

# 의사결정나무 모형을 이용한 주관적 음성장애 예측모형

변해원<sup>1\*</sup>

<sup>1</sup>남부대학교 언어치료청각학과, 남부대학교 언어치료센터

## The Prediction Model for Self-Reported Voice Problem Using a Decision Tree Model

Haewon, Byeon<sup>1\*</sup>

<sup>1</sup>Dept. of Speech Language Pathology & Audiology, Speech-Language Pathology Center, Nambu University

**요 약** 본 연구에서는 주관적 음성문제의 위험요인으로 구명된 주요 변수를 기반으로 주관적 음성장애를 예측할 수 있는 모형을 개발 하였다. 연구자료는 2008년도 국민건강영양조사이며, 이비인후검진을 완료한 전국의 19세 이상 지역사회 성인 3,600명(남 1,501명, 여 2,099명)을 분석대상으로 하였다. 분석방법은 주관적 음성장애 여부를 결과변수로 성, 연령, 흡연, 음주, 교육수준, 직업, 갑상선장애, 최근 2주간 급성 및 만성질환으로 인한 통증 및 불편감을 설명변수로 사용하였고, 예측모형은 의사결정나무 모형(Decision Tree)의 exhaustive CHAID(Chi Squared Automatic Interaction Detection) 알고리즘을 이용하였다. 주관적 음성 장애와 관련된 통계학적 분류 모형을 구축한 결과, 유의미한 예측 변수는 연령, 교육수준, 최장 직업, 갑상선 장애, 최근 2주 동안의 신체 불편 및 통증경험 여부였다. 이 연구의 모형을 기초로 음성장애 예방을 위해서 음성장애 고위험군에 대한 조기 관리의 필요성이 제기된다.

**Abstract** The purpose of this study was to analyze the risk factors of self-reported voice problem. Data were from the Korea National Health and Nutritional Examination Survey 2008. Subjects were 3,600 persons (1,501 men, 2,099 women) aged 19 years and older. A prediction model was developed by the use of a exhaustive CHAID (Chi Squared Automatic Interaction Detection) algorithm of decision tree model. In the decision tree analysis, pain and discomfort during the last 2 weeks, age, the longest occupation and thyroid disorders was significantly associated with self-reported voice problem. The findings of associated factors suggest potential ways of targeting counseling and prevention efforts to control self-reported voice problem.

**Key Words** : Decision tree, National survey, Prediction model, Risk factor, Self-reported voice problem

### 1. 서론

음성장애(voice disorder)는 질병학적으로 사망과의 관련성이 높은 질환이 아니라는 이유로 보건 정책에서 중요성이 간과되어 왔다. 그러나 음성장애는 지역사회에서 빈번하게 발생될 뿐만 아니라 의사소통과정에서 기능적 문제를 초래하여 일상생활 및 삶의 질에 큰 영향을 미치기 때문에 국민 건강 차원에서 중요한 현안이다.

음성장애는 수술 및 재활치료가 성공적으로 이루어지더라도 재발율이 약 75%로 매우 높기 때문에, 음성장애

의 발병 및 재발을 줄이기 위해서는 예방이 효율적이다 [1]. 지금까지 인구 집단을 대상으로 음성장애의 주요 위험 요인을 규명한 연구들에 따르면 주관적 음성문제가 음성질환의 독립적인 위험 요인으로 보고되었으며[1-3], 관련 요인들 중에서 가장 우선적으로 관여하는 개연성이 있는 예측 요인임이 확인되었다[4]. 그러나 선행 연구들은 연구마다 적용된 혼란변수에 차이가 있고, 위험요인을 예측하기 위한 방법으로 혼란변수를 단계적으로 보정하는 회귀모형을 사용하였기 때문에 독립적인 위험요인을 탐색하는 데 있어서는 효과적이지만 위험요인의 우선순

\*Corresponding Author : DrSc. Haewon, Byeon(Nambu Univ.)

Tel: +82-62-970-0227 email: [byeon@nambu.ac.kr](mailto:byeon@nambu.ac.kr)

Received April 30, 2013

Revised June 4, 2013

Accepted July 11, 2013

위를 고려하기가 어렵다. 또한, 선형 회귀모형은 선형성, 정규성, 등분산성 등의 가정을 필요로 하기 때문에, 일부 질병이나 장애 등의 자료의 경우 회귀모형의 가정이 위배될 가능성도 있다.

최근에는 사회학, 경제학, 보건학, 의학 등의 다양한 영역에서 관심이 있는 대상을 분류하거나 예측하는 방법으로 신경망분석(Neural Network Analysis), 의사결정나무 분석(Decision Tree), 군집분석(Cluster Analysis) 등의 데이터마이닝 분석(data mining)이 이용되고 있다. 이 중에서 의사결정나무 분석은 첫째, 모수적 자료와 비모수적 자료에 상관없이 분석이 가능하고, 둘째, 분석과정이 나무구조에 의해서 표현되기 때문에 신경망분석이나 군집분석에 비해서 음성장애의 속성에 대한 근거를 쉽게 이해할 수 있으며, 셋째, 결과변수에 우선적으로 관여하는 요인을 시각적으로 제시할 수 있기 때문에 임상이나 쉽게 해석해서 사용할 수 있다는 장점이 있다.

본 연구는 전국 조사의 이비인후과 검진자료를 이용하여 주관적 음성장애의 위험요인을 탐색하고, 이를 기반으로 예측 모형을 개발함으로써 음성장애의 예방을 위한 접근 방법을 모색하고자 하였다.

## 2. 연구방법

### 2.1 연구자료

질병관리본부가 2008년에 시행한 제4기 2차년도 국민건강영양조사(Korea National Health and Nutrition Examination Survey)의 자료를 공식적으로 승인을 받고 연구에 이용하였다. 국민건강영양조사는 질병관리본부의 주관 하에 매년 실시하는 전국단위의 국가통계조사로서 표본 추출은 전국의 행정구역과 주거 종류를 층화한 층화집락표본추출방법을 사용하였으며, 제4기 조사에서는 4,600가구 12,528명을 대상으로 하였다[5].

### 2.2 연구대상

건강설문조사와 이비인후과 검사를 모두 완료한 만 19세 이상 지역사회 성인 3,632명 중에서 이비인후과 검진의 음성문제 설문 무응답자 32명을 제외한 총 3,600명(남 1,501명, 여 2,099명)을 분석하였다.

### 2.3 변수의 측정과 처리

결과변수인 주관적 음성장애는 이비인후과 검진 결과를 기초로 이분형(있다, 없다)으로 분류되었다. 설명변수는 선행연구[1-3,8-11]를 참조하여 국민건강영양조사에서

추출 가능한 변수인 연령, 성, 최장직업, 흡연, 음주, 갑상선장애 유병, 최근 2주간 급성 및 만성 질환 등으로 인한 통증 및 불편감 여부로 정의하였다. 연령은 만 나이로 조사되었고, 연속형 변수로 처리하였다. 교육수준은 무학, 초등학교 졸업 이하, 중학교 졸업, 고등학교 졸업, 대학교 졸업 이상으로 분류하였고, 최장 직업은 육체노동자, 비육체노동자, 비경제활동인구로 분류하였다. 흡연 여부는 현재흡연자, 과거흡연자, 비흡연자로 분류하였다. 음주는 평균 음주 빈도를 기초로 주 1회 이하, 주 2~3회, 주 4회 이상으로 분석하였다. 갑상선장애 유병 및 최근 2주간 급성 및 만성 질환 등으로 인한 통증 및 불편감 여부는 각각 이분형(있다, 없다)으로 분류하였다.

### 2.4 주관적 음성장애 관련요인 분석 방법

대상자의 일반적 특성은 조사된 표본이 전체인구를 대표할 수 있도록 건강설문과 검진의 연관성가중치를 부여하였고, 기술 분석을 이용하여 가중치가 적용된 평균 및 표준 오차, 백분율을 각각 제시하였다. 국민건강영양조사의 가중치는 조사에 참여한 개인이 우리나라 전체 인구를 대표하도록 부여되어 있다[5]. 제4기 국민건강영양조사에서는 건강면접조사와 건강행태조사를 통합하여 실시한 건강설문조사 가중치와 검진조사의 가중치가 각각 제시되어 있는데, 이 연구에서는 건강설문과 검진의 자료를 모두 사용하였기 때문에 별도의 건강설문과 검진의 연관성가중치를 사용하였다. 주관적 음성장애 유형에 따른 집단 간의 차이는 연속형 변수의 경우에는 가중치가 적용된 일변량 분산분석(weighted one way ANOVA)으로, 명목형 변수는 라오-스콧 교차분석(Rao-Scott Chi-square test)으로 각각 분석하였다[6]. 이 때, 유의미수준 0.1 이하인 설명변수는 음성장애의 관련 요인으로 규정하고 의사결정나무모형에 포함하였다.

### 2.5 의사결정나무 모형을 이용한 분석방법

주관적 음성문제와 관련된 요인의 통계학적 분류 및 예측 모형 구축은 데이터마이닝 기법 중 하나인 의사결정나무 모형을 이용하였다. 관련요인의 예측은 음성질환 여부를 목표변수로 하여 exhaustive CHAID(Chi Squared Automatic Interaction Detection) 알고리즘을 사용하였다. exhaustive CHAID는 통계적 유의성에 근거하여 더 이상의 분할이 유의하게 필요하지 않으면 노드의 분할을 정지하는 알고리즘으로 분리기준(partitioning criterion)으로 카이제곱 적합성 검정을 사용하며 통계량은 식(1)과 같다 [7].

$$Q_p = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(n_{ij} - m_{ij})^2}{m_{ij}} \quad (1)$$

$m_{ij}$ :  $i$ 번째 그룹의  $j$  번째 결과에 대한 기대도수

$n_{ij}$ :  $i$ 번째 그룹의  $j$  번째 결과에 대한 기대도수

위의 통계량은 자유도 $(i-1)(j-1)$ 를 가지는 카이제곱분포를 따르기 때문에 유의수준을  $\alpha$ 로 할 때  $\chi^2$ 의 값이  $\chi^2_{\alpha}(i-1)(j-1)$ 에 비해서 매우 작다는 것은 예측변수의 각 범주에 따른 목표변수의 분포가 서로 동일하다는 것을 의미한다. 자유도에 대한 카이제곱 통계량 값의 크고 작음은  $p$ 값으로 표현될 수 있는데  $\chi^2$ 의 값이  $\chi^2_{\alpha}(i-1)(j-1)$ 에 비해서 작으면  $p$ 값은 커지게 되므로, 결국  $p$ 값이 가장 작은 예측변수와 최적 분리에 의해서 자식마디(child node)가 형성된다[7].

이 연구의 모형에서 의사결정규칙(decision rule)의 분리(splitting) 및 병합(merging) 기준  $\alpha$ 값은 0.05로 설정하였고, 부모마디(parent node)의 수는 200명, 자식마디 수

는 80명, 분지가지 개수는 5개로 제한하였다. 본 연구에서의 결과변수와 마찬가지로 질병의 유병률이나 불량제품의 선별 등과 같이 발생률이 드문 경우에는 분포의 불균형(unbalanced data)이 초래되어 다수의 결과범주로 분류가 진행되는 분류 알고리즘의 문제가 발생할 수 있다 [8]. 이러한 자료의 분포 불균형 문제를 보완하고자 본 연구에서는 한국인의 음성장애 유병률을 고려하여 오분류 비용(misclassification costs)의 가중치를 비대칭적으로 설정함으로써 자료의 균형을 조정하였다[9]. 모형의 타당성 평가는 10-fold 교차타당성 검증(10-fold cross-validation)을 이용하여 검증하였다.

### 3. 연구결과

#### 3.1 주관적 음성장애 관련요인

주관적 음성문제 여부에 따른 대상자의 일반적 특성 및 주관적 음성장애 관련 요인은 Table 1에 제시하였다.

[Table 1] Characteristics of Korean population according to Self reported voice problem

Variables	Self reported voice problem		P*
	No (n=3,361)	Yes (n=239)	
Age(weighted mean±standard error)	43.9±0.6	48.8±1.2	<0.001
Sex			0.008
Man	1,419(95.2±0.7)	82(4.8±0.7)	
Woman	1,942(92.7±0.7)	157(7.3±0.7)	
Education level			0.010
Elementary school	987(91.1±1.2)	87(8.9±1.2)	
Middle school	371(92.9±1.6)	31(7.1±1.6)	
High school	1,176(95.4±0.8)	62(4.6±0.8)	
≥ College	820(94.5±0.8)	58(5.5±0.8)	
The longest occupation			0.099
Economically inactive	79(87.4±4.2)	11(12.6±4.2)	
Non-manual	1,698(94.2±0.7)	125(5.8±0.7)	
Manual	1,498(94.3±0.8)	93(5.7±0.8)	
Smoking			0.637
Nonsmoker	1,976(93.7±0.7)	141(6.3±0.7)	
Past smoker	615(94.9±1.2)	41(5.1±1.2)	
Current smoker	764(94.3±1.0)	54(5.7±1.0)	
Alcohol drinking			0.872
≤1 time per week	2,618(94.0±0.6)	190(6.0±0.6)	
2-3 times per week	496(94.3±1.1)	34(5.7±1.1)	
>3 times per week	238(94.8±1.8)	12(5.2±1.8)	
Thyroid disorders			<0.001
No	3,250(94.3±0.5)	216(5.7±0.5)	
Yes	109(83.4±4.0)	22(16.6±4.0)	
Pain and discomfort during the last 2 weeks			<0.001
No	2,463(95.5±0.5)	126(4.5±0.5)	
Yes	892(89.1±1.2)	111(10.9±1.2)	

Values in parentheses denote weighted percent and standard error.

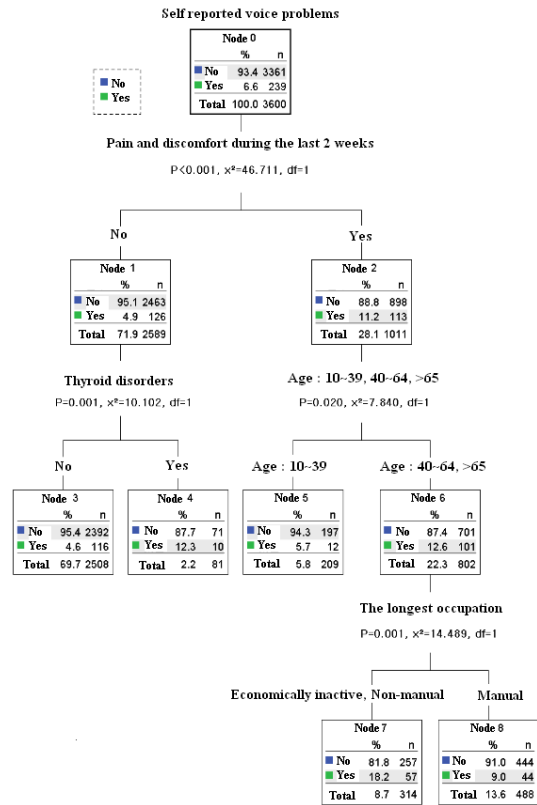
\* Rao-scott chi-squared test for weighted samples

전체 대상자 3,600명 중에서 정상 집단은 3,361명 (93.4%), 주관적 음성장애 집단은 239명(6.6%)이었다. 평균 연령은 정상집단 43.9세(표준오차 0.6), 주관적 음성장애 집단은 48.8세(표준오차 1.2)로서 가중치가 적용된 일변량 분산분석 검정 결과 주관적 음성장애 집단의 연령이 통계적으로 유의미하게 높았다( $p < 0.001$ ). 라오-스콧 교차 분석 결과, 주관적 음성장애 집단과 정상집단은 성, 교육수준, 최장 직업, 갑상선장애 유병, 최근 2주간 급성 및 만성 질환 등으로 인한 통증 및 불편감 여부에서 통계적으로 유의미한 차이가 있었다. 여성은 남성에 비해서 주관적 음성장애 유병률이 더 높았다( $p = 0.008$ ). 또한, 초등학교 졸업( $p = 0.01$ ), 비경제활동( $p = 0.09$ )에서 주관적 음성장애의 유병률이 가장 높았다. 갑상선 장애와 최근 2주 동안의 신체 불편 및 통증을 경험한 집단은 주관적 음성장애 유병률이 약 2.5배 더 높았다( $p < 0.05$ ). 이 연구에서 연령, 성, 교육수준, 최장 직업, 갑상선 장애, 최근 2주 동안의 신체 불편 및 통증경험 여부는 주관적 음성장애와 유의미한 관련이 있는 요인이었기 때문에 의사결정나무 모형의 설명변수에 포함하였다.

### 3.2 주관적 음성장애 예측모형

주관적 음성문제의 관련요인으로 설정된 변수들을 의사결정나무 모형에 포함한 후 exhaustive CHAID 알고리즘을 이용하여 통계학적 분류모형을 구축한 결과, 유의한 영향을 미치는 변수는 최근 2주간 급성 및 만성질환 등으로 인한 통증 및 불편감 여부, 갑상선 장애 유병, 연령, 최장 직업이었다[Fig. 1]. 가장 우선적으로 관여하는 예측요인은 최근 2주간 통증 및 불편감 여부였다. 다음으로 최근 2주간 통증 및 불편감이 없는 경우는 갑상선장애가, 통증 및 불편감이 있는 경우에는 연령이 관여하는 분류

변수였다. 마지막으로, 장년기(40~64세) 및 노년기(65세 이상)에서는 최장 직업 형태가 관여하는 예측변수였다.



[Fig. 1] Decision tree model for self-reported voice problem

다음으로, 각 노드에 대한 이익지표를 도출했을 때, 주관적 음성장애를 효과적으로 예측하는 경로로 3개의 노드가 확인되었다[Table 2]. 먼저, 주관적 음성문제예측에

[Table 2] Gains chart of predictor variable by decision tree model for self-reported voice problem

Node no	Node n (%) <sup>1</sup>	Gain n (%) <sup>2</sup>	Response % <sup>3</sup>	Gain Index % <sup>4</sup>	Description about node
7	314 (8.7)	57 (23.8)	18.2	273.4	Pain and discomfort during the last 2 weeks - yes; Age - >40; The longest occupation - Economically inactive, non manual
4	81 (2.3)	10 (4.2)	12.3	186.0	Pain and discomfort during the last 2 weeks - no; Thyroid disorders - yes
8	488 (13.6)	44 (18.4)	9.0	135.8	Pain and discomfort during the last 2 weeks - yes; Age - >40; The longest occupation - manual

<sup>1</sup> Node n(%); node number, % to 3,600

<sup>2</sup> Gain n(%); gain number, % to 239

<sup>3</sup> Response (%): The fraction of the self-reported voice problem in subjects

<sup>4</sup> Gain index (%):=273.4 in total 5 node

있어서 이익지표(Gain Index) 값이 가장 큰 경로는 최근 2주간 급성 및 만성 질환 등으로 인해 통증 및 불편감을 인지하는 40세 이상의 비경제활동인 또는 비육체노동자로서 18.2%가 주관적으로 음성문제를 인지하는 집단이었고, 이익지표가 273.4%였다. 주관적 음성문제 예측에 있어서 이익지표가 두 번째로 큰경로는 최근 2주 동안 통증 및 불편감을 인지하지는 않지만 갑상선 장애가 있는 집단으로서 12.3%가 주관적 음성문제를 인지하였고, 이익지표는 186%였다.

주관적 음성문제를 예측하는 세 번째 경로는 최근 2주간 급성 및 만성 질환 등으로 인해 통증 및 불편감을 인지하는 40세 이상의 육체노동자로 9.0%가 주관적으로 음성문제를 인지하는 집단이며, 이익지표가 135.8%였다. 도출된 모형의 타당성 및 안정성을 비교하기 위해서 10-fold 교차타당성 검정을 수행한 결과, 크로스 분류모형(cross validation)의 위험지수는 0.21 이었다.

#### 4. 고찰 및 결론

이 연구에서는 대표성 있는 국가통계자료를 이용하여만 19세 이상 지역사회 성인 인구를 대상으로 주관적 음성장애의 잠재적 위험요인의 경로를 파악하기 위한 예측모형을 구축하였다.

본 연구에서의 음성장애 예측을 위한 의사결정나무 모형 결과, 3개의 유의미한 경로가 확인되었다. 이 모형에 따르면, 최근 2주간 급성 및 만성 질환 등으로 인해 통증 및 불편감을 인지하는 40세 이상의 비경제활동인 및 직장인과 최근 2주간 통증이나 불편감을 없지만 갑상선장애가 있는 사람이 주관적 음성장애의 고위험 집단으로 예측되었다. 선행연구에서 확인된 바와 같이 연령, 성, 직업 등의 인구사회학적 요인과 갑상선장애 등은 음성장애의 주요 위험 요인으로 보고되고 있다[10-15]. 이와 같은 연구들에 따르면, 연령이 증가할수록 음성장애의 위험이 높고[13,15,16], 특히, 남성은 60대 이상, 여성은 40대와 50대의 연령에서 음성장애의 위험이 가장 높은 것으로 보고되었는데[17], 이러한 결과들은 주로 40대 이상에서 음성장애의 위험이 높은 것으로 확인된 본 연구의 예측모형을 지지한다.

한편, 직업도 음성장애에 영향을 미치는 주요 요인으로 확인되었는데, 음성을 많이 사용하는 직업에서 음성장애의 발생 비율이 높을 뿐만 아니라[10,12], 근로자는 비경제활동 인구에 비해서 음성문제의 위험이 유의미하게 높은 집단으로 보고되었다[16]. 이는 19세 이상 인구에서 직업이 음성장애의 예측을 위한 주요 분류 변수로 확인

된 본 연구의 결과와 일치한다.

한편, 본 연구에서는 음성문제의 또 다른 예측 경로로 갑상선장애가 확인되었다. 갑상선은 성대와 관련된 상후두신경(superior laryngeal nerve)과 반회후두신경(recurrent laryngeal nerve) 위치해 있기 때문에 갑상선 장애의 합병증으로 음성변화가 초래되기도 한다. 수술이 필요한 갑상선 장애 환자 500명을 대상으로 음성을 검사한 Nam 등 [18]에 따르면, 약 35%가 인후두 역류질환이나 성대결절 등의 후두질환이 발견된 것으로 보고되었는데, 이는 갑상선장애가 음성장애의 유의미한 예측경로라는 본 연구의 모형을 지지하는 결과이다.

나아가, 본 연구의 결과는 선행연구에서 확인되어진 연령, 직업, 갑상선장애 등의 위험 요인이 개별적으로 음성장애에 영향을 미치는 것뿐만 아니라 위험요인의 상호작용(interaction) 및 복합적 작용에 의해 시너지(synergy) 효과가 작용할 수 있다는 점을 시사한다[19]. 따라서 향후 음성질환의 요인을 주제로 한 연구들에서는 개별 요인뿐만 아니라 여러 가지 요인을 동시에 분석하는 다각적인 연구가 필요할 것이다.

이 연구는 대표성 있는 국가통계자료를 이용하였고, 음성장애 관련 요인의 분석에서 전체 인구를 대표할 수 있도록 가중치를 적용하여 위험 요인을 예측하였다는 장점이 있다. 또한, 잠재적 위험 요인을 복합적으로 고려하여 주관적 음성장애의 고위험 집단을 확인하였다는 강점이 있다. 연구의 제한점은 다음과 같다. 첫째, 이 연구의 모형에 포함된 위험요인 외에도 음성문제에 영향을 미칠 수 있는 잠재적 요인이 존재할 가능성이 있다. 둘째, 이 연구의 모형에서는 선행연구에서 확인된 위험요인을 설명변수로 포함하였지만, 선행연구들이 대부분 단면연구에 근거하였기 때문에 연구의 최종 모형에서 도출된 위험요인을 인과관계에 따른 위험요인으로 해석할 수는 없다. 시간적 선후관계에 의한 위험요인임을 입증하기 위해서는 추가로 전향적 연구(prospective study)가 필요할 것이다. 셋째, 발생률이 낮은 인구집단에서의 음성장애 유병률 결과변수로 이용하였기 때문에 분포의 불균형이 초래되었다. 이를 보완하고자, 본 연구에서는 음성장애 유병률을 고려하여 오분류 비용의 가중치를 비대칭적으로 설정하였는데, 이러한 방법은 모형의 타당성에 영향을 미칠 가능성이 있다.

본 연구를 종합하면, 최근 2주간 급성 및 만성 질환 등으로 인해 통증 및 불편감을 인지하는 40세 이상의 비경제활동인 및 직장인과 갑상선장애가 있는 사람은 향후 음성장애로 진행될 가능성이 높은 고위험 집단이므로 조기 관리 및 예방의 필요성이 제기된다.

References

[1] Cohen, S. M., "Self-reported impact of dysphonia in a primary care population: an epidemiological study", *Laryngoscope*, Vol. 120, No. 10, pp. 2022-2032, 2010. DOI: <http://dx.doi.org/10.1002/lary.21058>

[2] Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M., "Voice disorders in the general population: prevalence, risk factors, and occupational impact", *Laryngoscope*, Vol. 115 No.11, 1988-1995, 2005. DOI: <http://dx.doi.org/10.1097/01.mlg.0000179174.32345.41>

[3] Roy, N., Stemple, J., Merrill, R. M., & Thomas, L., "Epidemiology of voice disorders in the elderly: preliminary findings", *Laryngoscope*, Vol. 117, pp. 628-633, 2007. DOI: <http://dx.doi.org/10.1097/MLG.0b013e3180306da1>

[4] Byeon, H., "The Risk Factor of Voice Disorder", *The Korean Academy of Speech-Language Pathology and Audiology*, pp.284-286, 2011

[5] Ministry of Health and Welfare., "The Korea national health and nutritional examination survey 2007-2009", Seoul: Author. 2010.

[6] Byeon, H., "Comparative Analysis of Unweighted Sample Design and Complex Sample Design Related to the Exploration of Potential Risk Factors of Dysphonia", *The Korea academia industrial cooperation society*, Vol. 13, No. 5, pp.2251-2258, 2012.

[7] Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J., "Classification and Regression Trees", CA: Wadsworth. 1984.

[8] Tan, P., Steinbach, M., & Kumar, V., "Introduction to data mining", Boston: Addison Wesley, 2006.

[9] Cristianini, N., & Shawe-Taylor, J., "An introduction to support vector machins", Cambridge: Cambridge Univerity Press. 2000.

[10] Bermúdez de Alvear, R. M., Barón, F. J., & Martínez-Arquerom, A. G., "School teachers' vocal use, risk factors, and voice disorder prevalence: guidelines to detect teachers with current voice problems", *Folia Phoniatrica et Logopaedica*, Vol. 63, pp. 209-215, 2011. DOI: <http://dx.doi.org/10.1159/000316310>

[11] Best, S. R., & Fakhry, C. F., "The prevalence, diagnosis, and management of voice disorders in National Ambulatory Medical Care Survey(NAMCS) cohort", *Laryngoscope*, Vol. 121, No. 1, pp. 150-157, 2011. DOI: <http://dx.doi.org/10.1002/lary.21169>

[12] Chen, S. H., Chiang, S. C., Chung, Y. M., Hsiao, L. C., & Hsiao, T. Y., "Risk factors and effects of voice problems for teachers", *Journal of Voice*, Vol. 24, No. 2, pp.183-192, 2010. DOI: <http://dx.doi.org/10.1016/j.jvoice.2008.07.008>

[13] Herrington-Hall, B. L., Lee, L., Stemple, J. C., Niemi, K. R., & McHone, M. M., "Description of laryngeal pathologies by age, sex, and occupation in a treatment-seeking sample", *Journal of Speech, Hearing Disorders*, Vol. 53 No. 1, pp.57-64. 1988.

[14] Smith, E., Kirchner, H. L., Taylor, M., Hoffman, H., & Lemke, J. H., "Voice problems among teachers: differences in gender and teaching characteristics", *Journal of Voice*. Vol. 12, pp. 328 - 334, 1988. DOI: [http://dx.doi.org/10.1016/S0892-1997\(98\)80022-2](http://dx.doi.org/10.1016/S0892-1997(98)80022-2)

[15] Miller, M. K., & Verdolini, K., "Frequency and risk fators for voice problems in teachers of singing and control subjects", *Journal of Voice*, Vol. 9, No. 4, pp.348-362, 1995. DOI: [http://dx.doi.org/10.1016/S0892-1997\(05\)80197-3](http://dx.doi.org/10.1016/S0892-1997(05)80197-3)

[16] Verdolini, K., & Ramig, L.O., "Review: occupational risks for voice problems", *Logopedics, Phoniatrics, Vocology*. Vol. 26, No. 1, pp.37-46. 2001. DOI: <http://dx.doi.org/10.1080/140154301300109125> DOI: <http://dx.doi.org/10.1080/14015430119969>

[17] Byeon, H., & Hwang, Y., "Gender Differences in Risk Factors of Self-reported Voice Problems", *Phonetics and Speech Sciences*, Vol. 4, No. 1, pp.99-108, 2011.

[18] Nam, I. C., Bae, J. S., Shim, M. R., Hwang, Y. S., Kim, M. S., & Sun, D. I., "The importance of preoperative laryngeal examination before thyroidectomy and the usefulness of a voice questionnaire in screening", *World Journal of Surgery*, Vol. 36, No. 2, pp.303-309, 2012. DOI: <http://dx.doi.org/10.1007/s00268-011-1347-5>

[19] Byeon, H., & Lee, Y., "Laryngeal pathologies in older Korean adults and their association with smoking and alcohol consumption", *Laryngoscope*, Vol. 123, No. 2, pp.429-433, 2013. DOI: <http://dx.doi.org/10.1002/lary.23603>

변 해 원(Haewon Byeon)

[정회원]



- 2009년 8월 : 단국대학교 대학원 언어병리학 (이학석사)
- 2013년 2월 : 아주대학교 예방의 학교실 사회보건학 (이학박사)
- 2011년 3월 ~ 2013년 2월 : 대림대학교 언어재활과 조교수
- 2013년 2월 ~ 현재 : 남부대학교 언어치료청각학과 조교수 및 언어치료연구소장

<관심분야>

음성 의학, 치매 예측 모형, 노년기 인지장애