Original Paper

Folia Phoniatr Logop 2006;58:274–288
DOI: 10.1159/000093184

Folia Phoniatrica et Logopaedica

Regression Tree Approach to Studying Factors Influencing Acoustic Voice Analysis

Dimitar D. Deliyski[a] Heather S. Shaw[a] Maegan K. Evans[a] Roumen Vesselinov[b]

[a]Department of Communication Sciences and Disorders, and [b]Department of Statistics, University of South Carolina, Columbia, S.C., USA

Key Words

Acoustic voice analysis · Voice assessment · Fundamental frequency · Perturbation · Data acquisition · Environmental noise

Abstract

Multiple factors influence voice quality measurements (VQM) obtained during an acoustic voice assessment including: gender, intrasubject variability, microphone, environmental noise (type and level), data acquisition (DA) system, and analysis software. This study used regression trees to investigate the order and relative importance of these factors on VQM including interaction effects of the factors and how the outcome differs when the acoustic environment is controlled for noise. Twenty normophonic participants provided 20 voice samples each, which were recorded synchronously on five DA systems combined with six different microphones. The samples were mixed with five noise types at eight signal-to-noise ratio (SNR) levels. The resulting 80,000 audio samples were analyzed for fundamental frequency (F_0), jitter and shimmer using three software analysis systems: MDVP, PRAAT, and TF32 (CSpeech). Fifteen regression trees and their Variable Importance Measures were utilized to analyze the data. The analyses confirmed that all of the factors listed above were influential. The results suggest that gender, intrasubject variability, and microphone were significant influences on F_0. Software systems and gender were highly influential on measurements of jitter and shimmer. Environmental noise was shown to be the prominent factor that affects VQM when SNR levels are below 30 dB.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1021–7762/06/0584–0274$23.50/0

Accessible online at:
www.karger.com/fpl

Dimitar Deliyski, PhD
Communication Sciences and Disorders
University of South Carolina, 1621 Greene Street
Columbia, SC 29208 (USA), Tel. +1 803 777 2245
Fax +1 803 777 3081, E-Mail ddeliyski@sc.edu

Introduction

Accuracy of acoustic voice analysis is essential when used to assist clinicians in the appropriate diagnosis and treatment of vocal pathologies. Multiple factors influence voice quality measurements (VQM) obtained during an acoustic voice assessment. These factors include: gender [1, 2], intrasubject variability (token) [3, 4], microphone [5], environmental noise [6–9], data acquisition (DA) system [8, 10], and analysis software [11]. Knowledge related to the order and extent of influence of these factors has an immediate significance for speech-language pathologists who utilize computer-based acoustic VQM. It is important for clinicians to be aware of these factors and the degree of their influence in order to reduce their impact on VQM.

Gender

It is widely accepted that male and female voices are fundamentally different as a result of basic physiological characteristics, which cause significant differences in fundamental frequency (F_0) [2]. Fluctuations in F_0 and increased noise in female phonations are also related to anatomical characteristics [12]. According to Baken and Orlikoff [2], the relationship between gender and shimmer measurements is not clear.

Intrasubject Variability

A technique utilized to control for intrasubject variability is to obtain the mean VQM from multiple tokens. A study by Scherer et al. [4] investigated the number of tokens needed to acquire a representative perturbation value. For stable voices, at least 6 tokens were recommended and for voices with normal to high levels of instability, at least 15 tokens were recommended. The National Center for Voice and Speech (NCVS) recommends that 10 tokens are needed to obtain reliable perturbation measures [3].

Microphone

A study by Titze and Winholtz [5] demonstrated that the type of microphone used has a significant impact on acoustic voice analysis. The results demonstrated that: condenser microphones and microphones with a balanced output performed the best. Microphone sensitivity and distance were found to have the largest effect on perturbation measures. NCVS [3] recommends that a professional grade condenser microphone with a minimum sensitivity of –60 dB should be used; specifically, a miniature head-mounted microphone with balanced output and a mouth-to-microphone distance less than 10 cm. A recent study by Deliyski et al. [10], which investigated the effects of DA hardware on VQM, was in agreement with Titze and Winholtz [5] and Titze [3].

Environmental Noise

Noise added by the acoustic environment and data acquisition hardware becomes inseparable from noise originating from the larynx in perturbation measurements. According to NCVS [3], recordings should be made in a sound-treated room with 'ambient noise less than 50 dB'. Ingrisano et al. [6] investigated the effect of noise on computer-based analyses of voice samples and found that jitter and shimmer measurements increased as noise floors increased. Carson et al. [8] examined the effect of noise on computer-based analyses of voice samples and suggested that appropriate recording standards are needed to obtain valid and reliable results relative to voice production samples, recording processes, and analysis systems. Deliyski et al. [9] suggested that an instrumental measurement of the noise present in the acoustic environment is imperative when acoustic analysis is used for clinical voice assessment. A recommended signal-to-noise ratio (SNR) level of 42 dB was established, which allowed for errors due to noise in up to 1% of voice measurements. A level to produce reliable results was set at 25 dB SNR maintaining all noise-contributing factors in an acoustic environment within an error rate of 5%.

Data Acquisition Systems

Deliyski et al. [10] investigated the relative performance of different DA environments (microphone-to-DA-system combinations) and the relationship between their technical characteristics and VQM performance. The effective dynamic range, discretization error, and differences in VQM parameters as well as the relationship between these factors for 18 DA environments were compared. The effective dynamic range of the DA environments was the main factor influencing shimmer and had a significant influence on jitter. Discretization error was found to be the main factor influencing the accuracy of F_0 across DA environments. The importance of selecting high-quality DA systems combined with microphones that provide high SNR was emphasized.

Software

The analysis software is another factor that significantly affects VQM, but research on this topic is very limited and inconclusive. The average F_0 is a time-calibrated measurement. However, jitter and shimmer differences due to software are difficult to assess in the absence of a calibrated reference given that jitter and shimmer are acoustic estimates of the frequency and amplitude perturbations of the glottal velocity waveform. Depending on the pitch extraction algorithm these estimates can differ substantially [11].

Purpose and Research Questions

Regression trees derived from the Classification and Regression Trees statistical method (CART) [13] allowed for the representation of the interaction of the aforementioned factors and the order in which the factors influence F_0, frequency perturbation (jitter), and amplitude perturbation (shimmer), as measured by three software systems. As a logical extension of previous findings [9], the investigators were interested in whether the factors would differentially influence VQM over three noise level ranges. Range A included recommended noise levels only (from 66 to 42 dB SNR). Range B consisted of the recommended plus reliable levels (from 66 to 26 dB SNR). Range C included all noise levels studied (from 66 to 10 dB SNR).

The specific research questions were as follows:

(1) What are the interaction effects of gender, intrasubject variability, microphone, noise type, noise level, DA system, and analysis software on VQM? This

question was answered by building an optimal and validated model for each measurement: F_0, jitter and shimmer, using regression trees.

(2) What is the order and relative importance of gender, token, microphone, noise type, noise level, DA system and analysis software as factors influencing VQM? This question was answered by analyzing the results of the Variable Importance Measures (VIM) for each regression tree.

(3) How do these influences differ depending on whether the acoustic environment was controlled for noise or not? This question was answered by comparing the VIM for the three SNR ranges.

Method

The data utilized to answer the abovementioned research questions were obtained using the instrumentation and procedures previously detailed in Deliyski et al. [9].

Instrumentation and Procedures

The instrumentation included: computers, soundcards, microphones, and VQM software. Five DA systems, comprised of computers and soundcards, were utilized in this study. The DA systems included: three desktop computers, one with Computerized Speech Lab Model 4400 by Kay Elemetrics Corp. (CSL), one with an enhanced music soundcard (DT1), and one with a built-in soundcard (DT1), and two laptop computers, one of which (LT1) was coupled to a preamplifier, and the second one (LT2) utilized a built-in soundcard. The DA systems were coupled to six microphones and rotated to balance for microphone-to-DA-system interactions. The microphones included: three desktop condenser microphones (C0, C1, C2), of which C0 had balanced output, two desktop dynamic microphones (D1, D2), and a head-mount condenser microphone with balanced output (H0). Three voice analysis software systems were used in this study: MDVP [14, 15], PRAAT (version 4.1.5) [16, 17], and TF32 (previously known as CSpeech) [18, 19]. The DA hardware, microphones, and voice analysis software were selected based on their frequency of use worldwide for clinical and research voice applications.

Twenty normophonic individuals (10 men, 10 women) participated in the study. They represented a large age range (25–61 years for females and 23–64 years for males) with equivalent mean ages of 39 years. The participants seated in a sound-attenuated booth provided 20 sustained phonations of the vowel /ɑ/ in a habitual pitch for up to 10 s. All participants were positioned at a distance of 30 cm from the desktop microphones. The head-mount microphone, when used, was positioned at a distance of 4 cm and an angle of 45° from the participant's mouth. The intensity was measured at 30 cm distance using a digital sound level meter. The investigators maintained the intensity within 88 dB (A) ± 3 dB by cueing the participants to raise or lower their volume. This intensity was utilized to secure an acoustic environment with an invariant SNR of 66 dB at 22 dB (A) ambient noise level within the sound booth. These phonations were recorded in four sessions, 30 min apart, between which the microphones were rotated amongst the DA systems. The two microphones with balanced output, C0 and H0, were alternated for each rotation on CSL, the only DA system with balance input. The other four microphones, C1, C2, D1, and D2, were rotated amongst the remaining four DA systems, DT1, DT2, LT1, and LT2. Rotations allowed for producing all 18 possible microphone-to-DA-system combinations. Five tokens (1, 2, 3, 4, 5) were obtained from each participant during each recording session. In order to reduce the learning effect, a preliminary session was recorded and discarded. All recordings were made simultaneously on all five DA systems at a sampling rate of 44.1 kHz and a 16-bit quantization. Pulses were presented at the beginning and at the end of each recording session in order to synchronize the five DA systems. All tokens were viewed and a 4-second portion of each type 1 token, according to NCVS recommendations [3], was selected.

Five noise types, chosen to represent the most common noise interference possibilities in a realistic assessment environment, were recorded with a high-quality condenser desktop microphone. The noise types were inclusive of: fan noise from a computer (Fan), 60-Hz penetration noise from the A/C power (60 Hz), band-limited white noise (White), background speech (Talk), and noise from the city (City). The five types of noise were digitally mixed, using Matlab version 6.5, at precise SNR with the 66 dB (Clean) audio samples, resulting in a total of eight noise levels: 66, 58, 50, 42, 34, 26, 18, and 10 dB SNR.

As a result of recording 5 tokens in four separate rotation sessions for 20 participants simultaneously on five DA systems, a total of 2,000 Clean 4-second audio samples were obtained. These 2,000 audio samples were then mixed with five types of noise at eight SNR levels producing a total of 80,000 samples. Three types of VQM, F_0, jitter (RAP, %jit and jitter (rap)), and shimmer (Shim, %shm, and shimmer (local)) were measured using the three software analysis systems, MDVP, TF32, and PRAAT, respectively. The specific jitter and shimmer measures across the three programs were chosen based on the similarity of order of the perturbation function. The nine parameters resulted in a total of 720,000 output values for statistical analysis. The dataset was divided into three noise level ranges to answer the research questions. For range A, 40,000 acoustic samples were analyzed at four noise levels (66, 58, 50, 42 dB SNR). For range B, 60,000 acoustic samples were analyzed resulting from six noise levels (66, 58, 50, 42, 34, 26 dB SNR) and range C included 80,000 acoustic samples that resulted from all eight noise levels (66, 58, 50, 42, 34, 26, 18, 10 dB SNR).

Analysis

The CART approach was selected since it is a binary recursive partitioning method, which is a powerful discovery tool for data with a complex structure [13, 20, 21]. It has been shown that recursive partitioning improves the accuracy of conventional models by an average of 10–15% [20, 22]. The goal in a regression tree approach is 'to partition the data into relatively homogeneous (low standard deviation) terminal nodes, and obtain the mean value observed in each node as the node's predicted value' [20]. Tree building is initiated with the first binary split of the most important variable termed the 'parent' node. Subsequently, this 'parent' node is split repetitively until a terminal node is reached. For the purposes of this study, the terminal node was forced when the next split included less than 200 tokens. For each node, the variable that provides the best split, as determined by the split with the minimum error, is included in the tree. The splitting criterion utilized for this study was the Gini impurity coefficient. For each split, a list of 'competitors' and 'surrogates' are generated. Competitors might replace the 'best' variable at a particular split with a change in the structure of the split. A surrogate is a variable that might replace or 'mimic' the 'best' variable without much change in the split. At any split, the model attempts to predict the dependent variable, and the mean for each node (within each node in fig. 1–3) is the predicted value.

Since trees may be sensitive to random noise/error in the data, a validation process is commonly employed when building regression trees. The validation technique used by CART, cross-validation, allows for building very robust models that are superior to other tree and standard regression models [20]. For this study, '10-fold' cross-validation was employed. The data was divided randomly into 10 equal subsamples (each containing 10% of the data) and the tree building process was repeated 10 times. First, a tree was built using the first 9/10 of the data, while the remaining 1/10 (subsample No. 1) was used to estimate the error rate. Then, the process was repeated on another 9/10 of the data, until each of the 10 subsamples had been utilized to provide an error rate. The resulting subsample error rates for the 10-fold sample iterations were combined to form the error rates for each tree. This procedure ensured the estimation of the independent predictive accuracy of the tree and the confidence that the resulting tree could be generalized to a completely different set of data.

Once the tree is built and validated, the importance of each variable can be ascertained. At each splitting point, both the most influential variable and its surrogates are determined by CART. With conventional models, the importance of one variable is often 'masked' by another variable. For example, in a model using a stepwise procedure, the surrogates would be dropped out of the equation and their actual importance would be obscured. CART solves this 'masking' problem by taking into account the improvement measure not only for the primary splitter, but also for the surrogates. The variable importance score is calculated by 'looking at the improve-

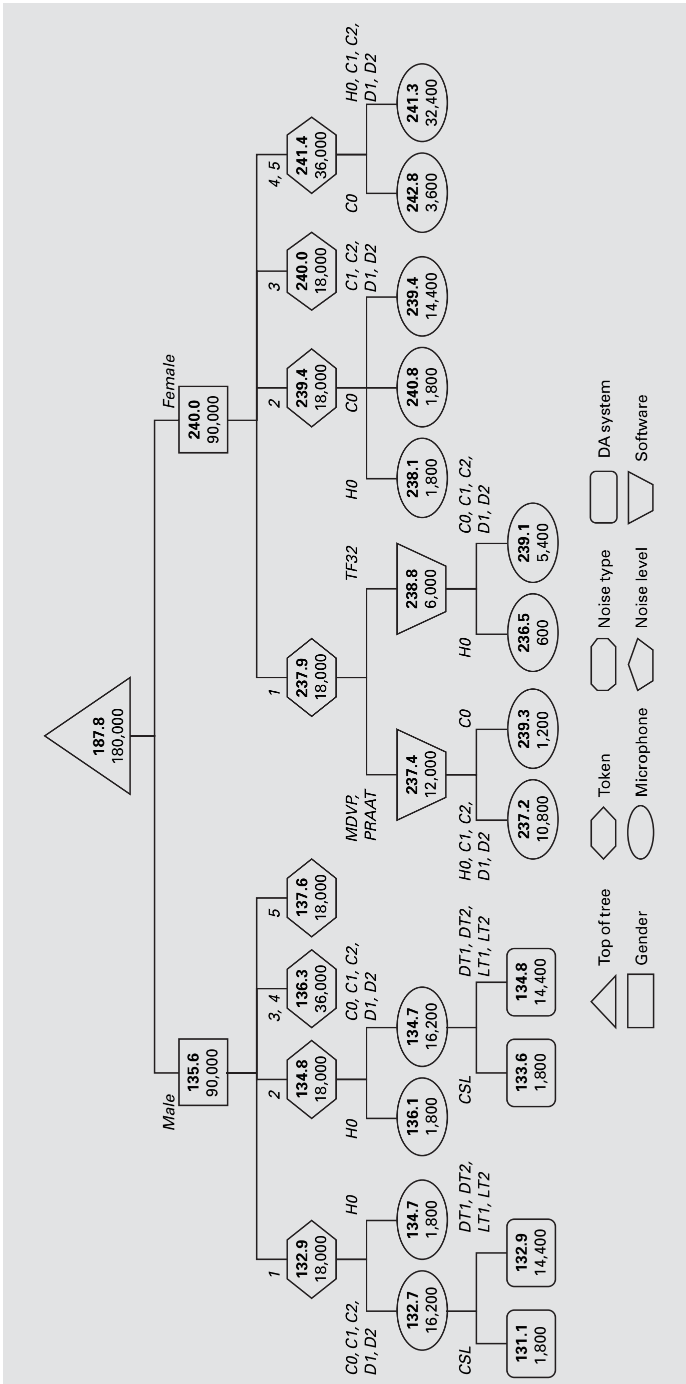

Fig. 1. Schematic of the regression tree for F_0 based on range B with the reported mean F_0 values and number of samples at each node.

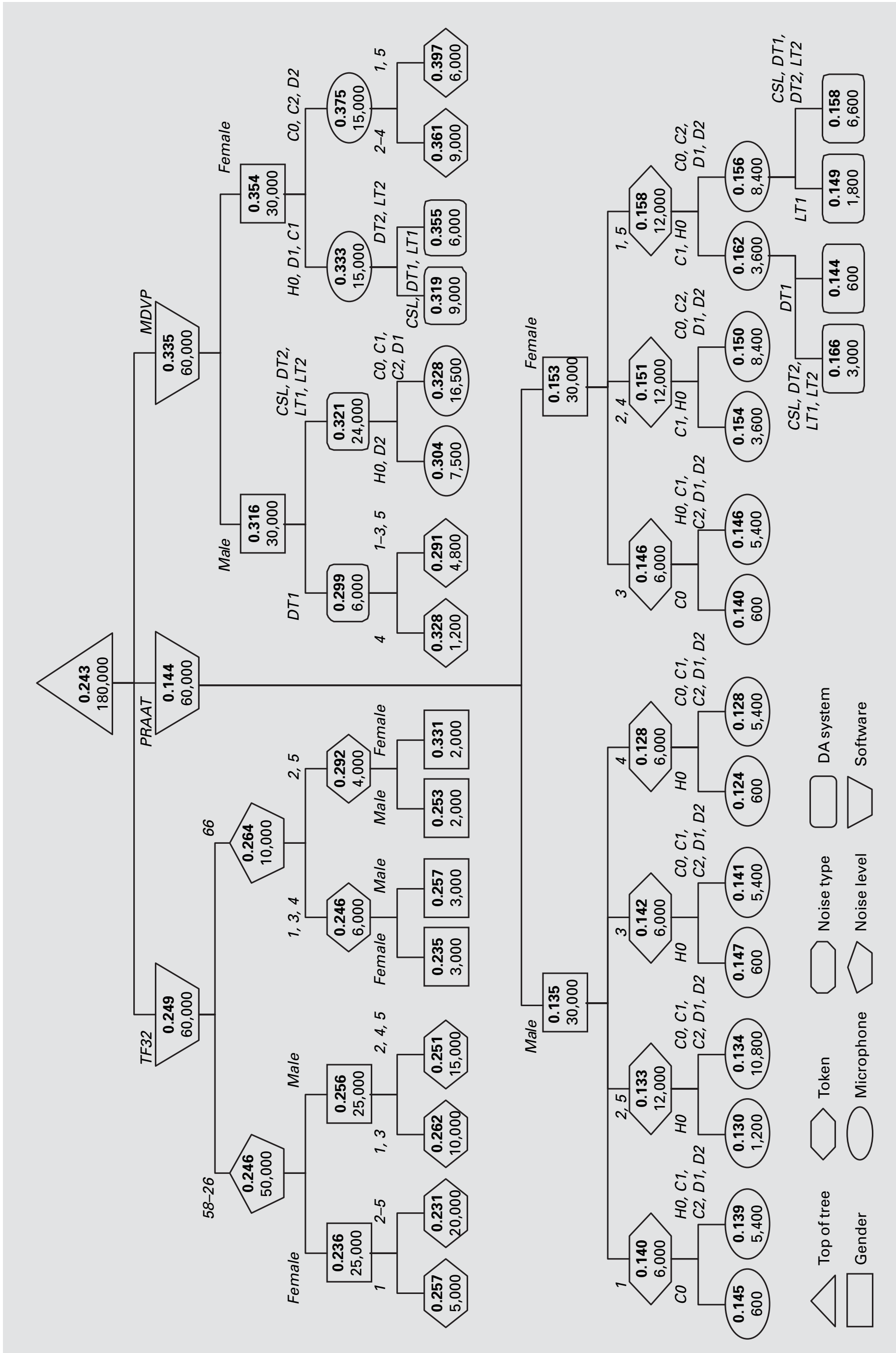

Fig. 2. Schematic of the regression tree for jitter based on range B with the reported mean jitter values and number of samples at each node.

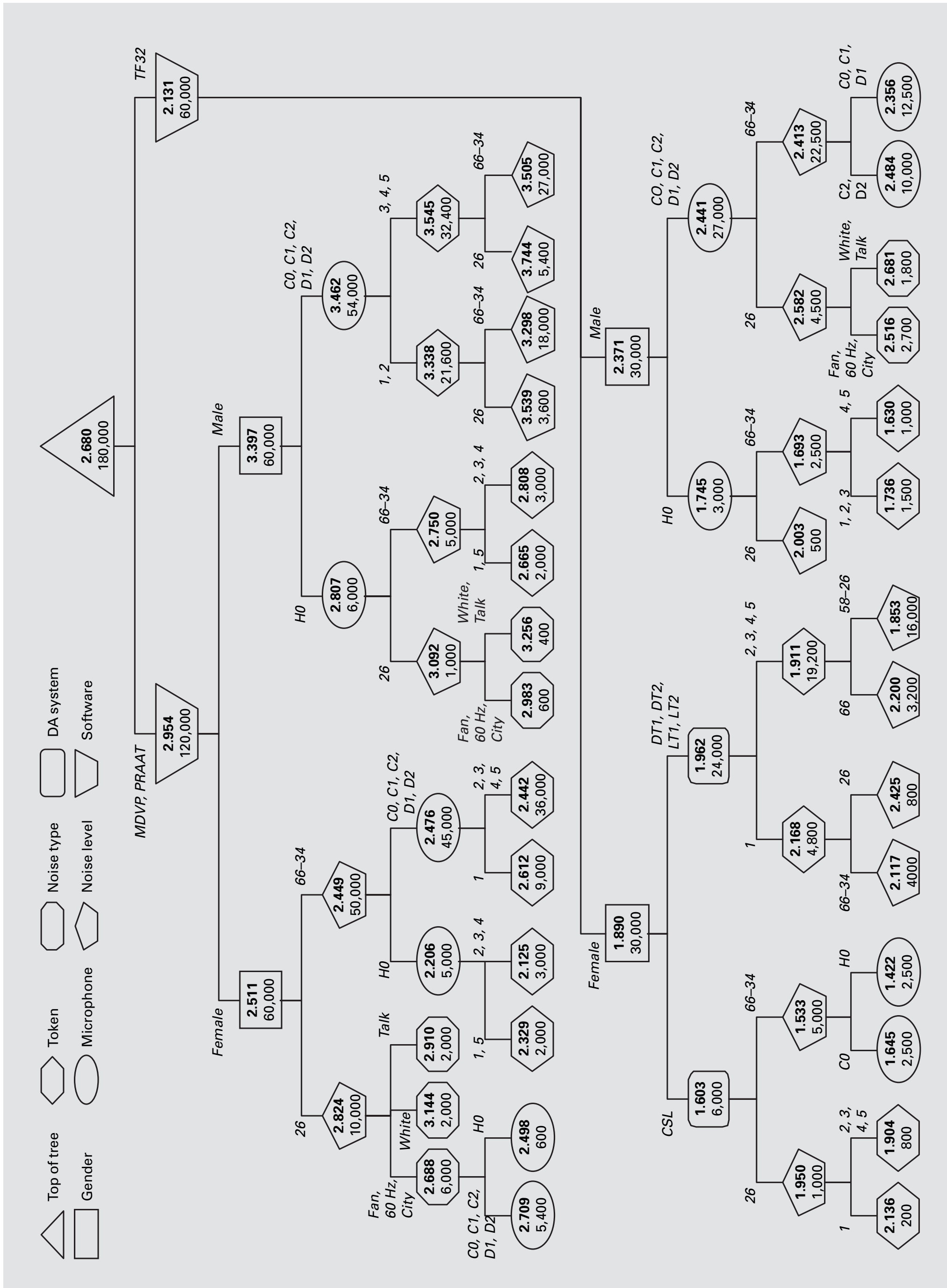

Fig. 3. Schematic of the regression tree for shimmer based on range B with the reported mean shimmer values and number of samples at each node.

ment measure attributable to each variable in its role as a surrogate to the primary split. The values of these improvements are summed over each node, totaled for the tree, and scaled relative to the best performing variable' [20] forming the VIM. That is, the variable with the highest score is set to 100% and the remaining variables are scaled relative to this variable. The ranking is relative to the tree structure, thus, it is very important to validate each tree to obtain results that can be generalized.

The 15 regression trees discussed below were generated by the CART program version 5 [21], and provide a hierarchy of the influential factors. Three noise level ranges for the three types of VQM parameters resulted in nine trees. Six additional trees were utilized to compare measurements of F_0 across gender for each range. The factors included in the first nine trees were: gender, token, microphone, noise type, noise level, DA system, and software system. The factors included in the six trees for F_0 were the same with the exception of gender. After the tree was completed, VIM were obtained for all factors included in the tree. The tree structure and VIM allowed for the assessment of the relative influence of each factor on VQM in the presence of the other variables in the model.

Results and Discussion

Regression Trees

The interpretations of the CART statistical models provide valuable information related to the interaction effects of the factors on VQM. The regression trees for each VQM parameter are shown in figures 1–3. These trees are based on results utilizing range B noise levels. This range was shown to provide information on the impact of noise level and type, while restricting the possible dominance of these factors due to invalid VQM [9].

Fundamental Frequency. The optimal regression tree for F_0 range B resulted in 13 levels and 130 end nodes. This tree was reduced to five levels to provide a schematic and to limit the complexity of interactions investigated (fig. 1). For males, there were no visible variations in the interaction terms amongst the factors with a standard order: gender, token, microphone, and DA system. For females, a greater variation in the interaction terms was observed. The first two factors were gender and token. Then, the difference between tokens influenced the next factor. For tokens 2 through 5, the following and final factor was microphone. However for token 1, the next factor was software followed by microphone. Apparent differences between token 1 and the rest of the tokens for female samples caused TF32 to measure F_0 differently from MDVP and PRAAT, thus software became an important variable. Software might be an important factor for measuring F_0 of female voices because of the higher probability of pitch extraction gross errors compared to males. Some pitch extraction algorithms may be more vulnerable to gender than others. The cause of the different interactions following token 1 for females remains unexplained.

Jitter. The optimal regression tree for jitter range B resulted in 13 levels and 175 end nodes. As a dependent variable, jitter appears 35% more complex than F_0 when measured by end nodes. This tree was reduced to five levels to provide a schematic and to increase the relevance of interactions investigated (fig. 2). Software system was the first variable. For PRAAT, no variation in the interaction terms was observed. The factors sequentially were gender, token, microphone, and DA system. For TF32, the next factor was noise level. For the clean samples (66 dB SNR), the next factor was token followed by gender. For the other levels, the order was reversed, gender was the next factor followed by token. This difference between the order of factors for 66 dB SNR versus the other SNRs highlights the impact that even small amounts of noise may have on VQM. For MDVP, the second factor was gender. Microphone was the next factor for females, while DA system was next in importance for males. Thus, gender has an impact on the importance of microphone and DA system. For females, half of the microphones (C0, C2, D2) influenced the occurrence of token as the next factor, while the other half (H0, D1, C1) was followed by DA system. Thus, for females the type of microphone is a factor in determining the contribution of intrasubject variability and DA system. For males, DT1 was followed by token. The remaining DA systems were followed by microphone. For males, when the DT1 hardware with the MDVP software was utilized, microphone did not appear to be a significant factor. However, when other DA systems were used, intrasubject variability was not significant. The interpretation of this regression tree suggests that the data structure of jitter is highly dependent on the software used. This finding is due to the significant differences in the algorithms of the three programs, which likely do not measure the same voice feature. Therefore, frequency perturbation values obtained with different analysis software cannot be compared.

Shimmer. The optimal regression tree for shimmer range B resulted in 14 levels and 392 end nodes. The number of levels and end nodes illustrates the complexity of shimmer when compared to F_0 (202% more complex) and jitter (124% more complex), concurring with previous findings that shimmer is the VQM most vulnerable to influences. This tree was reduced to five levels to provide a schematic and to limit the complexity of interactions investigated (fig. 3). The first factor for shimmer, like jitter, was software system. However, unlike jitter, MDVP and PRAAT followed the same course throughout the tree, while TF32 diverged. All software systems had the same second factor, gender. For TF32, females split on DA system. The CSL differed from the rest of the systems and was followed by noise level. For SNR levels between 34 and 66 dB, microphone was the following and final factor. However, for 26 dB SNR, token was the final factor. Thus, gender influences the importance of DA systems, noise level, and microphone. For TF32, males split on microphone. The H0 microphone was followed by noise level and token as factors. The remaining microphones split on noise level. For SNRs 34 dB and above, microphone was the final factor with C2 and D2 differing from the other three. For 26 dB SNR, noise type emerged as a final factor with White and Talk noises splitting from Fan, 60 Hz and City. For MDVP and PRAAT, females split on noise level. For SNR 34 dB and above, microphone and then token were the following factors. For 26 dB SNR, noise type and microphone were the following factors. For MDVP and PRAAT, males split on microphone. For H0 noise level was the next factor, while for the desktop microphones, token and noise level were the following factors. These results suggest that amplitude perturbation data obtained with MDVP and PRAAT is fully compatible, while shimmer computed by TF32 has significantly different values and data structure. The regression tree model also provided an independent confirmation of earlier findings [9] that: (1) shimmer is most sensitive to environmental noise influences in comparison to F_0 and jitter; (2) noise can invalidate shimmer values at SNR below 30 dB, and that (3) noise type influences VQM differentially at noise levels below 30 dB SNR.

Variable Importance Measures

The results of the CART statistical models provide valuable information related to the overall importance of the factors that affect VQM. The following results indicate the order of influence of the factors for each VQM. The VIM of each factor is shown in figures 4 and 5. The factors are arranged from most influential to least influential.

Fundamental Frequency. Figure 4a demonstrates how the VIM change for F_0 across the three different noise level ranges. Gender remains the most influential factor across all three noise level ranges. When environmental noise is controlled, at least within recommended or acceptable levels, token (intrasubject variability), microphone, and software are the most important factors. The factors listed after gender have VIM values close to zero. The most prominent changes for all three models occur in the transition from range B to range C. As lower SNRs are included in the data to be analyzed, noise type, software, and noise level become the most influential factors, with the exception of males (fig. 4c) where an increase in importance for noise type was not shown. In agreement with earlier findings [9], the results shown for range C reconfirm the importance of controlling for the type and level of environmental noise. That is, SNR levels below 25 dB are not adequate to perform reliable and valid acoustic voice analyses. These results stress the importance of carefully selecting software systems and microphones as well as using multiple tokens in order to obtain accurate results. To further explore the importance of these factors when gender was not included in the model, F_0 was separately analyzed for males and females. For females (fig. 4b) and males (fig. 4c), there is no change in the order of relative importance of factors between range A and range B. However, for females the VIM values change significantly between range A and range B, demonstrating that microphone, software, and DA system become increasingly important as SNRs decrease from 42 to 26 dB. Results for males suggest that token, microphone, and DA system are the main factors to consider when measuring F_0 in an environment that has an SNR above 26 dB. As shown in figure 4c, token remains the most influential factor for males across all ranges. When the lower SNRs are included (range C), the results show that software and noise level become the most important factors to consider. For females, when SNR decreases below 26 dB, the model illustrates that noise type, software, and noise level become the most important factors. These results further demonstrate the importance of controlling the noise in the environment and selecting appropriate software and hardware when measuring F_0.

Jitter. As seen in figure 5a, for measurements of jitter, the model reveals that software is the most influential factor and remains so across range B. Similarly, gender remains the second most influential factor. For range B, environmental noise (noise level and noise type) does not yet have a significant effect. In fact, DA system and microphone have increased relative importance for range B, suggesting that these factors have more influence on VQM within this SNR range. When data from SNR levels deemed as unacceptable are included in the analysis (range C), noise level, noise type, and software system become the most prominent factors. These results further suggest that when performing an acoustic voice analysis, noise level, noise type, software system, and microphone significantly affect the reliability and validity of the jitter values obtained.

Shimmer. As shown in figure 5b, software, gender, and microphone are the most influential factors when obtaining shimmer measurements within acceptable SNR

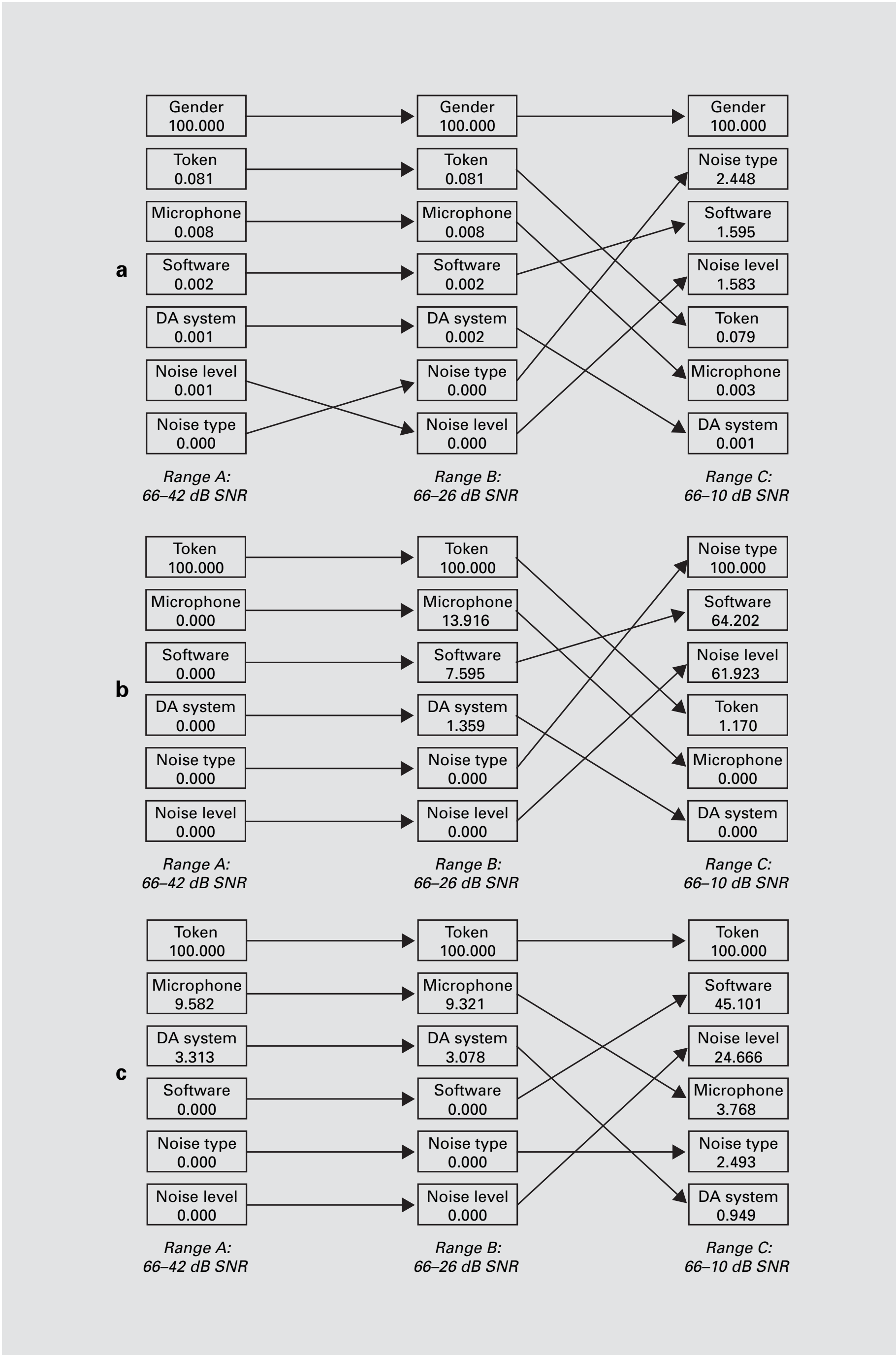

Fig. 4. Block diagram of the factor order and specific VIM for F_0 across the three SNR ranges for all (**a**), female (**b**), and male (**c**) participants.

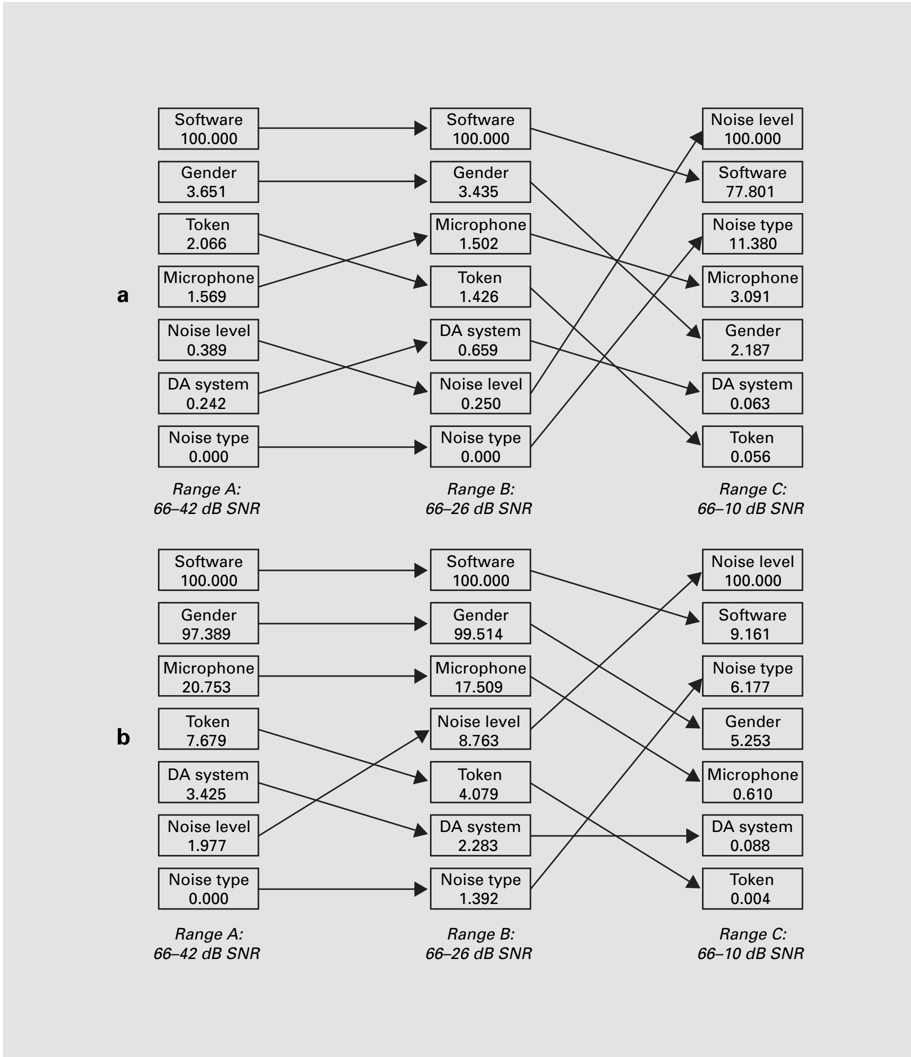

Fig. 5. Block diagram of the factor order and specific VIM for jitter (**a**) and shimmer (**b**), across the three SNR ranges.

levels. Shimmer is affected by noise level as demonstrated in the transition from range A to range B, which displays an increase in the importance of noise level. The most prominent changes overall occur between range B and range C, where the inclusion of lower SNR causes the factors of noise level and noise type to significantly increase in importance. All other factors, with the exception of DA system, decreased in relative importance, demonstrating the degree of influence that noise has on measurements of shimmer.

Conclusion

The results of the study confirm that all of the factors included were significant at varying degrees. Gender strongly influences VQM, which demonstrates the importance of considering gender differences when comparing male and female voices to normative data. Intra-subject variability was shown to be a significant influence for all measurements, especially for F_0 measurements, concurring with the results of Scherer et al. [4] and with the NCVS requirement of implementing multiple tokens [3]. In agreement with Titze and Winholtz [5], it was shown that for range B, the effect of microphone never fell below the third place of importance for all three measurements. The effect of environmental noise, including noise level and noise type, was shown to be the most prominent factor that affects VQM when SNR levels drop below 30 dB. This result concurs with earlier findings [9] that: (1) environmental noise below 30 dB SNR distorts voice samples to the point where normal voices are recognized as abnormal, and (2) the impact of environmental noise below 25 dB SNR exceeds that due to intrasubject variability also causing pitch extraction algorithms to incur gross errors. Although DA system was not shown to be one of the most important factors, it was certainly shown to become highly significant when environmental noise was controlled within acceptable SNR levels. The results demonstrated the effect of differing computational algorithms of the software systems on VQM and further support the observation that more research is needed on this topic. The results reinforce the role of the software system as a highly influential factor making it a critical variable when performing acoustic voice analysis and comparing results across software systems, particularly for measurements of jitter and shimmer. Other factors, such as sampling rate, may have a significant effect on comparing VQM and should be further studied. The results from this experiment demonstrate the importance of controlling for factors that affect VQM in order to obtain accurate and reliable VQM necessary to appropriately diagnose and treat individuals with vocal pathologies.

Acknowledgments

The authors express their appreciation to the University of South Carolina faculty, staff, students, and family members who volunteered to participate in this study; to the Department of Communication Sciences and Disorders for supporting the project; to Dr. Paul Milenkovic of the University of Wisconsin-Madison and Dr. Paul Boersma of the University of Amsterdam for contributing the TF32 and PRAAT software systems, respectively.

References

1 Kent RD, Ball MJ: Voice Quality Measurement. San Diego, Singular Publishing Group, 2000.
2 Baken RJ, Orlikoff RF: Clinical Measurement of Speech and Voice, ed 2. San Diego, Singular Publishing Group, 2000.
3 Titze IR: Workshop on Acoustic Voice Analysis: Summary Statement. Iowa City, National Center for Voice and Speech, 1995.
4 Scherer RC, Vail VJ, Guo CG: Required number of tokens to determine representative voice perturbation values. J Speech Hear Res 1995;38:1260–1269.
5 Titze IR, Winholtz WS: Effects of microphone type and placement on voice perturbation measurements. J Speech Hear Res 1993;36:1177–1190.

6 Ingrisano DR, Perry CK, Jepson KR: Environmental noise: a threat to automatic voice analysis. Am J Speech Lang Pathol 1998;7:91–96.
7 Perry CK, Ingrisano DR, Palmer MA, McDonald EJ: Effects of environmental noise on computer-derived voice estimates from female speakers. J Voice 2000;14:146–153.
8 Carson CP, Ingrisano DR, Eggleston KD: The effect of noise on computer aided measures of voice: a comparison of CSpeechSP and the Multi-Dimensional Voice Program software using the CSL 4300B module and Multi-Speech for Windows. J Voice 2003;17:12–20.
9 Deliyski DD, Shaw HS, Evans MK: Adverse effects of environmental noise on acoustic voice quality measurements. J Voice 2005;19:15–28.
10 Deliyski DD, Evans MK, Shaw HS: Influence of data acquisition environment on accuracy of acoustic voice quality measurements. J Voice 2005;19:176–186.
11 Titze IR, Liang H: Comparison of F_0 extraction methods for high-precision voice perturbation measurements. J Speech Hear Res 1993;36:1120–1133.
12 Kent RD: The Speech Sciences. San Diego, Singular Publishing Group, 1997.
13 Breiman L, Friedman J, Olshen R, Stone C: Classification and Regression Trees. Pacific Grove, Wadsworth, 1984.
14 Deliyski DD: Acoustic model and evaluation of pathological voice production. Proc 3rd Eur Conf on Speech Communication and Technol, EuroSpeech, Berlin, 1993, vol 3, pp 1969–1972.
15 Kay Elemetrics Corp: Multi-Dimensional Voice Program (MDVP) model 5105, version 2.0: Software Instruction Manual. Lincoln Park, Kay Elemetrics Corp, 1999.
16 Boersma P: Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proc Inst Phonet Sci, Amsterdam, 1993, vol 17, pp 97–110.
17 Boersma P, van Heuven V: PRAAT, a system for doing phonetics by computer. Glot Int 2001;5:341–345.
18 Milenkovic P: Least mean square measures of voice perturbation. J Speech Hear Res 1987;30:529–538.
19 Milenkovic P: TF32 User's Manual. Madison, Milenkovic, 2001.
20 Steinberg D, Colla P: CART: Tree-Structured Non-Parametric Data Analysis. San Diego, Salford Systems, 1995.
21 Steinberg D, Colla P: CART: Classification and Regression Trees. San Diego, Salford Systems, 1997.
22 Steadman H, Silver E, Monahan J, Appelbaum P, Robbins P, Mulvey E, Grisso T, Roth L, Banks S: A classification tree approach to the development of actuarial violence risk assessment tools. Law Hum Behav 2000;24:83–100.