

Article

Investigation of the Clinical Effectiveness and Prognostic Factors of Voice Therapy in Voice Disorders: A Pilot Study

Ji-Yeoun Lee ¹, Ji-Hye Park ², Ji-Na Lee ³ and Ah-Ra Jung ^{2,*}

¹ Department of Bigdata Medical Convergence, Eulji University, 553 Sanseong-daero, Sujeong-gu, Seongnam-si 13135, Republic of Korea; jylee@eulji.ac.kr

² Department of Otorhinolaryngology, Nowon Eulji Medical Center, Eulji University School of Medicine, 68 Hangeulbiseok-Ro, Nowon-gu, Seoul 01830, Republic of Korea

³ Division of Global Business Languages, Seokyeong University, Seogyong-ro, Seongbuk-gu, Seoul 02173, Republic of Korea

* Correspondence: jungara2030@eulji.ac.kr; Tel.: +82-2-970-8276; Fax: +82-2-970-8265

Abstract: Examining the relationship between the prognostic factors and the effectiveness of voice therapy is a crucial step in developing personalized treatment strategies for individuals with voice disorders. This study recommends using the multilayer perceptron model (MLP) to comprehensively analyze the prognostic factors, with various parameters, including personal habits and acoustic parameters, that can influence the effectiveness of before-and-after voice therapy in individuals with speech disorders. Various methods, including the assessment of personal characteristics, acoustic analysis, statistical analysis, binomial logistic regression analysis, and MLP, are implemented in this experiment. Accuracies of 87.5% and 85.71% are shown for the combination of optimal input parameters for female and male voices, respectively, through the MLP model. This fact validates the selection of input parameters when building our model. Good prognostic indicators for the clinical effectiveness of voice therapy in voice disorders are jitter (post-treatment) for females and MPT (pre-treatment) for males. The results are expected to provide a foundation for modeling research utilizing artificial intelligence in voice therapy for voice disorders. In terms of follow-up studies, it will be necessary to conduct research that utilizes big data to analyze the optimal parameters for predicting the clinical effectiveness of voice disorders.



Citation: Lee, J.-Y.; Park, J.-H.; Lee, J.-N.; Jung, A.-R. Investigation of the Clinical Effectiveness and Prognostic Factors of Voice Therapy in Voice Disorders: A Pilot Study. *Appl. Sci.* **2023**, *13*, 11523. <https://doi.org/10.3390/app132011523>

Academic Editor: Gang Wei

Received: 14 September 2023

Revised: 15 October 2023

Accepted: 19 October 2023

Published: 20 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: voice disorder; prognostic factor; clinical effectiveness; voice therapy; multilayer perceptron model; GRBAS; acoustic parameters; binomial logistic regression analysis

1. Introduction

Speech is an essential means of communication in everyday life, and speech disorders can have a negative impact on an individual's vocal health and social participation [1–7]. Individuals with speech disorders aim to improve their vocal function through speech therapy [8,9]. However, not all patients show the same results from speech therapy. This suggests that personal characteristics and prognostic factors can influence the effectiveness of speech therapy [10,11]. Examining the relationship between the factors and the effectiveness of speech therapy is a crucial step in developing and implementing personalized treatment strategies for individuals with speech disorders [12]. By conducting a comprehensive analysis of prognostic factors, we can identify which patients are likely to derive the greatest benefit from speech therapy. Furthermore, the development of personalized treatment approaches by considering prognostic factors is an important strategy to enhance the effectiveness of speech therapy and improve patients' vocal function and quality of life.

Past studies have explored the association between the prognostic factors of individuals with speech disorders and the outcomes of speech therapy. Zoltan Galaz et al. conducted a partial correlation analysis to explore the connection between initial phonatory characteristics and fluctuations in clinical scores. The XGBoost models demonstrated an

ability to forecast alterations in clinical scores with an error margin ranging from 11% to 26%. This research proposed a potential method for forecasting the advancement of Parkinson's disease (PD) through the acoustic examination of speech patterns [13]. Phan Huu Ngoc Minh et al. examined the extent of the relationship between aerodynamic assessments, acoustic measurements, and auditory-perceptual factors. They observed a robust correlation between local jitter and shimmer and parameters G, R, B, and S, all with a significance level of $p < 0.001$. This investigation found noteworthy associations between these vocal evaluations, highlighting the potential of combined analyses using a multi-parametric approach to provide a comprehensive and unbiased assessment of pathological voice conditions in their early stages [14]. Chang Bin Yun et al. studied the predictive factors for the efficacy of voice therapy for pediatric vocal fold nodules. They found that good prognostic factors for voice therapy in vocal fold nodules were gender and older age. However, acoustic and perceptual measures of the voice before treatment could not predict the effectiveness of voice therapy [15]. In the most recent notable paper, the authors focused on forecasting vocal recovery three months after thyroid surgery using deep neural networks applied to spectrograms. The approach involved utilizing a pretrained model, based on the GRBAS framework, and training it on preoperative and two-week-postoperative voice spectrograms using the EfficientNet architecture combined with long short-term memory (LSTM) in order to predict vocal outcomes at the three-month mark. The results of the correlation analysis for the predicted grade, breathiness, and asthenia scores were 0.741, 0.766, and 0.433, respectively. This research showed the potential for predicting vocal recuperation three months post-surgery via spectrogram analysis [16]. Patrick Schlegel et al. conducted a study to identify clinical parameters that are sensitive to functional voice disorders using boosted decision stumps. Their findings indicated that a smaller subset of parameters (specifically, 4 out of 13) was effective in distinguishing between three groups: one healthy group and two groups with voice disorders [17]. However, these studies have been limited in terms of considering the association between prognostic factors and the outcomes of speech therapy. Some studies have focused only on one prognostic factor or used subjective evaluation tools to assess the effectiveness of speech therapy [14,15]. These limitations make it difficult to derive reliable results from previous studies and have resulted in the lack of a comprehensive understanding of speech therapy.

To overcome these limitations, in this study, we aim to comprehensively analyze prognostic factors in relation to various parameters, including personal habits and acoustic parameters, which influence the effectiveness of speech therapy in individuals with speech disorders. This paper also evaluates changes in vocal quality before and after speech therapy to elucidate the relationship between prognostic factors and the effectiveness of speech therapy. To achieve this, we employ various methods including assessments of personal characteristics, acoustic analysis, perceptual evaluations statistical analysis, and the multilayer perceptron model. By adopting this comprehensive approach, we clarify the correlation between prognostic factors and the effectiveness of speech therapy.

The results of this study are expected to provide useful guidelines to clinical physicians and speech and language therapists. The development of personalized treatment strategies considering prognostic factors can greatly improve the effectiveness of speech therapy and enhance the vocal function and quality of life of individuals with speech disorders. The contributions of this study are summarized as follows:

- Based on a gender analysis reflecting the vocal characteristics of males and females, we investigate the relationship between the effectiveness of voice therapy and predictive factors in relation to each gender's voice.
- This paper introduces correlation analysis before and after treatment based on effectiveness (+) and non-effectiveness (−) in women's and men's voices.
- New parameters proposed according to the characteristics of acoustic parameters predict the effectiveness of speech therapy through binomial logistic regression analysis.
- Multiple experiments were conducted to validate the utility of the proposed parameters, employing the multilayer perceptron model.

- The results highlight the superiority of this system, which predicts the effectiveness of voice therapy by combining gender-analysis-based perceptual modeling and the new parameters.

2. Materials and Methods

2.1. Materials

This study retrospectively reviewed the medical records of 206 patients diagnosed with voice disorders at the Otorhinolaryngology Department of Nowon Eulji Medical Center, Eulji university, from March 2020 to February 2023. The study was conducted with the approval of the Institutional Review Board (IRB No. 2022-04-014), following the research ethics guidelines. Excluding patients who had not completed voice therapy and those who were lost to follow up, a total of 81 patients with voice disorders were included in the study. Table 1 presents the dataset details employed in this study. This dataset comprises recordings of the /a/ vowels obtained from 55 female and 26 male voices, each exhibiting more than 10 distinct pathologies. The purpose of this dataset was to facilitate a gender-based analysis. The subjects were categorized into two groups based on the effectiveness of voice therapy, specifically concerning improvements in the G scale within the GRBAS (grade, roughness, breathiness, asthenia, strain) assessment [14]. The first group was the responsive group, which showed an improvement of at least 1 point, and the second group was the non-responsive group, which did not show any improvement. These scores are represented as an effectiveness metric (+ or −) in both groups, and the numbers are shown in Table 1. Finally, the variables used in this study are shown in Table 1.

Table 1. Information related to the samples in the experimental dataset.

	Female	Male
Number of samples	55 (27 Voice users)	26 (12 voice users)
Average age	51	48
Types of voice disorders (numbers)	Vocal fold polyp (15), vocal nodule (16), thyroid nodule (2), hoarseness (1), muscle tension dysphonia (5), sulcus vocalis (1), dysphonia (8), presbyphonia (3), vocal cyst (4)	Vocal fold polyp (11), vocal cord paralysis (2), mutational dysphonia (1), vocal nodule (1), vocal cyst (2), leukoplakia (1), dysphonia (2), sulcus vocalis (1), muscle tension dysphonia (2), presbyphonia (2), vallecular cyst (1)
Number of responsive samples (effectiveness, +)	41	18
Variables	Smoking status, alcohol status, voice user status, coffee status, fundamental frequency before and after treatment (Hz), jitter before and after treatment (%), shimmer before and after treatment (%), noise to harmonic ratio before and after treatment (NHR, dB), speaking fundamental frequency before and after treatment (SFF, Hz), maximum phonation time before and after treatment (MPT, s)	

2.2. Acoustic Analysis

A microphone was positioned 5 cm from the subject's mouth, and the subject was instructed to produce and sustain the vowel /a/ at their most comfortable pitch and volume [18]. The digitally recorded data were then transferred to a computer, with a sampling frequency of 44.1 kHz, for subsequent analysis using the multidimensional voice program (MDVP) software version 2.3 from Kay Elemetrics [18,19]. Then, measurements such as jitter (%), shimmer (%), the noise–harmony ratio (NHR, dB), fundamental frequency (F0, Hz), speaking fundamental frequency (SFF, Hz), and maximum phonation time (MPT, s) were obtained [9,20]. For this experiment, the measurements were taken before and after the voice therapy process, as shown in Table 1.

2.3. Perceptual Analysis

A speech and language pathologist utilized the grade–roughness–breathiness–asthenia–strain (GRBAS) scale for the perceptual evaluation of the patient’s voice. To assess the severity of dysphonia perceptually, the rater followed instructions for evaluating the G component (overall dysphonia) of the GRBAS using a four-point ordinal scale, as recommended by the Japanese Society of Logopedics and Phoniatrics [21]. This scale ranged from 0, indicating normal voice quality, to 3, signifying severe dysphonia.

Voice therapy was conducted using both direct and indirect methods. The therapy sessions were scheduled once a week, and the number of sessions varied depending on the patient, ranging from one to eight sessions. On average, four therapy sessions were conducted. The therapy sessions had a duration of approximately 40 min per session. The patients themselves evaluated their own voices using the voice handicap index (VHI) and the Korean version of the voice-related quality-of-life (KVQOL) questionnaire before and after the treatment [22,23].

2.4. Statistical Analysis

The statistical analysis was conducted using software packages including IBM SPSS, version 21.0 (IBM Corp., Armonk, NY, USA). Initially, descriptive analyses, such as histograms and box plots, were executed as part of the analysis process. Therefore, means, maximums, minimums, medians, etc., were calculated for all measures to show the difference before and after treatment in female and male voices. Next, we calculated the correlation coefficient and p -value by means of gender and effectiveness for all acoustic parameters. Subsequently, we employed the Pearson square correlation coefficient (r^2), which quantifies the shared variance between two variables. A correlation analysis was conducted to ascertain the statistical significance of the correlation coefficients, with a significance threshold set at $p < 0.05$ for all findings in our study [24]. The normality of the data distribution was examined using the Kolmogorov–Smirnov test. When it came to comparing the distributions of the two groups in this study, the two-sample t -test was employed under the assumption of normality, considering means and standard deviations. Conversely, if the data did not meet the normality criteria, the Mann–Whitney U-test was utilized. The predetermined significance level was set at $p < 0.05$ [25].

Binomial logistic regression analysis is a statistical method used when dealing with categorical data for the dependent variable [26]. It is used when the outcome variable is dichotomous, meaning it has only two possible outcomes [27]. Unlike linear regression, logistic regression uses the logit function to model the data, allowing the prediction of the probability of a certain class or event by fitting data to a logistic curve. The logistic regression model computes the log odds of the dependent variable, which represents the likelihood of an event occurring. It estimates the probability of the occurrence of a binary event based on one or more predictor variables. Logistic regression typically employs maximum likelihood estimation to determine the model parameters and evaluate the impact of independent variables on the dependent variable. It is commonly used to understand the influence of specific variables on an outcome and to make predictions for future observations [26,27].

2.5. Multilayer Perceptron Model

Numerous artificial neural network models have been created for diverse applications across various fields [28–31]. Within this array of network models, the multilayer feed-forward artificial neural network (MLP) stands out as being one of the most frequently employed, and it was also the model of choice in our study [32–35]. All MLP models were created using Python (version 3.8.0). Prior to entering the networks, the inputs underwent normalization using the min–max method [32,33]. The training of the MLP involved supervised learning, wherein a sequence of input and output variables from the training dataset was provided [34,35]. By undertaking iterative adjustments of connection weights, an optimal input–output mapping function was developed. During model training, various

factors, such as the choice of the optimization algorithm, the model structure parameters, and the maximum number of training iterations, were examined. The neural network architecture and connection parameters were fine-tuned until the model's loss function stabilized and achieved the best fitting performance. Following successful training, the model's generalization performance was assessed using an external testing dataset [32,33].

The input characteristics of the MLP models comprised clinical features derived from the patients' medical history, which were linked to the target output attributes, namely, age, gender, smoking, alcohol, coffee, and voice user status. They also included acoustic features such as fundamental frequency, jitter, shimmer, NHR, SFF, and MPT [36–40]. Within the training dataset, the chosen features were employed to construct an MLP model. Each neuron's activation function was configured as the sigmoid function. The learning process employed the adaptive momentum (Adam) algorithm, with cross-entropy serving as the loss function. Utilizing the training dataset, the model's performance underwent assessment through a five-fold cross-validation procedure, which helped to determine the optimal number of hidden layer units and the maximum iteration limit for the model. To prevent overfitting, a regularization coefficient of 0.001 was set. Subsequently, the final MLP model was trained using the entire training dataset, utilizing the best model parameters identified during the evaluation phase.

3. Results

3.1. Descriptive Statistics

Figure 1 shows histogram distributions of the different categories based on the status of the participants. Males tend to smoke more than females, while the patterns regarding alcohol consumption are similar for both genders. Additionally, females tend to consume more coffee than males. Moreover, both females and males frequently confirmed that they were not voice users.

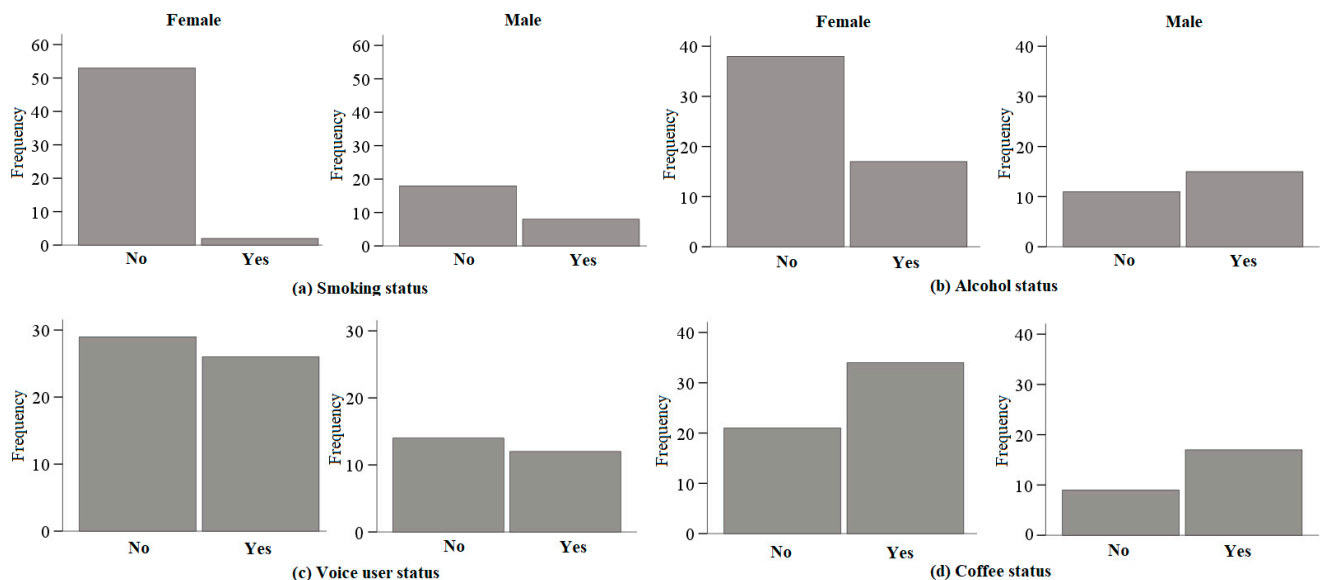


Figure 1. Histogram distributions of different categories based on the status of the participants.

Figure 2 presents distributions of various acoustic parameters in the form of box plots, which provide better visualization before and after treatment for females and males. As shown in Figure 2a, for females, F0 showed a distribution ranging from 80 Hz to 316 Hz before treatment (Pre-tx), with a median of 194 Hz, and a distribution ranging from 131 Hz to 242 Hz after treatment (Post-tx), with a median of 204 Hz. For males, F0 showed a distribution ranging from 85 Hz to 230 Hz before treatment, with a median of 144 Hz, and a distribution ranging from 87 Hz to 235 Hz after treatment, with a median of 147 Hz. As shown in Table 2, the pre- and post-treatment F0 values for both genders were not

significantly different and exhibited a typical positive linear correlation. Based on the results, as F0 exhibits differences based on gender, it is essential to examine the patterns according to gender when analyzing variables related to F0. Jitter reflects the consistency of the oscillation cycle and the variation in the F0 mean, and it is associated with the level of roughness. As shown in Figure 2b, jitter values showed significant differences both before and after treatment for both female and male voices ($p < 0.001$), although jitter extracted from male voices showed a weak correlation. For females, jitter showed a distribution ranging from 0.2% to 7.854% before treatment (Pre-tx), with a median of 2.166%, and a distribution ranging from 0.296% to 2.920% after treatment (Post-tx), with a median of 1.144%. For males, it showed a distribution ranging from 0.318% to 4.340% before treatment, with a median of 2.173%, and a distribution ranging from 0.304% to 2.444% after treatment, with a median of 0.628%. The diagram presented in Figure 2c illustrates the perturbation of glottic vibration, signifying the amplitude of the sound wave. This perturbation is linked to alterations in the level of voice breathiness and variations in intensity. Similar results were observed for shimmer ($p < 0.001$ and $p = 0.013$) as seen for jitter, although shimmer showed weak correlations for both females and males. For females, shimmer showed a distribution ranging from 1.075% to 16.681% before treatment (Pre-tx), with a median of 4.15%, and a distribution ranging from 0.237 to 7.394% after treatment (Post-tx), with a median of 2.792%. For males, it showed a distribution ranging from 1.861% to 28.831% before treatment, with a median of 4.10%, and a distribution ranging from 1.662% to 11.995% after treatment, with a median of 3.107%. In Figure 2d, NHR represents the quantity of noise present within the harmonics of the waveform. A higher NHR value indicates a lower overall sound quality level. SFF and MPT in Figure 2e,f refer to the fundamental frequency at which a person speaks and the maximum duration of sustained phonation or vocalization that a person can produce in a single breath, respectively. The results show that, as shown in Figure 2d–f and Table 2, female voices exhibited significant differences in NHR, while male voices showed significant differences in MPT. Both genders exhibited a strong positive correlation of 0.7 or higher in terms of SFF and MPT.

Table 2. Correlation and p -value analysis before and after treatment in female and male voices.

	Women		Men	
	Pre vs. Post Treatment		Pre vs. Post Treatment	
	Correlation Coefficient	p -Value	Correlation Coefficient	p -Value
F0 (Hz)	0.399 *	0.256	0.587 *	0.647
Jitter (%)	0.307 **	<0.001 **	0.111	<0.001 **
Shimmer (%)	0.154	<0.001 **	0.325	0.013 **
NHR (dB)	−0.015	0.004 **	0.208	0.884
SFF (Hz)	0.792 *	0.635	0.771 *	0.182
MPT (s)	0.695 *	0.061	0.688 *	0.004 **

* and ** mean that the correlation is significant at the 0.01 and 0.05 levels, respectively.

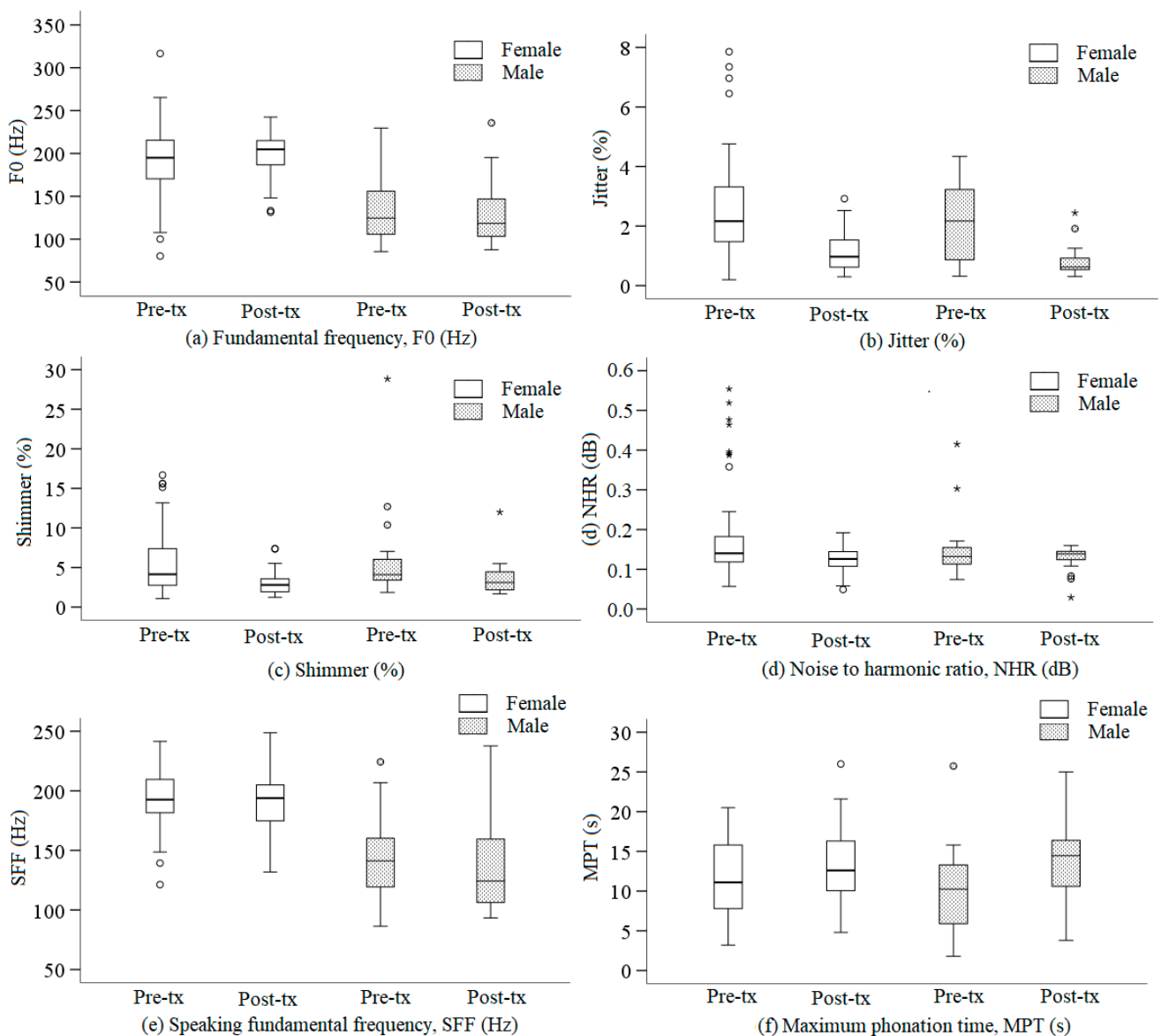


Figure 2. Distributions of acoustic parameters before and after treatment in female and male voices. * Extreme, ° Outlier.

In Figure 3, effectiveness is assessed based on the GRBAS scale, with “G” being the reference point. When comparing the results from before and after voice therapy, if the G score of the post-treatment improves, it is assumed to indicate effectiveness (+). The distributions of various acoustic parameters before and after voice therapy were examined in female and male voices, using effectiveness as the criterion in Figures 3 and 4. They provide an overview of the changes in various acoustic parameters based on the presence (+) or absence (−) of effectiveness. When comparing the distribution of acoustic parameters in female voices between Figures 2 and 3, they exhibited similar patterns. In particular, there was a noticeable tendency toward clear treatment effects in the distribution of jitter, shimmer, and NHR in Figures 3b, 3c and 3d, respectively.

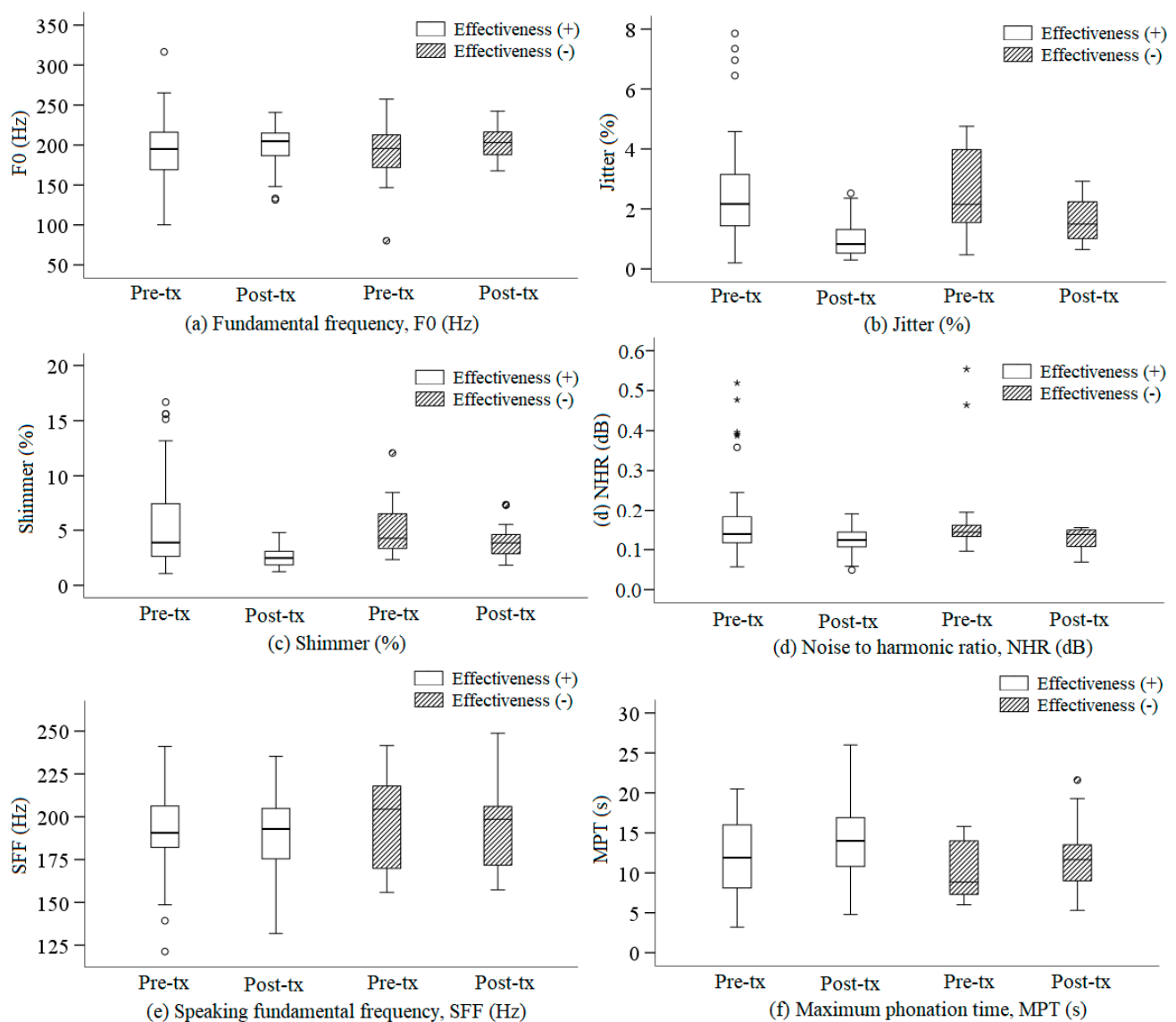


Figure 3. Comparison of acoustic parameters before and after treatment for the effectiveness and non-effectiveness groups in female voices. * Extreme, ° Outlier.

As shown in Table 3, for females, significant results were observed for jitter, shimmer, and NHR when effectiveness was demonstrated after voice therapy. However, when there was no improvement after therapy, none of the parameters showed significant results. Regarding the correlation analysis, when the voice therapy showed effectiveness, strong positive correlations were observed in SFF and MPT.

Table 3. Correlation and *p*-value analysis before and after treatment based on effectiveness in female and male voices.

	Female				Male			
	Effectiveness (+)		Effectiveness (−)		Effectiveness (+)		Effectiveness (−)	
	Before vs. after Treatment		Before vs. after Treatment		Before vs. after Treatment		Before vs. after Treatment	
	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value
F0 (Hz)	0.484 *	0.359	0.101	0.462	0.149	0.291	0.984 *	0.916
Jitter (%)	0.223	<0.001 **	0.818 *	0.129	0.068	<0.001 **	0.661	0.753
Shimmer (%)	0.475	<0.001 **	0.682 *	0.270	0.257	0.002 **	0.681	0.916

Table 3. Cont.

	Female				Male			
	Effectiveness (+)		Effectiveness (−)		Effectiveness (+)		Effectiveness (−)	
	Before vs. after Treatment		Before vs. after Treatment		Before vs. after Treatment		Before vs. after Treatment	
	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value	Correlation Coefficient	<i>p</i> -Value
NHR (dB)	−0.033	0.018 **	0.013	0.108	0.330	0.338	−0.485	0.114
SFF (Hz)	0.791 *	0.735	0.781 *	0.678	0.634 *	0.129	0.950 *	0.674
MPT (s)	0.737 *	0.104	0.507	0.382	0.635 *	0.002 **	0.975 *	0.916

* and ** mean that the correlation is significant at the 0.01 and 0.05 levels, respectively.

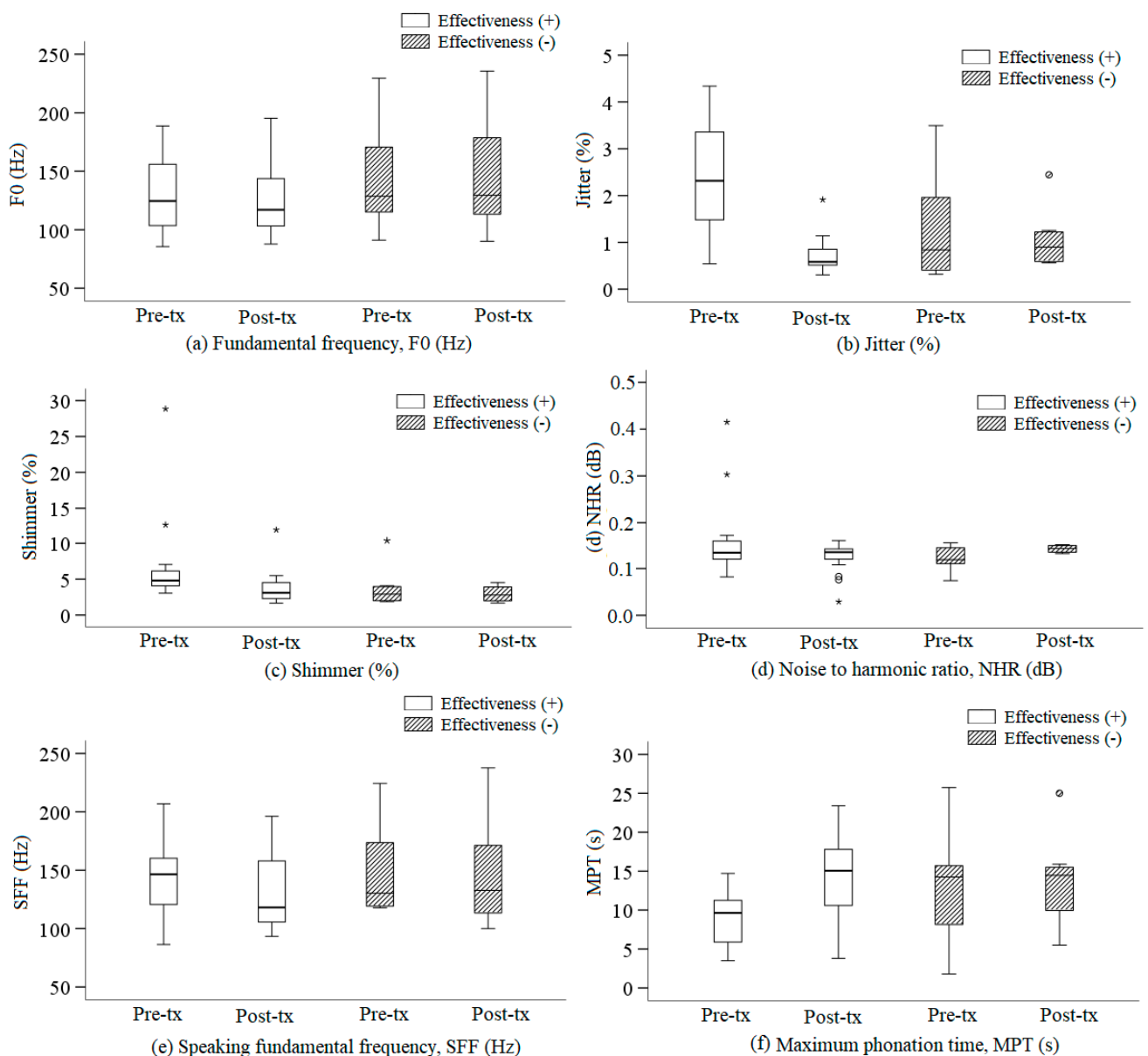


Figure 4. Comparison of acoustic parameters before and after treatment for the effectiveness and non-effectiveness groups in male voices. * Extreme, ° Outlier.

When comparing the distribution of acoustic parameters in male voices between Figures 2 and 4, they exhibited similar patterns. In particular, there was a noticeable tendency of clear treatment effects to be seen in the distribution of jitter in Figure 4b. Based on Table 3, after voice therapy, there were significant improvements in terms of jitter,

shimmer, and MPT, as indicated by the p -values. Additionally, SFF and MPT showed a strong positive correlation. This suggests that voice therapy had a significant positive impact on these parameters. In both females and males, unless there was an improvement in the G scale, no significant differences in the acoustic parameters were observed.

3.2. Binomial Logistic Regression Analysis

Binomial logistic regression analysis was conducted to examine the effectiveness of voice therapy in relation to various acoustic and subjective habit-related variables. The model estimation results are presented in Table 4. The analysis is conducted by starting with a full model that includes all independent variables and then removing insignificant variables using Wald backward elimination. The dependent variable was effectiveness and the independent variables were age, smoking, alcohol, coffee, and the voice user status, F0 before and after treatment (Pre-tx and Post-tx), F0 difference between Pre-tx and Post-tx, jitter before and after treatment, jitter difference, shimmer before and after treatment, shimmer difference, NHR before and after treatment, NHR difference, SFF before and after treatment, SFF difference, MPT before and after treatment, F0 comparison, jitter comparison, shimmer comparison, NHR comparison, and SFF comparison. That is, a total of 27 dependent variables were used in the experiment in Table 4. In this paper, we propose new parameters for predicting the effectiveness of voice therapy using the characteristics of acoustic parameters. The detailed criteria are shown in Table 5. By employing these criteria, we aim to confirm the impact of interventions on speech disorders in relation to F0, jitter, shimmer, NHR, and SFF. For each acoustic parameter, the effectiveness was categorized according to a binary scale of 1 or 0 based on whether the difference between the pre- and post-treatment values was positive or negative. The categorization aligns with the effectiveness based on the G scale. Additionally, the normal thresholds for determining the presence or absence of effectiveness were based on clinical guidelines and prior research in the field [41]. It is important to note that the proposed criteria should be further validated through empirical studies and expert consensus.

Table 4. Model estimation results of binomial logistic regression using all parameters in female and male voices.

Female							Model
Dependent Variable	Independent Variable	B	S.E.	Wald	p Value	Exp(B)	
Effectiveness	Alcohol	2.154	1.204	3.202	0.074	8.622	−2 log likelihood = 34.834 Cox and Snell R ² = 0.502 Nagelkerke R ² = 0.682 $p < 0.001$
	Coffee	2.794	1.243	5.049	0.025 *	16.340	
	Jitter (Post-tx)	−1.951	0.947	4.243	0.039 *	0.142	
	Jitter comparison	4.471	2.008	4.957	0.026 *	87.431	
	Shimmer (Pre-tx)	0.494	0.224	4.871	0.027 *	1.640	
	Constant	−2.155	1.914	1.267	0.260	0.116	
	Male						−2 log likelihood = 21.60 Cox and Snell R ² = 0.332 Nagelkerke R ² = 0.468 $p = 0.005$
	Jitter (Pre-tx)	1.282	0.606	4.470	0.034 *	3.603	
	Jitter (Post-tx)	−2.358	1.150	4.206	0.040 *	0.095	
	MPT(Pre-tx)	−1.202	0.533	5.089	0.024 *	0.301	
	Constant	0.619	1.093	0.320	0.571	1.857	

* $p < 0.05$.

Table 5. The criteria of new parameters based on effectiveness (+).

	Female	Male
F0 comparison	(1) If the difference between the before and after values is negative (i.e., the after value is higher than the before value), it is considered to indicate a positive effect (+). (2) If the difference between the before and after values is positive, the before value should exceed a normal threshold ¹ . Otherwise, it is considered to indicate effectiveness (−).	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ² . (2) If the difference between the before and after values is negative, the before value should be within a normal threshold ² . Otherwise, it is considered to indicate effectiveness (−).
Jitter comparison	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ³ . (2) If the difference between the before and after values is negative, both before and after values should be within a normal threshold ³ (acceptable range of ± 0.02 – 0.03). Otherwise, it is considered to indicate effectiveness (−).	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). (2) If the difference between the before and after values is negative, both before and after values should be within a normal threshold ⁴ (acceptable range of ± 0.01). Otherwise, it is considered to indicate effectiveness (−).
Shimmer comparison	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ⁵ . (2) If the difference between the before and after values is negative, the before value should be within a normal threshold ⁵ . Otherwise, it is considered to indicate effectiveness (−).	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ⁶ . (2) If the difference between the before and after values is negative, both the before and after values should be within a normal threshold ⁶ . Otherwise, it is considered to indicate effectiveness (−).
NHR comparison	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ⁷ (acceptable range of ± 0.01). (2) If the difference between the before and after values is negative, the before value should be within a normal threshold ⁷ (acceptable range of ± 0.01).	(1) If the difference between the before and after values is positive (i.e., the after value is lower than the before value), it is considered to indicate a positive effect (+). Then, the before value should exceed a normal threshold ⁸ (acceptable range of ± 0.01). (2) If the difference between the before and after values is negative, both before and after values should be within a normal threshold ⁸ (acceptable range of ± 0.02).
SFF comparison	Same as F0 comparison	Same as F0 comparison

¹ 200 Hz, ² 118 Hz, ³ 30.82 %, ⁴ 0.58 %, ⁵ 2.97 %, ⁶ 2.72 %, ⁷ 0.12 dB, ⁸ 0.13 dB [41].

In Table 4, for females, the Cox and Snell R^2 value, estimated at 0.502, and the Nagelkerke R^2 value, estimated at 0.682, were relatively high compared to the R^2 values in the regression analysis. Therefore, the model, which represents the relationship between the explanatory variables included in the model and their effectiveness, is statistically significant and considered appropriate. The factors of alcohol status, coffee status, jitter after treatment (Post-tx), jitter comparison, and shimmer before treatment (Pre-tx) were identified as important parameters influencing the effectiveness of voice therapy. Among these, coffee status, jitter after treatment (Post-tx), jitter comparison, and shimmer before treatment (Pre-tx) were found to have a significant impact. For males, the estimated model's Cox and Snell R^2 value is 0.232, and the Nagelkerke R^2 value is 0.330; these values are relatively low compared to the R^2 values in regression analysis. However, based on the p -value, this model was considered statistically significant and appropriate. It was found that jitter before and after treatment (Pre-tx and Post-tx and MPT before treatment (Pre-tx) significantly

influence the effectiveness of voice therapy. Detailed verification results are presented in Table 4. Through the estimated coefficients and Wald statistics of the constructed model, along with the odds-ratio, it is possible to estimate how much each explanatory variable influences the effectiveness of voice therapy.

The regression model presented in Table 6 includes the new parameters representing the scale of the difference between pre- and post-voice-therapy assessments as independent variables, as shown in Table 5. In both cases of females and males, jitter and shimmer comparison were identified as important parameters influencing the effectiveness of voice therapy, but they were found to have no significant impact on the effectiveness of voice therapy. Additionally, the Cox and Snell R^2 value, indicating the explanatory power of the estimated model, was 0.162 and 0.185, while the Nagelkerke R^2 value was 0.220 and 0.261, suggesting a relatively low level. Therefore, it can be concluded that the model is not suitable for conducting regression analyses with new parameters for male and female voices.

Table 6. Model estimation results of binomial logistic regression using new parameters in female and male voices.

Female								
Dependent Variable	Independent Variable	B	S.E.	Wald	Odds Ratio	p Value ¹	Exp(B)	Model
Effectiveness	Jitter comparison	2.151	1.169	3.386	1	0.066	8.596	−2 log likelihood = 63.417 Cox and Snell R ² = 0.162 Nagelkerke R ² = 0.220 p = 0.008
	NHR comparison	1.335	0.686	3.794	1	0.051	3.801	
	Constant	−2.424	1.228	3.897	1	0.048	0.089	
	Male							
	Jitter comparison	2.303	1.378	2.790	1	0.095	10.000	−2 log likelihood = 26.769 Cox and Snell R ² = 0.185 Nagelkerke R ² = 0.261 p = 0.07
	NHR comparison	2.015	1.111	3.292	1	0.070	7.500	
	Constant	−2.708	1.653	2.683	1	0.101	0.067	

¹ $p < 0.05$.

3.3. Multilayer Perceptron Model

This study conducted an analysis of the MLP method with the dataset of voice disorders, and effectiveness factors that may be linked to voice therapy were identified. Using the statistically significant parameters identified through the various regression models shown in Tables 4 and 6, multilayer perception modeling was conducted as shown in Table 7. For females, jitter (Post-tx), shimmer (Pre-tx), coffee-drinking status, and the jitter comparison were found to be statistically significant. For males, jitter (Pre-tx), MPT (Pre-tx), and jitter (Post-tx) were found to be statistically significant.

Table 7. Information related to MLP modeling and results.

		Female	Male
Input layer	Input factors	Coffee status, jitter (Post-tx), shimmer (Pre-tx), jitter comparison,	Jitter (Pre-tx), jitter (Post-tx), MPT (Pre-tx)
	Number of units	6	3
Hidden layer	Number of hidden layers	1	1
	Number of units	2	1
	Activation function	Hyperbolic tangent	Hyperbolic tangent
Output layer	Dependent variable	Effectiveness	Effectiveness
	Number of units	2	2
	Rescaling of scale-dependent variables	Standardized	Standardized
	Activation function	Softmax	Softmax
	Error function	Cross entropy	Cross entropy

Approximately 70% of the total dataset was allocated for the training phase, while the remaining 30% was reserved for testing, as outlined in Table 7. For the female subgroup, the input layer consisted of seven units, the hidden layer included three units, the activation function for the hidden layer was a hyperbolic tangent, the output layer comprised one unit with an identity activation function, and the error function was based on the sum of squares. The optimization of the model's hyperparameters was accomplished using the scaled conjugate gradient method. In the case of the male subgroup, the MLP modeling approach mirrored that of the female subgroup, with the exception of the number of units in the input and output layers.

The MLP model was evaluated using performance metrics including accuracy, sensitivity, specificity, the F-score, and the area under the curve (AUC). AUC-ROC curves were compared as shown in Figure 5. The predictive capacity of the prognostic models was similar according to the AUC-ROC values, with 0.853 and 0.861 for female and male voices, respectively. This finding validates the choice of input parameters used to build our model. Table 8 shows the confusion matrix produced by each MLP, using feature parameters to predict effectiveness in female and male voices. From Table 9, accuracies of 87.5% and 85.71% are shown for the combination of the input variables of Table 10. Additionally, as shown in Table 10, Jitter (Post-tx) and MPT (Pre-tx) were identified as the most important parameters in the MLP models for female and male voices, respectively.

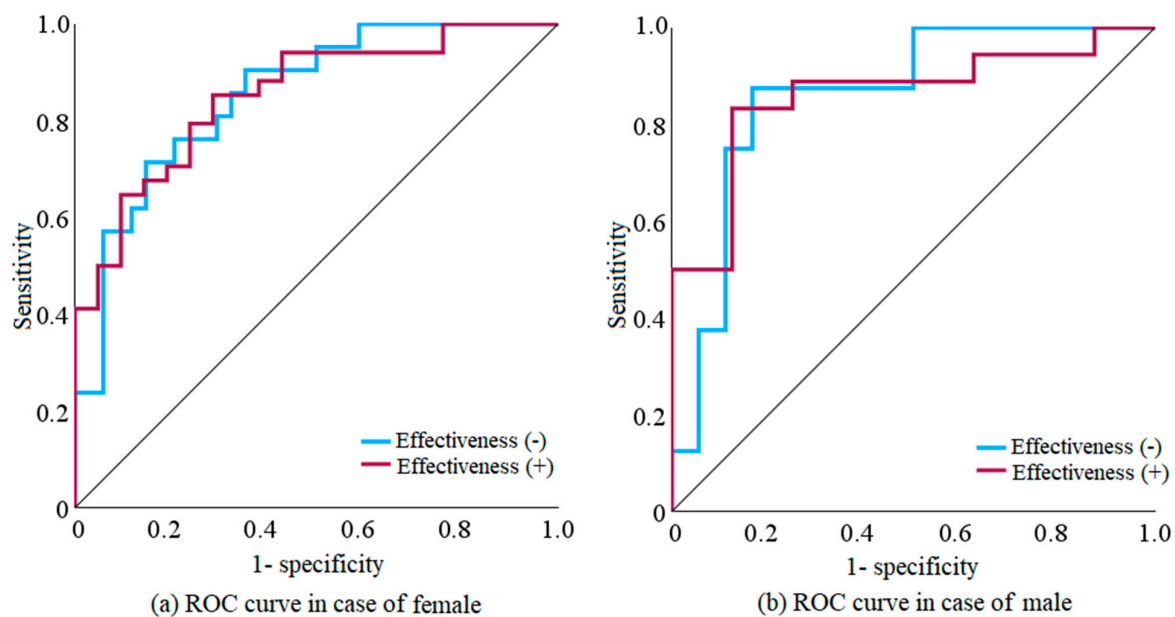


Figure 5. ROC curves of MLP modeling for males and females.

Table 8. Confusion matrices.

Female		Reference		
		Effectiveness (+)	Effectiveness (−)	Total
Predicted	Effectiveness (+)	11	1	12
	Effectiveness (−)	1	3	4
	Total	11	5	16
Male		Reference		
		Effectiveness (+)	Effectiveness (−)	Total
Predicted	Effectiveness (+)	4	1	5
	Effectiveness (−)	0	2	2
	Total	4	3	7

Table 9. Classification performance matrices of the MLP model.

Performance Metrics	Female	Male
	Values	
Accuracy (%)	87.5%	85.71%
Precision	0.92	1.00
Specificity	0.75	1.00
Recall	0.92	0.67
G value	0.83	0.82
F score	0.92	0.80
AUC	0.853	0.861

Table 10. Importance of the input variables.

Input Variable	Female	Input Variable	Male
	Importance		Importance
Jitter (Post-tx)	0.363	MPT (Pre-tx)	0.395
Shimmer (Pre-tx)	0.260	Jitter (Post-tx)	0.361
Coffee-drinking status	0.226	Jitter (Pre-tx)	0.244
Jitter comparison	0.151		

4. Discussion

The main purpose of this study was to identify the prognostic factors that influence the effectiveness of speech therapy in individuals with speech disorders. In the comparison of pre- and post-voice-therapy aspects, for females, significant differences were observed in terms of jitter ($p < 0.001$), shimmer ($p < 0.001$), and NHR ($p = 0.004$), whereas, for males, significant differences were observed for jitter ($p < 0.001$), shimmer ($p = 0.013$), and MPT ($p = 0.004$). In terms of the effectiveness of voice therapy before and after treatment, for females, the jitter ($p < 0.001$), shimmer ($p < 0.001$), and NHR ($p = 0.018$) parameters demonstrated positive efficacy. Similarly, for males, the jitter ($p < 0.001$), shimmer ($p = 0.002$), and MPT ($p = 0.002$) parameters showed favorable effectiveness. Based on this experiment, we were able to identify gender-specific efficacy factors for assessing the effectiveness of voice therapy before and after treatment. The results indicated that the jitter, shimmer, and NHR parameters were significant efficacy indicators for both males and females. This suggests the necessity of conducting gender-specific experiments to accurately evaluate the effectiveness of voice therapy. Two binomial logistic regression analyses were conducted to examine the effectiveness of voice therapy in relation to various acoustic and subjective habit-related variables. For females, the binomial logistic regression model consisting of coffee status, jitter after treatment, jitter comparison, and shimmer before treatment ($p = 0.025$, $p = 0.039$, $p = 0.027$, and $p = 0.026$, respectively) effectively explained its efficacy. According to the real values of the model, the regression model was found to be statistically significant ($p < 0.001$); the Cox and Snell R^2 value was 0.502 and the Nagelkerke R^2 value was 0.682, indicating that this model has an explanatory power of approximately 68% for the dependent variables included in the study. For males, the regression model was found to be statistically significant ($p = 0.005$), and the model had an explanatory power of approximately 33% (Cox and Snell R^2 value = 0.232, Nagelkerke R^2 value = 0.330), which is relatively low compared to the R^2 values in regression analysis. It was found that jitter before and after treatment and MPT before treatment ($p = 0.034$, $p = 0.040$, $p = 0.024$) significantly influence the effectiveness of voice therapy.

In this paper, we propose new parameters for predicting the effectiveness of voice therapy by using the characteristics of acoustic parameters such as F0, jitter, shimmer, NHR, and the SFF comparison. Although the model using the proposed parameters did not show significant results in both males and females, there are plans to retest the utility of the suggested parameters using big data in the future. This paper emphasizes the significance

of introducing multiple parameters. Finally, the MLP model was evaluated to compare its performance, including accuracy, sensitivity, specificity, the F-score, and AUC. Accuracies of 87.5% and 85.71% are shown for the combination of input variables for female and male voices, respectively. This result validates the input parameters selected to build our model. For females, a combination of jitter (Post-tx), shimmer (Pre-tx), coffee-drinking status, and jitter comparison was used. For males, a combination of jitter (Pre-tx), MPT (Pre-tx), and jitter (Post-tx) was found to be statistically significant. Therefore, jitter (Post-tx) and MPT (Pre-tx) were identified as the most important parameters in the MLP models for female and male voices, respectively.

A notable finding in this study's results is that, among the acoustic variables, jitter (Post-tx) for females and MPT (Pre-tx) for males have been identified as factors capable of predicting the effectiveness of voice therapy. Furthermore, the methodology proposed in this paper is believed to be well suited to identifying factors predicting the effectiveness of voice therapy before and after treatment, centered around an effectiveness metric based on the G scale from the GRBAS scale, which exhibits a strong correlation with acoustic variables. Our study has certain limitations. First, our dataset consisted of 81 individuals, including 55 females and 26 males, with more than 10 different pathologies. This sample size may not be sufficient for training a deep neural network effectively. Consequently, we are in the process of planning a validation study, aiming to recruit a larger number of patients with long-term follow-up data. Furthermore, we did not take into account factors that might have influenced interpersonal phonetic outcomes, such as the patient's surgical range, method, or age. It is important to note that the variability in the GRBAS scores can vary depending on the individual conducting the assessment, and prediction of these scores may not fully encapsulate the participant's vocal condition.

5. Conclusions

This study recommends the MLP model for comprehensive analysis of prognostic factor with various parameters, including personal habits and acoustic parameters that can influence the effectiveness of before-and-after speech therapy in individuals with speech disorders. Good prognostic indicators for speech therapy in voice disorders are jitter values (Post-tx) for females and MPT (Pre-tx) for males. Most pre-treatment acoustic and perceptual indices cannot predict the effectiveness of speech therapy. The results of this study are expected to serve as a foundation for promising modeling research utilizing artificial intelligence in the context of speech therapy for voice disorders.

In terms of follow-up studies, it is necessary to conduct further research that utilizes big data to analyze the optimal parameters for predicting the effectiveness of voice disorder treatment.

Author Contributions: Data collection and analysis, J.-H.P. and A.-R.J.; conceptualization, J.-Y.L. and A.-R.J.; methodology, J.-Y.L. and A.-R.J.; software, J.-Y.L.; validation, J.-Y.L.; original draft preparation, J.-Y.L. and J.-N.L.; writing—review and editing, J.-Y.L. and A.-R.J.; visualization, J.-N.L.; funding acquisition, J.-Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) (No. 2022R1H1A209259811).

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Nowon Eulji Medical Center, Eulji University (IRB No. 2022-04-014).

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The sponsor had no involvement in the study design; in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

- Seok, J.; Kwon, T. Artificial Intelligence for Clinical Research in Voice Disease. *J. Korean Soc. Laryngol. Phoniatr. Logop.* **2022**, *33*, 142–155. [\[CrossRef\]](#)
- Remacle, A.; Lefèvre, N. Which teachers are most at risk for voice disorders? Individual factors predicting vocal acoustic parameters monitored in situ during a workweek. *Int. Arch. Occup. Environ. Health* **2021**, *94*, 1271–1285. [\[CrossRef\]](#) [\[PubMed\]](#)
- Naranjo, L.; Perez, C.J.; Martin, J.; Campos-Roca, Y. A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. *Comput. Methods Prog. Biomed.* **2017**, *142*, 147–156. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lopez-de-Ipina, K.; Satue-Villar, A.; Faundez-Zanuy, M.; Arreola, V.; Ortega, O.; Clave, P.; Sanz-Cartagena, M.; Mekyska, J.; Calvo, P. Advances in a multimodal approach for dysphagia analysis based on automatic voice analysis. In *Advances in Neural Networks*; Springer International Publishing: Cham, Switzerland, 2016; pp. 201–211. ISBN 978-3-319-33746-3.
- Gupta, R.; Chaspari, T.; Kim, J.; Kumar, N.; Bone, D.; Narayanan, S. Pathological speech processing: State-of-the-art, current challenges, and future directions. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 6470–6474.
- Zheng, K.; Padman, R.; Johnson, M.P.; Diamond, H.S. Understanding technology adoption in clinical care: Clinician adoption behavior of a point-of-care reminder system. *Int. J. Med. Inform.* **2005**, *74*, 535–543. [\[CrossRef\]](#)
- Sim, I.; Gorman, P.; Greenes, R.A.; Haynes, R.B.; Kaplan, B.; Lehmann, H.; Tang, P.C. Clinical Decision Support Systems for the Practice of Evidence-based Medicine. *J. Am. Med. Inform. Assoc.* **2001**, *8*, 527–534. [\[CrossRef\]](#)
- Andrews, M. *Voice Treatment for Children and Adolescents*; Singular Publishing Group: San Diego, CA, USA, 2002.
- Lee, A.R.; Huh, M.J. Auditory Perceptual Factors of Voice Disorders for by Laypeople. *J. Speech-Lang. Hear. Disord.* **2016**, *25*, 103–111.
- Lee, Y.S.; Lee, D.H.; Jeong, G.E.; Kim, J.W.; Roh, J.L.; Choi, S.H.; Kim, S.Y.; Nam, S.Y. Treatment Efficacy of Voice Therapy for Vocal Fold Polyps and Factors Predictive of Its Efficacy. *J. Voice* **2017**, *31*, 120.e9–120.e13. [\[CrossRef\]](#)
- Henry, L.R.; Helou, L.B.; Solomon, N.P.; Howard, R.S.; Gurevich-Uvena, J.; Coppit, G.; Stojadinovic, A. Functional Voice Outcomes after Thyroidectomy: An Assessment of the Dysphonia Severity Index (DSI) after Thyroidectomy. *Surgery* **2010**, *147*, 861–870. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chhetri, S.S.; Gautam, R. Acoustic Analysis Before and After Voice Therapy for Laryngeal Pathology. *Kathmandu Univ. Med. J.* **2015**, *13*, 323–327. [\[CrossRef\]](#) [\[PubMed\]](#)
- Galaz, Z.; Mekyska, J.; Zvoncak, V.; Mucha, J.; Kiska, T.; Smekal, Z.; Eliasova, I.; Mrackova, M.; Kostalova, M.; Rektorova, I.; et al. Changes in Phonation and Their Relations with Progress of Parkinson's Disease. *Appl. Sci.* **2018**, *8*, 2339. [\[CrossRef\]](#)
- Minh, P.H.N.; Yun, E.M.; Hong, K.H. A Study of the Correlation between Phonetic Parameters during Sustained Vowel and Speech Production with Benign Laryngeal Disorders. *Int. Arch. Commun. Disord.* **2020**, *3*, 1–6.
- Yun, C.B.; Kim, Y.-M.; Choi, J.-S.; Kim, J.W. Predictive Factors for the Efficacy of Voice Therapy for Pediatric Vocal Fold Nodule. *J. Korean Soc. Laryngol. Phoniatr. Logop.* **2021**, *32*, 130–134. [\[CrossRef\]](#)
- Lee, J.H.; Lee, C.Y.; Eom, J.S.; Pak, M.; Jeong, H.S.; Son, H.Y. Predictions for Three-Month Postoperative Vocal Recovery after Thyroid Surgery from Spectrograms with Deep Neural Network. *Sensors* **2022**, *22*, 6387. [\[CrossRef\]](#)
- Schlegel, P.Y.; Kist, A.M.; Semmler, M.; Dollinger, M.; Kunduk, M.; Durr, S.; Schutzenberger, A. Determination of Clinical Parameters Sensitive to Functional Voice Disorders Applying Boosted Decision Stumps. *IEEE J. Transl. Eng. Health Med.* **2020**, *22*, 2100511. [\[CrossRef\]](#)
- Smitsm, I.; Ceuppens, P.; Bodt, M.S.D. A Comparative Study of Acoustic Voice Measurements by Means of Dr. Speech and Computerized Speech Lab. *J. Voice* **2005**, *19*, 187–196. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lovato, A.; De Colle, W.; Giacomelli, L.; Piacente, A.; Righetto, L.; Marioni, G.; de Filippis, C. Multi-Dimensional Voice Program (MDVP) vs Praat for Assessing Euphonic Subjects: A Preliminary Study on the Gender-discriminating Power of Acoustic Analysis Software. *J. Voice* **2016**, *30*, 765.e1–765.e5. [\[CrossRef\]](#) [\[PubMed\]](#)
- Silva, W.J.; Lopes, L.; Galdino, M.K.C.; Almeida, A.A. Voice Acoustic Parameters as Predictors of Depression. *J. Voice* **2021**. *online ahead of print.* [\[CrossRef\]](#) [\[PubMed\]](#)
- Choi, S.H.; Yu, M.; Choi, C. Comparisons of 4-Point GRBAS, 7-Point-GRBAS, and CAPE-V for Auditory Perceptual Evaluation of Dysphonia. *Audiol. Speech Res.* **2021**, *17*, 206–219. [\[CrossRef\]](#)
- Youn, Y.S.; Kim, H.H.; Son, Y.-I.; Choi, H.S. Validation of the Korean Voice-Handicap Index(K-VHI) and the clinical usefulness of Korean VHI-10. *Commun. Sci. Disord.* **2008**, *13*, 216–241.
- Lee, Y.J.; Hwang, Y.J. Comparative Studies on the Self Voice Assessment of Voice Disorder Patients and the Hearer Voice Assessment of a Comparative Group of normal subjects. *Phon. Speech Sci.* **2012**, *4*, 105–114. [\[CrossRef\]](#)
- Choi, H.-J.; Lee, J.-Y. Comparative Study between Healthy Young and Elderly Subjects: Higher-Order Statistical Parameters as Indices of Vocal Aging and Sex. *Appl. Sci.* **2021**, *11*, 6966. [\[CrossRef\]](#)
- Kwak, S.G.; Park, S.-H. Normality Test in Clinical Research. *J. Rheum. Dis.* **2019**, *26*, 5–11. [\[CrossRef\]](#)
- Kim, S.H.; Jeong, G.H. An Analysis for Influencing Factors in Purchasing Electric Vehicle using a Binomial Logistic Regression Model (Focused on Suwon City). *KSCE J. Civ. Environ. Eng. Res.* **2018**, *38*, 887–894.
- Kim, M.J. A Study on WLB (Work-Life Balance) Attributes Affecting Job Satisfaction by Gender by using a Logistic Regression. *Inst. Bus. Manag.* **2018**, *41*, 213–229.

28. Byun, H.W. The Prediction Model for Self-Reported Voice Problem Using a Decision Tree Model. *J. Korea Acad.-Ind. Coop. Soc. (JKAIS)* **2013**, *14*, 3368–3373.
29. Verde, L.; Pietro, G.D.; Sannino, G. Voice Disorder Identification by Using Machine Learning Techniques. *IEEE Access* **2018**, *6*, 16246–16255. [\[CrossRef\]](#)
30. Yoo, J.-H.; Heo, E.-J.; Kim, N.-Y.; Lee, Y.-J.; Kim, G.-W. Predictors of Clinical Efficacy of Oriental Medical Treatment in Patients with Panic Disorder. *J. Orient. Neuropsychiatry* **2015**, *26*, 293–305. [\[CrossRef\]](#)
31. Yun, J.; Shim, H.J.; Seong, C. Classification of muscle tension dysphonia (MTD) female speech and normal speech using cepstrum variables and random forest algorithm. *Phon. Speech Sci.* **2020**, *12*, 91–98. [\[CrossRef\]](#)
32. Mehmet, K. Performance Evaluation of Multilayer Perceptron Artificial Neural Network Model in the Classification of Heart Failure. *J. Cogn. Syst.* **2021**, *6*, 35–38.
33. Gholamreza, P.; Maryam, M.Z. Comparison of Artificial Neural Network and SPSS Model in Predicting Customers Churn of Iran's Insurance Industry. *Int. J. Comput. Appl.* **2020**, *176*, 14–21.
34. Lee, J.; Choi, J. Alcohol Dependence Screening Test Using Artificial Neural Network Analysis: The Sensitivity and Specificity Stud. *J. Korean Acad. Addict. Psychiatry* **2005**, *9*, 102–109.
35. Zhang, Z.; Zhou, D.; Zhang, J.; Xu, Y.; Lin, G.; Jin, B.; Liang, Y.; Geng, Y.; Zhang, S. Multilayer perceptron-based prediction of stroke mimics in prehospital triage. *Sci. Rep.* **2022**, *12*, 17994.
36. Jeong, K.; Kim, S.-T.; Kim, S.-Y.; Roh, J.-L.; Nam, S.-Y.; Choi, S.-H. Factors Predictive of Voice Therapy Outcome in Patients with Unilateral Vocal Fold Paralysis. *J. Korean Soc. Laryngol. Phoniatr. Logop.* **2010**, *21*, 121–127.
37. Tafiadis, D.; Tatsis, G.; Ziavra, N.; Toki, E.I. Voice Data on Female Smokers: Coherence between the Voice Handicap Index and Acoustic Voice Parameters. *AIMS Med. Sci.* **2017**, *4*, 151–163. [\[CrossRef\]](#)
38. Kim, S.; Lee, Y.C.; Kwon, O.E.; Eun, Y. Factors Predicting the Outcome of Voice Therapy in Patients with Polyp or Nodule. *Am. J. Otolaryngol. Head Neck Surg.* **2022**, *5*, 1202.
39. Giuliano, M.; García-López, A.; Pérez, S.; Pérez, F.D.; Spositto, O.; Bossero, J. Selection of voice parameters for Parkinson's disease prediction from collected mobile data. In Proceedings of the 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), Bucaramanga, Colombia, 24–26 April 2019; pp. 1–3.
40. Sachdeva, K.; Shrivastava, T. Dysphonia and its Correlation with Acoustic Voice Parameters. *Int. J. Phonosurg. Laryngol.* **2018**, *8*, 6–12. [\[CrossRef\]](#)
41. Kim, J.O. Acoustic characteristics of the voices of Korean normal adults by gender on MDVP. *J. Korean Soc. Speech Sci.* **2009**, *1*, 147–157.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.