5. 연장발성 /a/에 대한 동음집단과 실제지각집단의 분석 결과

|         |        | 평균(M)   | SD      | t     |
|---------|--------|---------|---------|-------|
| F0(Hz)  | 동음집단   | 227     | 20.17   | 2.74* |
|         | 실제지각집단 | 201     | 11.99   |       |
| jitter  | 동음집단   | 0.00394 | 0.00150 | -.307 |
|         | 실제지각집단 | 0.00423 | 0.00176 |       |
| shimmer | 동음집단   | 0.00219 | 0.00177 | -.48  |
|         | 실제지각집단 | 0.00259 | 0.00128 |       |
| HNR(dB) | 동음집단   | 26.53   | 4.43    | .139  |
|         | 실제지각집단 | 26.14   | 5.22    |       |

\* P<.05

t-검정을 통하여 두 집단의 음향학적 파라미터들 간의 차이를 살펴 본 결과 F0의 차이가 통계적으로 유의미하였다. 그러나 jitter와 shimmer, HNR은 통계적으로 유의미한 차이가 나타나지 않았지만 동음집단이 실제지

각 집단보다 jitter와 shimmer에서 더 낮은 것을 알수 있다. 이러한 결과는 젊은 여성 화자와 나이든 여성 화자의 음향학적 파라미터를 분석했을 때 SFF와 HNR에서 두 집단간에 통계적으로 유의미한 차이가 나며 SD-SFF와 jitter, shimmer에서는 차이가 나지 않는다는 Debruyne의 연구와는 차이가 있다.[3]

### (2) 읽기 및 문장발화에 대한 동음집단과 실제지각집단의 분석결과

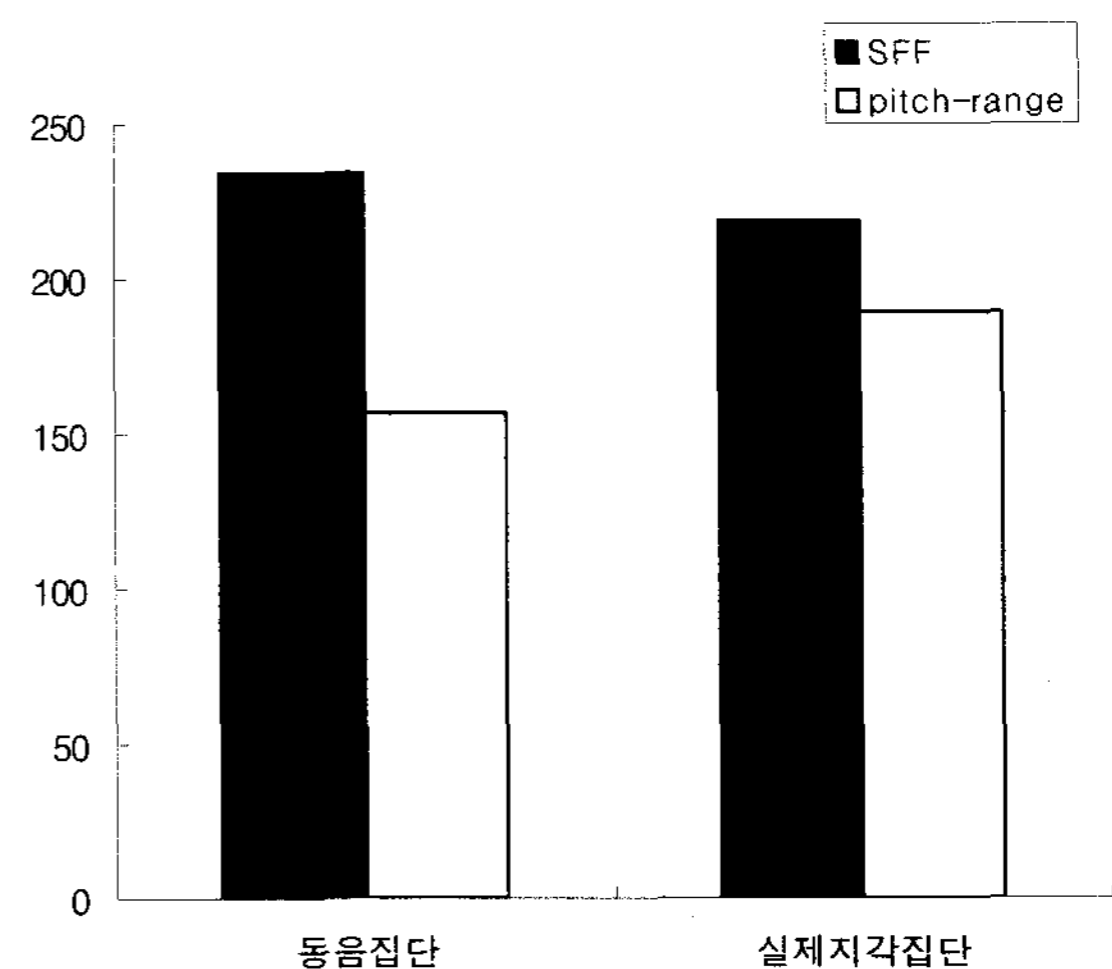
동음집단과 실제지각집단의 음성이 음향학적으로 차이가 있는지 알아보기 위하여 산책문단 낭독과 문장 발화를 녹음한 뒤 SFF와 pitch-range, HNR을 측정하고, 두 집단의 측정값을 독립 t-검정을 사용하여 그 차이가 유의미한지 통계적으로 살펴보았다. 결과는 다음과 같다(표 5).

표 6. 산책문단 읽기 및 문장 발화에 대한 동음집단과 실제지각집단의 분석 결과

|             |        | 평균    | 표준편차   | t     |
|-------------|--------|-------|--------|-------|
| SFF         | 동음집단   | 235   | 16.16  | 2.63* |
|             | 실제지각집단 | 128   | 21.00  |       |
| pitch-range | 동음집단   | 175   | 69.03  | 0.65  |
|             | 실제지각집단 | 185   | 108.91 |       |
| HNR         | 동음집단   | 24.95 | 4.91   | 0.59  |
|             | 실제지각집단 | 24.15 | 4.32   |       |

\* P<.05

그림 1. 산책문단 읽기 및 문장 발화에 대한 동음집단과 실제지각집단의 분석 결과



두 집단의 SFF의 차이를 살펴본 결과 동음집단이 실제지각집단 보다 읽기 및 문장 과업 모두에서 높음을

알 수 있었다. 이를 t-검정을 통하여 두 집단의 음향학적 파라미터들 간의 차이를 살펴 본 결과 SFF의 차이가 통계적으로 유의미하였다. 그러나 pitch-range와 HNR은 통계적으로 유의미한 차이가 나타나지 않았다.

이는 청자가 화자의 연령을 지각 할 때 주로 기본주파수 관련 파라미터를 중심으로 판단한다고 추정할 수 있다.

#### IV. 결론 및 고찰

이 연구는 동일한 연령집단 내에서 청각적으로 실제 나이보다 젊은 나이로 지각되는 집단과 실제 나이로 지각되는 집단을 구별하고 두 집단이 음향학적으로 보이는 특징을 살펴보고자 하였다. 연령이 다른 집단을 지각적으로 구별하고 그 정확도를 알아보거나, 음향학적으로 특징을 분석한 연구는 많이 이루어져 왔으나 같은 연령 집단 내에서 다른 연령으로 지각되는지에 대한 연구는 이루어지지 않았다.[1],[2],[3] 나이가 들더라도 더 젊어 보이는 것을 선호하고 자신의 나이보다 젊고 건강하게 살고자 하는 최근의 사회 현상으로 볼 때 앞으로는 얼굴의 생김새뿐만 아니라 더 젊게 지각되는 음성에 대한 일반인들의 욕구 또한 증가될 것이다. 이에 따라 음성치료사 역시 단순히 음성 장애를 치료하는 것뿐만 아니라 젊게 지각되는 동음 음성에 대한 카운슬링과 치료를 실시한다면 음성 치료의 영역 또한 확대될 것이다.

본 연구 결과에서, 두 집단 간의 음향학적 파라미터들의 측정치의 차이에서 흥미로운 결과가 산출되었다. 동음 집단이 실제지각 집단 보다 F0와 SFF에서 통계적으로 유의미하게 더 높았으며, jitter와 shimmer에서 동음 집단이 실제지각 집단 보다 낮은 것으로 나타났다. 이는 젊은 집단과 나이든 집단을 지각하는데 SFF와 jitter가 중요한 요인이 된다는 Hollen 등의 연구들과 부분적으로 일치한다.[8],[9]

이 연구의 결과와 선행 연구의 고찰을 통하여 본 연구자는 다음과 같은 결론을 내릴 수 있었다. 첫째, 청자들은 화자의 나이를 결정하는데 기본주파수를 주요한 요인으로 삼는 것 같다. 둘째, 동음으로 지각되는 집단은 실제로 젊은 연령 집단과 유사한 형태의 음향학적 특징을 보인다. 즉, 동음 집단의 음성은 음향학적으로 노화되지 않았다는 것을 반증한다. 그러나 본 연구는 연령집단과 대상자의 수가 제한되었고, 제한된 수의 음향학적 파라미터를 분석했으므로 결과를 일반화하기에는 무리가 있다. 또한 동음으로 지각되는 사람들의 음성이 어떠한 이유로 젊은 연령 집단과 유사한 패턴으로 보이는지에 대한 요인도 연구되어 있지 않은 실정이다.

따라서 본 연구자는 성인기부터 노년기까지의 다양한 연령집단을 대상으로 운율이나 파형의 안정성 측정과 같은 다양한 분석기법을 사용하여 좀 더 다차원적인 음향학적 형태를 밝히고 설문지 조사와 인터뷰를 통하여 형태를 가지는 원인에 대해 조사하고, 요인분석을 통해 청자들이 화자들의 음성 연령을 지각하는 가장 큰 요인에 대해 조사할 것을 제안한다.

#### 참고 문헌

- [1] Ptacek, P. H. "Age recognition from voice." *J Speech Hearing Research*, 9, 273-277, 1966.
- [2] Shipp, T., & Hollien, H. "Perception of the aging male voice." *Journal of Speech Hearing Research*, 12, 703-710, 1969.
- [3] Debruyne, F. & Decoster, W. "Acoustic differences between sustained vowels perceived as young or old." *Log Phon Vocol*, 24, 1, 1999.
- [4] Huntley, R., Hollien, H., & Shipp, T. "Influences of listener characteristics on perceived age estimations." *Journal of Voice*, 1, 49-52, 1987.
- [5] Jacques, R. D & Rastatter, M. R. "Recognition of speaker age from selected acoustic features perceived by normal young and old listeners." *Floia Fhoniatr* Vol. 42, pp.118-124, 1990.
- [6] Hartman, D, E. "The perceptual identify and characteristics of aging in normal male adult speakers." *Journal of Communication Disorders*, 12, 53-61, 1979.
- [7] Linville, S, E. "The aging voice." In Kent, R. D. & Ball, M. J. (Ed). *Voice quality measurement* (pp. 359-376), Singular; San Diego, USA, 2000.
- [8] 정옥란. "신경언어장애 검사도구." 미간행, 1993.
- [9] 양병곤. "프라트(Praat)를 이용한 음성분석의 이론과 실제." 만수출판사, 2003.
- [10] Hollien H. & Shipp T. "Speaking fundamental frequency and chronological age in males." *Journal of Speech and Hearing Resarch*, 15, 155-159, 1972.
- [11] Wilcox S. A. & Horii Y. "Age and changes in vocal jitter." *Journal of Gerontology*, 35, 194-198. 1980.