The effects of physiological adjustments on the perceptual and acoustical characteristics of vibrato as a model of vocal tremor

Rosemary A. Lester-Smith, and Brad H. Story

Citation: The Journal of the Acoustical Society of America **140**, 3827 (2016); doi: 10.1121/1.4967454
View online: https://doi.org/10.1121/1.4967454
View Table of Contents: https://asa.scitation.org/toc/jas/140/5
Published by the Acoustical Society of America

The effects of physiological adjustments on the perceptual and acoustical characteristics of vibrato as a model of vocal tremor

Rosemary A. Lester-Smith[a] and Brad H. Story
Speech, Language, and Hearing Sciences, University of Arizona, Tucson, Arizona 85721-0071, USA

(Received 26 June 2016; revised 17 October 2016; accepted 28 October 2016; published online 18 November 2016)

The purpose of this study was to investigate the effects of physiological adjustments on listeners' perception of the magnitude of modulation of voice and to determine the characteristics of the acoustical modulations that explained listeners' judgments. This research was carried out using singers producing vibrato as a model of vocal tremor. Twenty healthy adults participated in a perceptual study involving pair-comparisons of the magnitude of "shakiness" with singers' samples, which differed by fundamental frequency, vocal quality, and vowel. Results revealed that listeners perceived a higher magnitude of voice modulation when female samples had a pressed vocal quality. Acoustical analyses were performed with voice samples to determine the features that predicted listeners' judgments. Based on regression analyses, listeners' judgments were predicted to some extent by modulation information in frequency bands across the spectrum.
 [http://dx.doi.org/10.1121/1.4967454]

[JFL] Pages: 3827–3833

I. INTRODUCTION

Vocal tremor is a neurogenic voice disorder characterized by atypical modulation of the fundamental frequency (f_o) and intensity of the voice (Brown and Simonson, 1963; Ludlow *et al.*, 1986). These acoustical modulations lead to the perception of a "shaky" sounding voice (Lederle *et al.*, 2012) and increased effort when speaking (Sulica and Louis, 2010). Behavioral treatment of vocal tremor primarily focuses on training voluntary voice adjustments that reduce the perceptual severity of vocal tremor (Barkmeier-Kraemer *et al.*, 2011). But it is unclear why these physiological adjustments have an effect on listeners' perception of vocal tremor severity. The aims of this study are (1) to investigate the effects of physiological adjustments of voice on listeners' perception of the magnitude of modulation of voice, and (2) to determine the characteristics of the acoustical modulations that explain listeners' judgments.

A recent study of simulated laryngeal vocal tremor involving sinusoidal modulation of the f_o revealed that adjustments to the voice source and vocal tract filter affected listeners' perception of the magnitude of modulation of voice, even when the extent of f_o modulation was the same (Lester and Story, 2015). In the most natural simulations, a greater degree of vocal fold adduction (corresponding to a pressed vocal quality) and a vocal tract shape representing /i/ were perceived as being shakier than a lower degree of vocal fold adduction (corresponding to a breathy vocal quality) and a vocal tract shape representing /ɑ/. Acoustical analyses indicated that listeners' judgments were predicted to some extent by the mean amplitude and the coefficient of variation of amplitude in both low and high frequency bands.

The benefit of using a computational model in the study of vocal tremor was the ability to precisely control the rate, extent, and physiologic source of modulation. However, individuals with vocal tremor do not always exhibit regular modulation of the rate and extent of f_o and intensity modulation (Dromey *et al.*, 2002; Ramig and Shipp, 1987). Furthermore, oscillation within one component of the vocal mechanism may result in adjustments within another component (Lester and Story, 2013) likely due to the interactions between the respiratory system, larynx, and vocal tract during voice production. As a result, computational models of vocal tremor may not represent all of the complexities of this voice disorder. These complex interactions, as well as the fact that tremor may affect the respiratory system, larynx, vocal tract or a combination of these components (Bové *et al.*, 2006; Hachinski *et al.*, 1975; Koda and Ludlow, 1992; Warrick *et al.*, 2000), make it difficult to study the relation of physiological oscillations, acoustical modulations, and listener perception with individuals who have vocal tremor. Fortunately, singers producing vibrato could also serve as a model of vocal tremor, which would allow for some level of isolation of the sources of modulation while preserving more human-like interactions within the vocal mechanism and natural irregularities in the voice modulation.

Vibrato is a singing technique that is characterized by volitional modulation of voice and has acoustical and physiological similarities to vocal tremor. Both vocal tremor and vibrato are characterized by f_o and intensity modulation, with similar rates and extents of modulation (Ramig and Shipp, 1987). Studies of vocal tremor and vibrato using electromyography indicate that the physiological bases of these phenomena are also similar. Koda and Ludlow (1992) found that two intrinsic laryngeal muscles, the thyroarytenoid (TA)

[a]Current address: Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL 60611, USA. Electronic mail: rosemary.lester@northwestern.edu

and cricothyroid (CT) muscles, were most frequently affected by tremor. Similarly, the TA and CT as well as the lateral cricoarytenoid (LCA) muscles exhibited modulated activity when singers produced vibrato (Niimi *et al.*, 1988a; Niimi *et al.*, 1988b; Vennard *et al.*, 1970). Although tremor may affect the respiratory musculature in some individuals with vocal tremor, Appelman and Smith (1985) found that the respiratory muscles were not involved in production of vibrato. These acoustical and physiological similarities between vocal tremor and vibrato make singers producing vibrato a reasonable model of laryngeal vocal tremor.

The purpose of this study was to determine if the findings of Lester and Story (2015) would hold true with the human voice, when the acoustical modulations are not sinusoidal and when the components of the vocal mechanism have natural interactions. The aim of Study 1 was to determine if trained singers producing vibrato with modulation of the f_o could volitionally modify voice source and vocal tract filter characteristics (i.e., f_o, vocal quality, vocal tract configuration for vowels) to reduce listeners' perception of the magnitude of modulation of voice. It was hypothesized that listeners would perceive a higher magnitude of modulation of voice when samples had a higher mean f_o, pressed vocal quality, and vocal tract shape with clustering of resonant peaks in higher frequency regions (i.e., /i/), based on Lester and Story (2015). The aim of Study 2 was to determine the acoustical characteristics that were associated with differences in perception of the magnitude of modulation of voice. It was expected that listeners' judgments would be predicted by modulations in low and high frequency regions of the voice spectrum, as was seen in the previous study. The finding of this study will have implications for behavioral treatment of vocal tremor and will clarify the acoustical bases for listeners' perception of vocal tremor severity.

II. STUDY 1

A. Method

1. Singers

Four classically trained singers (two female, two male), who reported and exhibited a good understanding of written and spoken English and denied a history of cognitive, neurological, upper respiratory tract, speech, and hearing disorders, participated in this study. Their mean age was 30 yr (range = 19–37 yr). Prior to the experimental trials, all singers passed a bilateral hearing screening using pulsed pure tones at 500, 1000, 2000, and 4000 Hz at 25 dB (American Speech-Language-Hearing, 1997). This study was approved by the University of Arizona Institutional Review Board (UA IRB).

2. Data collection

Audio recordings were obtained for each singer in a sound-treated booth using an AKG C420 head-mounted condenser microphone. Simultaneously, electroglottographic (EGG) signals were obtained using the EGGs for Singers electroglottograph with surface electrodes placed bilaterally on the lamina of the thyroid cartilage. EGG recordings were obtained to document the rate of vocal fold vibration and degree of vocal fold adduction (Fourcin, 1981). The audio and EGG signals were routed through a Tascam US-122 MK II audio/MIDI interface to a desktop computer with VoceVista Pro software (VoceVista, 2013, Version 3.2). The sampling frequency of the audio recordings was 44 100 Hz. A spectrogram was displayed in VoceVista, and singers were oriented to the information represented in the spectrogram for visual feedback during audio recordings.

Singers were asked to produce vibrato with sustained vowels /i/ and /ɑ/ at two different target notes (A2 and E3 for males, A3 and C4 for females) and with two different vocal qualities (breathy and pressed). Target notes were selected to correspond with the fundamental frequencies used in Lester and Story (2015). At least twelve audio recordings were obtained for each singer with all combinations of vowel, pitch, and quality, and one repetition of each combination. Singers were asked to sustain the vowels for as long as they comfortably could; therefore, signals varied in duration.

3. Audio stimuli

One second segments from approximately the mid-point of the audio recordings were analyzed for the rate and extent of f_o and intensity modulation using a custom-written Praat script (Boersma & Weenink, 2011, Version 5.2.14). The first male singer's recordings and the second female singer's recordings were selected for the perceptual and acoustical studies because their rate and extent of f_o modulation was relatively consistent within and between recordings and they denied difficulty producing vibrato. The male's recordings had an f_o modulation rate ranging between 5.2 and 5.8 Hz and an extent ranging between 1.2% and 3.3%. The female's recordings had an f_o modulation rate ranging between 5.0 and 5.4 Hz and an extent ranging between 3.0% and 5.6%. A 0.1 s long silence was included before and after each one-second-long vowel and a 0.05 s long fade in and out was imposed for each vowel to avoid a click at the onset and offset.

4. Listeners

Twenty listeners (17 female, three male) with a mean age of 21 yr (range = 18–28 yr) were included in the perceptual study. Prior to the experiment, all listeners passed a bilateral hearing screening using pulsed pure tones at 500, 1000, 2000, and 4000 Hz at 25 dB (American Speech-Language-Hearing, 1997). All listeners reported and exhibited a good understanding of written and spoken English and denied a history of cognitive disorders. This study was also approved by the University of Arizona Institutional Review Board (UA IRB).

5. Listening task

The methods for the current listening experiment were identical to the methods used in Lester and Story (2015). Listeners were seated individually in a sound-treated booth. Audio stimuli were presented with an intensity of 60 dB sound pressure level (SPL) over a loudspeaker placed 1 m in front of the listener. An ALVIN interface (Hillenbrand and Gayvert, 2005) was used to present stimuli and record listeners' responses. Paired stimuli were presented with an inter-stimulus interval of 0.5 s and an inter-trial interval of 1 s. There was a 0.1 s-long silence included at the beginning and end of each stimulus. Therefore, a 0.7 s silence was included between each vowel within a pair and a 1.1 s silence between the second vowel in one trial and the first vowel in the next trial. Each trial advanced when listeners made their selection of which vowel in the pair was "shakier."

Prior to the experiment, each listener participated in a brief training protocol consisting of 16 pairs of simulated vowels that were produced using a kinematic model of the vocal folds (Titze, 1984, 2006) coupled to a wave-reflection model of the trachea and vocal tract (Liljencrants, 1985; Story, 2005). These simulated vowels had different extents of modulation (1%, 5%, 10%, 20%), one neutral vowel (/ə/), and one f_o (either 123 Hz for the male simulations or 247 Hz for the female simulations). Each pair was presented in AB and BA orders. A pair-comparison approach was selected over a perceptual rating scale because pair comparisons were used in a previous study of the perception of respiratory-induced vocal tremor and had high intra-rater agreement (Farinella *et al.*, 2006). Listeners were provided with verbal and written instructions to identify which stimulus in each pair was "shakier," to replay the pairs of stimuli as many times as needed, and to take as many rest breaks as needed. Participants were instructed to "guess" if they were unsure which item in the pair was shakier. Shaky was selected as the descriptor based on Lederle *et al.* (2012) and Lester and Story (2015).

For the experimental trials, participants were again presented with pairs of stimuli and were asked to identify which stimulus in the pair was shakier. The stimuli in these pairs were either identical or differed by one characteristic only (i.e., f_o, vocal quality, or vowel). With all possible combinations of f_o, vocal quality, and vowel for the female singer or the male singer, counterbalancing for order using AB and BA presentations, and identical pairs, 64 trials were presented to each listener. Listeners were given same instructions as they were during the training for replaying stimuli, taking rest breaks, and guessing. The order of presentation of the female and male simulations was randomly assigned and counterbalanced.

6. Data analysis

For each contrast (f_o, vocal quality, and vowel), the percentage of trials (out of 16) that the participants selected high f_o over low f_o, pressed vocal quality over breathy vocal quality, and /i/ over /ɑ/ as being shakier was determined. Intra-rater reliability was calculated by determining the percentage of trials (out of four) when each participant's judgments were consistent across items using the aggregate function with the break variables of participant and stimulus pair in SPSS (IBM, Version 22). Inter-rater reliability was calculated by determining the percentage of trials when all participants' judgments were the same across items using an aggregate function with the break variable of stimulus pair only.

7. Reliability and agreement

Intra-rater reliability was 83% for the female stimuli and 84% for the male stimuli. Inter-rater agreement was 70% for the female stimuli and 78% for the male stimuli.

8. Statistical analyses

The data were analyzed using repeated measures analysis of variance (ANOVA) with sex of the singer (male, female) and contrast (f_o, vocal quality, vowel) as within-subjects factors in SPSS (IBM, 2013, Version 22). Planned one-tailed, one-sample t-tests with a test value of 50 (to account for chance) were performed in SPSS. *Post hoc* analyses of effect size were calculated using Cohen's d with a mean H0 of 50, mean of the actual data, and standard deviation of the actual data in G*Power (Faul, Erdfelder, Lang, & Buchner, 2014, Version 3.1.9.2)

B. Results

The repeated measures analysis of variance (ANOVA) test of within-subjects effects revealed that there were significant main effects of sex [$F(1,19) = 55.586$, $p < 0.001$] and contrast [$F(2,38) = 9.084$, $p < 0.005$] and a significant interaction of sex and contrast [$F(2,38) = 24.408$, $p < 0.001$]. The means and standard errors are presented in Fig. 1. T-tests revealed a significant effect of quality for the female stimuli in the predicted direction ($M = 84.689$, $SD = 19.710$); $t(19) = 7.871$, $p < 0.0005$, $d = 1.760$). All other t-tests were not significant in the predicted direction: f_o for male stimuli ($M = 40.938$, $SD = 10.434$); $t(19) = -3.884$, $p < 0.0010$, $d = 0.869$); f_o for female stimuli ($M = 59.750$, $SD = 25.228$); $t(19) = 1.728$, $p < 0.1000$, $d = 0.386$); quality for male stimuli ($M = 35.625$, $SD = 18.596$); $t(19) = -3.457$, $p < 0.0050$, $d = 0.773$), vowel for male stimuli ($M = 42.188$, $SD = 11.267$); $t(19) = -3.101$, $p < 0.0050$, $d = 0.693$), vowel for female stimuli ($M = 41.563$, $SD = 18.393$); $t(19) = -2.052$, $p < 0.0500$, $d = 0.459$).

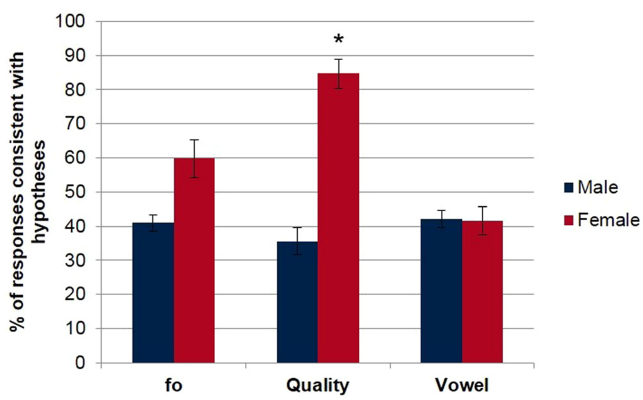

FIG. 1. (Color online) Mean percentage of responses consistent with hypotheses for male and female productions with contrasts of f_o, vocal quality, and vowel with error bars representing the standard error.

C. Discussion

The results of this study revealed that the vocal quality contrast with the female stimuli was the only contrast that significantly affected listeners' perception of the magnitude of modulation of voice in the direction that was predicted. The trends in the perceptual patterns differed between the male and female stimuli for the vocal quality and the f_o contrasts. With the vocal quality contrasts, a breathy voice was more often perceived as being shakier than a pressed voice for the male stimuli. However, this finding was not significant based on the one-tailed t-test because the direction of the pattern differed from the hypothesized direction. With the f_o contrasts, a high f_o was more often perceived as being shakier than a low f_o with the female stimuli, although this effect was not significant. A low f_o was more often perceived as being shakier than a high f_o with the male stimuli. The difference in the direction of the pattern from the hypothesized direction led to a non-significant finding. The patterns of perception of vowel contrasts were the same with female and male stimuli, with the vowel /ɑ/ being perceived more often as shakier than /i/ with both the male and female stimuli. The direction of this pattern differed from the hypothesis, leading to a non-significant finding.

The results of the current study with singer-simulated vocal tremor differed from the perceptual study with computer-simulated vocal tremor reported by Lester and Story (2015). It is presumed that these differences in the perceptual results are related to some acoustical differences in the output of the female singer compared to the male singer and in the vocal tract shaping of the singers compared with the simulated vocal tract shape in the previous study. Acoustical analyses were needed to determine the bases for these differences in perception.

III. STUDY 2

A. Method

1. Acoustical and perceptual data

The acoustic signals for Study 2 were the experimental audio stimuli used in Study 1. All of the stimulus pairs from Study 1, except for the matched pairs, were included in the acoustical analyses for Study 2.

2. Acoustical analyses

The methods for the current acoustical analyses were identical to the methods used in by Lester and Story (2015). Acoustical analyses were performed using custom-written Matlab (The Mathworks, 2011, Version 7.13.0.564 [R2011b]) scripts that were based on the American National Standard Specification for Octave-Band and Fractional-Octave-Band Analog and Digital Filters [ANSI S1.1-1986 (ASA 65-1986), 1993]. The root mean square (RMS) amplitudes within each octave band were normalized to the total RMS amplitude (on a linear scale) for each signal. The acoustic signals were then analyzed for the mean amplitude and the standard deviation of the amplitude within the six standard octave bands with the center frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz. These bands are shown in Fig. 2. The coefficient of variation of the amplitude was then calculated by dividing the standard deviation by the mean for each octave band to represent the extent of amplitude modulation within each band.

3. Statistical analyses

For each contrast (f_o, vocal quality, and vowel), the participants' perceptual judgments (i.e., stimulus 1 is "shakier" or stimulus 2 is "shakier") were modeled using binary logistic stepwise regression with potential predictor variables representing: (1) the ratios of the mean amplitudes in each of the six octave bands for each of the stimulus pairs, and (2) the ratios of the coefficients of variation in each of the six octave bands for each of the stimulus pairs. These two models were calculated separately for female and male stimuli to more precisely model the perceptual results of Lester and Story (2015).

The order of entry of the variables into the forward conditional stepwise regression model was based on a conditional likelihood ratio test with criterion of variable entry of $p < 0.05$ and removal of $p > 0.10$. The variables entered into the model were the best predictors that a listener would select one stimulus over another as sounding shakier. These models were calculated using SPSS (IBM, Version 22).

B. Results

Examples of the results of the acoustical analyses are presented in Figs. 3 and 4 [all acoustical analyses are

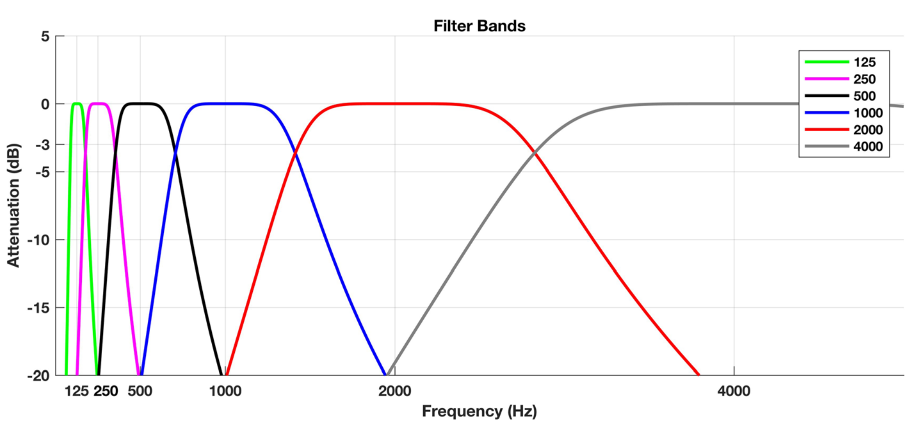

FIG. 2. (Color online) Six octave bands with center frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz used for acoustical analyses of simulated vowels with f_o modulation.

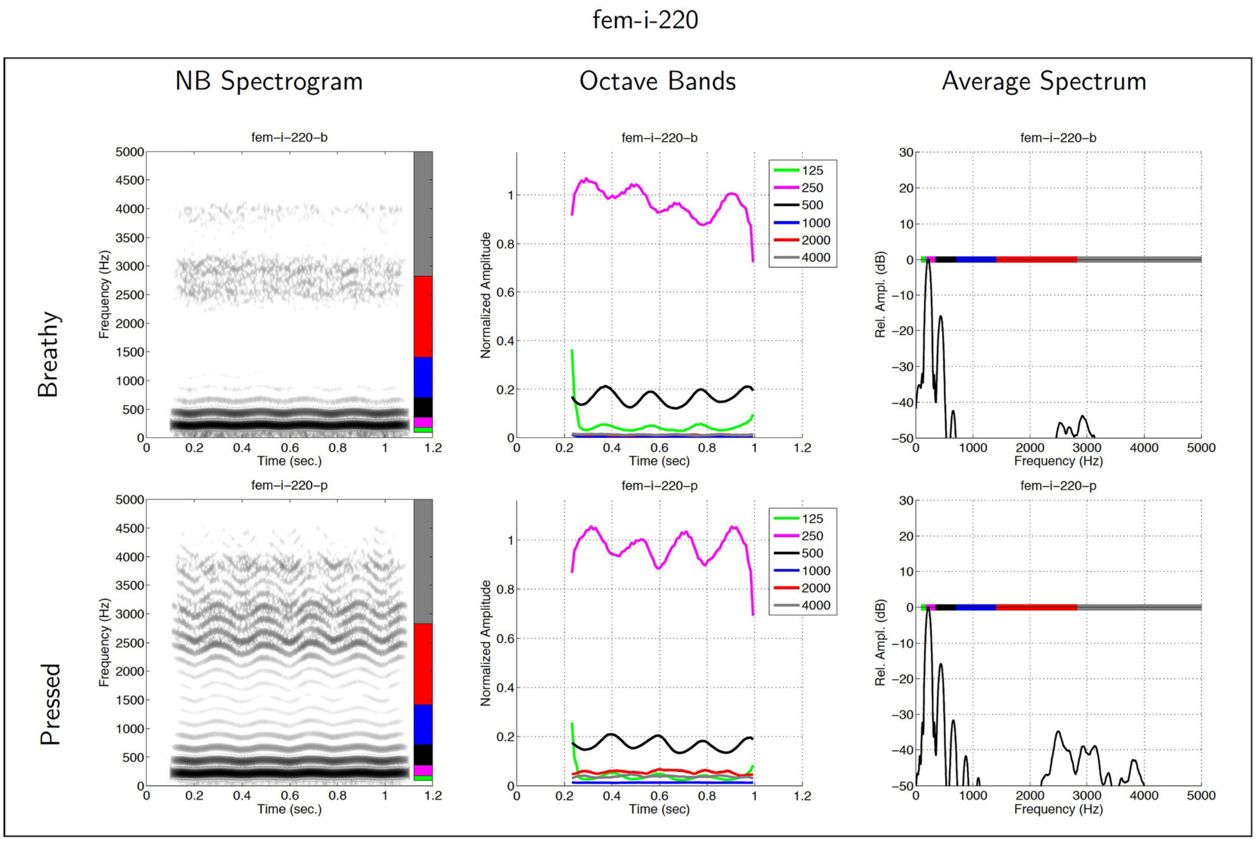

FIG. 3. (Color online) Narrowband spectrogram (left panel), octave band analyses (middle panel), and average spectrum (right panel) for the female /i/ with a target note of A3 (220 Hz) and a breathy vocal quality (upper row) and pressed vocal quality (lower row).

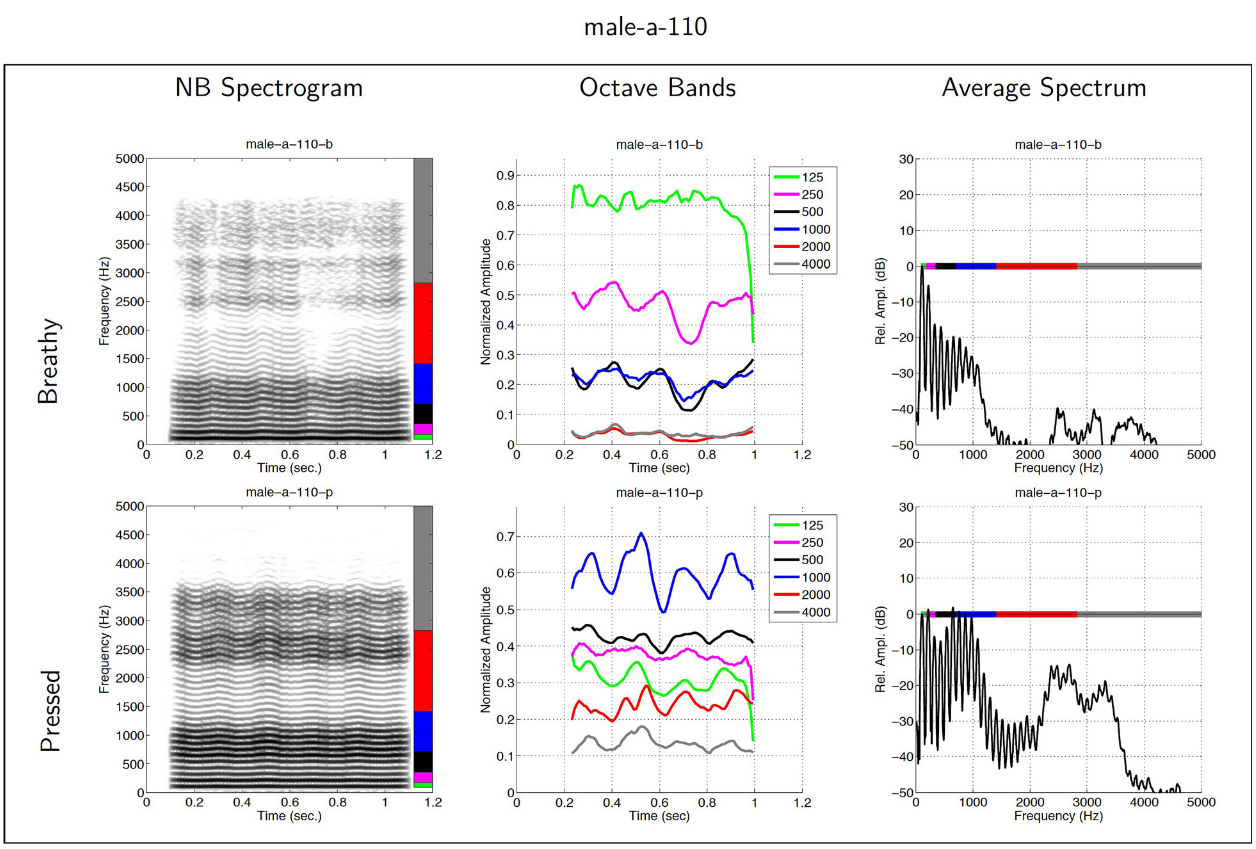

FIG. 4. (Color online) Narrowband spectrogram (left panel), octave band analyses (middle panel), and average spectrum (right panel) for the male "ah" with a target note of A2 (110 Hz) and a breathy vocal quality (upper row) and pressed vocal quality (lower row).

presented in Lester (2014)]. The results of the regression analyses are presented in Table I.

C. Discussion

The results of this study indicate that information in frequency bands across the spectrum predicted, to some extent, listeners' judgments about the magnitude of modulation of voice. This is consistent with the results of the previous study using a computational model. For female samples in the present study, the mean amplitude within lower frequency bands was used with f_o contrasts, and the mean amplitude within the lower and higher frequency bands was used with quality contrasts and with vowel contrasts. The coefficient of variation of amplitude of the lower and higher frequency bands was used with all three contrasts. For male voices, the mean amplitude and the coefficient of variation of amplitude of the lower and higher frequency bands were used with all contrasts. Although there were patterns in the variables used in the regression models, the goodness of fit of the models was only fair based on the R-square values.

The pattern in the perceptual findings of the current study unexpectedly differed from the results of the previous study. That is, in Lester and Story (2015), the pressed voice was perceived more often as being shakier than the breathy voice, and the vowel /i/ was perceived more often as being shakier than the /ɑ/ for both the female and male stimuli. In the current study, the pressed voice was significantly more often perceived as being shakier than the breathy voice only for the female stimuli. With the male stimuli, there was a trend for the breathy voice to be perceived as shakier than the pressed voice. In addition, there was a trend for the /ɑ/ to be perceived as shakier than the /i/ with both the female and male stimuli in this study. The differences in the direction of the patterns from the hypothesized direction led to non-significant findings for vocal quality with the male stimuli and for vowel with the male and female stimuli.

The differences in perception of the breathy vs pressed voices for the male and female singers might be due to the irregularity in the rate and extent of f_o modulation that can be seen in the male productions; whereas, the rate and extent of f_o modulation in the female productions appears to be quite regular. Listeners in this study were trained to make judgments about the magnitude of modulation of voice using stimuli with sinusoidal modulation of the f_o, which may have inadequately primed listeners to make judgments about voices with irregular rates and extents of f_o modulation. In addition, a previous study of the perception of vocal tremor revealed that the regularity of f_o modulation had an effect on listeners' perception of vocal tremor (Kreiman *et al.*, 2003).

The differences in the perception of the breathy vs pressed voices for the male and female singers might also be due to differences in the extent of f_o modulation between the female and male productions. That is, acoustical analyses revealed that the female's f_o modulation extent ranged between 3.0 and 5.6 Hz, whereas the male's f_o modulation extent ranged between 1.2 and 3.3 Hz. No differences were seen in the patterns of participants' judgments based on the extent of f_o modulation in Lester and Story (2015). However, a pattern might be have been revealed with a greater number of trials. It is unclear what might have caused the differences in perception of the magnitude of modulation of voice with the vowel contrasts in this study compared with the previous study. It is possible that the formant frequencies of the singers differed from the formant frequencies that were produced by the computational model, which based vocal tract shapes on spoken vowel productions.

Future studies should investigate the mean amplitude and coefficient of variation of the amplitude in frequency bands that are higher than the bands used in this study. Moore and Tan (2003) found that judgments of the "naturalness" of voices were significantly affected by high frequency energy between 7000 and 10 900 Hz. In addition, Liss *et al.* (2010) found that the 8000 Hz octave band was the best at predicting dysarthria type for different speakers with neuromotor diseases. Monson *et al.* (2014) have suggested that voice therapy to address vocal quality might be improved by understanding the contribution of high frequency energy (i.e., energy above 5000 Hz) to the perception of vocal quality. As mentioned in the previous study, a small percentage of f_o modulation results in a wide range of

TABLE I. Results of Study 2 binary logistic stepwise regression for each contrast (f_o, vocal quality, vowel) with potential predictor variables representing (1) the ratios of the mean amplitudes in each of the six octave bands for each of the stimulus pairs and (2) the ratios of the coefficients of variation of amplitude in each of the six octave bands (center frequencies: 125, 250, 500, 1000, 2000, and 4000 Hz) for each of the stimulus pairs for the female and male stimuli.

			125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	Cox and Snell R-square
Female	Mean amplitude	f_o	2	-	1	-	-	-	0.099, χ^2 (2, N = 320) = 33.414, p < 0.001
		Quality	2	3	-	-	1	-	0.386, χ^2 (3, N = 320) = 158.103, p < 0.001
		Vowel	1	2	-	-	-	3	0.103, χ^2 (3, N = 320) = 34.632, p < 0.001
	Coefficient of variation	f_o	-	2	-	-	-	1	0.112, χ^2 (1, N = 320) = 38.036, p < 0.001
		Quality	4	2	3	-	1	-	0.413, χ^2 (4, N = 320) = 170.723, p < 0.001
		Vowel	2	1	-	-	3	-	0.109, χ^2 (3, N = 320) = 37.029, p < 0.001
Male	Mean amplitude	f_o	-	2	-	-	1	-	0.336, χ^2 (2, N = 320) = 131.077, p < 0.001
		Quality	1	-	-	-	-	2	0.265, χ^2 (2, N = 320) = 98.428, p < 0.001
		Vowel	2	4	-	-	1	3	0.270, χ^2 (4, N = 320) = 100.610, p < 0.001
	Coefficient of variation	f_o	4	1	2	3	-	-	0.345, χ^2 (4, N = 320) = 136.036, p < 0.001
		Quality	3	1	2	-	-	-	0.310, χ^2 (3, N = 320) = 118.583, p < 0.001
		Vowel	3	2	1	4	6	5	0.288, χ^2 (6, N = 320) = 108.893, p < 0.001

frequency modulation in the higher harmonics. Therefore, it is possible that modulation of frequencies above 5000 Hz, which were not measured in the present study, could contribute most to the perception of the severity of vocal tremor.

D. Clinical applications

Based on the results of these studies with singers producing vibrato as a model of vocal tremor, it appears that production of a breathy voice could be effective in reducing the perception of the magnitude of modulation of voice for speakers with laryngeal vocal tremor. However, the effectiveness of this strategy might depend on the sex of the speaker, the regularity of the rate and extent of f_o modulation, or the extent of f_o modulation. These studies highlight the need for further perceptual and acoustical studies of vocal tremor to better understand the compensatory strategies that are effective in reducing the perception of the severity of vocal tremor and to identify the patients who would benefit most from using these compensatory strategies.

IV. CONCLUSIONS

Simulations of laryngeal vocal tremor using singers producing vibrato revealed that adjustments to the vocal quality affected the perception of magnitude of modulation in the predicted direction for female voices. The perceptual patterns differed for vocal quality based on the sex of the singer. The perceptual patterns for f_o and vowel contrasts differed in this study compared with the previous study using a computational model. The acoustical measures used in these studies did not fully account for listeners' perceptual judgments. Further studies are needed to determine if these differences in the perceptual results might be due to differences in the extent of f_o modulation or the regularity in the rate and extent of f_o modulation.

ACKNOWLEDGMENTS

This research was funded by National Institutes of Health F31 DC012697 and R01 DC011275. We would like to thank Andrew Lotto and Mark Borgstrom for their assistance with the statistical analyses for this study.

American Speech-Language-Hearing Association (**1997**). "Guidelines for audiological screening," Rockville, MD.

ANSI S1.1-1986 (ASA 65-1986). (**1993**). "Specifications for octave-band and fractional-octave-band analog and % digital filters" (American National Standards Institute, New York).

Appelman, D., and Smith, E. (**1985**). "Cinefluorographic and electromyographic observations of abdominal muscular function in its support of vibrato," in *Transcripts of the Fourteenth Symposium: Care of The Professional Voice, Part I: Scientific*, edited by V. L. Lawrence (The Voice Foundation, New York), pp. 79–82.

Barkmeier-Kraemer, J., Lato, A., and Wiley, K. (**2011**). "Development of a speech treatment program for a client with essential vocal tremor," Semin. Speech Lang. **32**(1), 43–57.

Bové, M., Daamen, N., Rosen, C., Wang, C. C., Sulica, L., and Gartner-Schmidt, J. (**2006**). "Development and validation of the vocal tremor scoring system," Laryngoscope **116**(9), 1662–1667.

Brown, J. R., and Simonson, J. (**1963**). "Organic voice tremor: A tremor of phonation," Neurology **13**(6), 520–525.

Dromey, C., Warrick, P., and Irish, J. (**2002**). "The influence of pitch and loudness changes on the acoustics of vocal tremor," J. Speech, Lang. Hear. Res. **45**(5), 879.

Farinella, K. A., Hixon, T. J., Hoit, J. D., Story, B. H., and Jones, P. A. (**2006**). "Listener perception of respiratory-induced voice tremor," Am. J. Speech-Lang. Pathol. **15**(1), 72–84.

Fourcin, A. J. (**1981**). "Laryngographic assessment of phonatory function," ASHA Rep. **11**, 116–127.

Hachinski, V. C., Thomsen, I. V., and Buch, N. H. (**1975**). "The nature of primary vocal tremor," Can. J. Neurol. Sci. **2**(3), 195–197.

Hillenbrand, J. M., and Gayvert, R. T. (**2005**). "Open source software for experiment design and control," J. Speech, Lang. Hear. Res. **48**(1), 45–60.

Koda, J., and Ludlow, C. L. (**1992**). "An evaluation of laryngeal muscle activation in patients with voice tremor," Otolaryngol.–Head Neck Surg. **107**(5), 684–696.

Kreiman, J., Gabelman, B., and Gerratt, B. R. (**2003**). "Perception of vocal tremor," J. Speech, Lang., Hear. Res. **46**(1), 203–214.

Lederle, A., Barkmeier-Kraemer, J., and Finnegan, E. (**2012**). "Perception of vocal tremor during sustained phonation compared with sentence context," J. Voice **26**(5), 668.e1–668.e9.

Lester, R. A. (**2014**). "Modulation of voice related to tremor and vibrato," Ph.D., University of Arizona, Tucson, AZ.

Lester, R. A., and Story, B. H. (**2013**). "Acoustic characteristics of simulated respiratory-induced vocal tremor," Am. J. Speech-Lang. Pathol. **22**(2), 205–211.

Lester, R. A., and Story, B. H. (**2015**). "The effects of physiological adjustments on the perceptual and acoustical characteristics of simulated laryngeal vocal tremor," J. Acoust. Soc. Am. **138**(2), 953–963.

Liljencrants, J. (**1985**). *Speech Synthesis with a Reflection-Type Line Analog* (Royal Institute of Technology, Department of Speech Communication and Music Acoustics, Stockholm, Sweden).

Liss, J. M., LeGendre, S., and Lotto, A. J. (**2010**). "Discriminating dysarthria type from envelope modulation spectra," J. Speech, Lang., Hear. Res. **53**(5), 1246–1255.

Ludlow, C. L., Bassich, C. J., Connor, N. P., and Coulter, D. C. (**1986**). "Phonatory characteristics of vocal fold tremor," J. Phon. **14**, 509–515.

Monson, B. B., Hunter, E. J., Lotto, A. J., and Story, B. H. (**2014**). "The perceptual significance of high-frequency energy in the human voice," Front. Psychol. **5**, 587.

Moore, B. C. J., and Tan, C.-T. (**2003**). "Perceived naturalness of spectrally distorted speech and music," J. Acoust. Soc. Am. **114**(1), 408–419.

Niimi, S., Horiguchi, S., and Kobayashi, N. (**1988a**). "The physiological role of the sternothyroid muscle in phonation. An electromyographic observation," Ann. Bull. Res. Inst. Logoped. Phoniatr. **22**, 163–168.

Niimi, S., Horiguchi, S., Kobayashi, N., and Yamada, M. (**1988b**). "Electromyographic study of vibrato and tremolo in singing," in *Voice Production, Mechanisms and Functions* (Raven, New York), pp. 403–414.

Ramig, L. A., and Shipp, T. (**1987**). "Comparative measures of vocal tremor and vocal vibrato," J. Voice **1**(2), 162–167.

Story, B. H. (**2005**). "Synergistic modes of vocal tract articulation for American English vowels," J. Acoust. Soc. Am. **118**(6), 3834–3859.

Sulica, L., and Louis, E. D. (**2010**). "Clinical characteristics of essential voice tremor: A study of 34 cases," Laryngoscope **120**(3), 516–528.

Titze, I. R. (**1984**). "Parameterization of the glottal area, glottal flow, and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570.

Titze, I. R. (**2006**). *The Myoelastic Aerodynamic Theory of Phonation* (National Center for Voice and Speech, Iowa City, IA), pp. 197–214.

Vennard, W., Hirano, M., and Ohala, J. (**1970**). "Laryngeal synergy in singing chest, head and falsetto," Natl. Assoc. Teachers Singing Bull. **27**, 16–21.

Warrick, P., Dromey, C., Irish, J. C., Durkin, L., Pakiam, A., and Lang, A. (**2000**). "Botulinum toxin for essential tremor of the voice with multiple anatomical sites of tremor: A crossover design study of unilateral versus bilateral injection," Laryngoscope **110**(8), 1366–1374.