Social and Stylistic Correlates of Vocal Fry in *a cappella* Performances

[a]Whitney Chappell, [b]John Nix, and [c]Mackenzie Parrott, [ab]*San Antonio, Texas*

Summary: Objective. To determine the social and stylistic correlates of vocal fry in *a cappella* performances.
Study Design. A matched-guise experiment was used to measure listener evaluations of fry and non-fry guises.
Methods. Four singers, two male and two female, sang "The Star-Spangled Banner" with onset vocal fry. These recordings were used to create the two guises: (i) an unmodified recording with onset vocal fry on vowel-initial words and (ii) a recording in which the fry had been removed. In total, 253 participants listened to the recordings and evaluated the singers' social and stylistic attributes along a Likert scale, e.g., how confident, sexy, and sincere each singer sounded. A factor analysis was used to conflate correlated variables, and mixed effects linear regression models ($n = 1{,}012$) were fitted to each lone or joint factor to determine whether vocal fry significantly influenced listeners' responses to the singers.
Results. Vocal fry significantly altered listener evaluations of the singers' sincerity/commitment, maturity/sophistication, naturalness, and confidence ($P < 0.05$). Unlike male singers, who were rated as significantly less sincere/committed with vocal fry, female singers were seen as more sincere/committed with vocal fry and younger listeners also found them less natural, suggesting vocal fry is associated with emotional intensity in female voices. Younger listeners perceived singers with fry as less mature/sophisticated, suggesting an association with youth. Finally, listeners with more musical training rated singers with fry as less confident, while less trained listeners did not exhibit this difference.
Conclusions. Listeners are highly attuned to vocal fry in music but respond to it differently based upon their age, musical training, and the singer's sex. Vocal fry is evaluated more positively among younger, less musically trained listeners, and it is better received in women's voices, suggesting that the use of fry strategically targets a specific audience, i.e., younger and less trained listeners, who interpret fry as a marker of youth and emotional earnestness. These findings show that a single stylistic feature like vocal fry can be imbued with multiple meanings depending on the singer and audience, and its use can serve to include or exclude particular listener groups.
Key Words: Vocal fry−Matched guise−Social correlates−Stylistic correlates−*a cappella*.

INTRODUCTION

Vocal fry has been recognized as a type of phonation with a very low fundamental frequency,[1] where each flow pulse decays completely prior to the next one commencing.[2] Fry is often distinguished from 'creaky voice,' which is found at higher frequencies and is characterized by a perception of roughness and the existence of subharmonics in the glottal waveform.[3]

A number of recent studies have examined the prevalence, use, and meaning of vocal fry in American English. In speech, vocal fry tends to occur at the ends of sentences and paragraphs, at word boundaries, syllable boundaries, and in unstressed syllables.[4-5] It also is more prevalent among female speakers than male speakers.[6-7] The use of fry in speech has been associated with a variety of social constructs, including toughness in Chicana gang girls;[8] masculinity, authority, and whiteness;[9] professionalism and upward mobility in young women;[10] and, conversely, evaluations of lower competence, education, trustworthiness, attractiveness, and hireability in young women, especially among older listeners.[11]

Vocal fry in Western singing styles has not been as widely explored. Pedagogical opinions regarding its utility in singing training are available,[12-21] but relatively little work has been done to study fry's use as a stylistic element in popular singing,[22] which has been observed in many American commercial singing genres, such as country, blues, jazz, pop, and R&B. Thompson[23] suggests that vocal fry is used as an embellishment to add color to certain notes, while also evoking intimacy, and Pecknold[24] argues that vocal fry is associated with the inherent in-betweenness of teenage girls' bodies and identities. Finally, in contrast to its previously mentioned phrase-final use in speech, the authors (a linguist and two singing teachers) have observed that fry in popular singing styles appears most commonly at phonation onset rather than at termination, especially on vowel-initial words and some voiced initial consonants, such as /l/, /m/, /n/, and /r/, in both male and female performances.

The use of fry as a stylism in singing led the authors to ask several questions: Why is this particular stylistic feature being widely used in so many different genres (i.e., what is driving its use)? What is the fry expressing? Is vocal fry an amplifying device for emotional content that is already present in the vocal performance? Ultimately, we narrowed the question to the following: What are the social and stylistic correlates of vocal fry in popular *a cappella* performances?

Accepted for publication June 8, 2018.
From the [a]The University of Texas at San Antonio, Texas; [b]The University of Texas at San Antonio, Texas; and the [c]Independent Scholar.
Address correspondence and reprint requests to Whitney Chappell, The University of Texas at San Antonio, Department of Modern Languages and Literatures, 1 UTSA Circle, San Antonio, Texas 78249. E-mail: whitney.chappell@utsa.edu
Journal of Voice, Vol. 34, No. 1, pp. 156.e5−156.e13
0892-1997

https://doi.org/10.1016/j.jvoice.2018.06.004

By exploring the question of what the fry might be expressing, we hoped to also answer the question of why it was being used. Drawing on prior work by two of the authors,[22] we hypothesized that older individuals and those with formal vocal training would view fry more negatively when listening to performances with and without fry. We also predicted that listeners would evaluate vocal fry in male and female voices differently.

MATERIALS AND METHODS

This study employs the matched-guise technique to indirectly measure listener attitudes toward vocal fry. The technique, which is widely used in sociolinguistics and social psychology, was first introduced by Wallace Lambert and his colleagues in 1960,[25] and it asks listeners to evaluate a series of voices along a matrix of properties, such as how attractive, intelligent, or likeable the speakers sound. Unbeknownst to the listeners, they hear the same speakers two times, differing only in a predefined feature. If the same speakers are given significantly different evaluations, it can be safely assumed that this is attributable to the difference between the guises, as all other aspects of the recordings are identical. In this case, listeners heard the same singers with and without onset vocal fry. Listeners' reactions to these guises are then interpreted as reactions to the vocal fry present or absent in the recordings, as outlined in more detail below.

Recordings and pilot study

To establish the matrix of social and stylistic properties to be included in the matched-guise test and to determine whether accompaniment or *a cappella* performances should be used, a pilot study was created before the large-scale study, using portions of songs from pop and country genres of different decades (1970s, 1980s, 1990s, 2000s, and 2010s) and portions of the traditional songs "Amazing Grace" and "The Star-Spangled Banner." The matched-guise method was chosen as an improvement upon a previous study undertaken by the second and third authors,[22] in which singers performed songs in as similar a fashion as possible in separate recordings with and without vocal fry. Four different singers (two male and two female) who self-reported as being in good health were used in making the recordings, which all featured the singers employing vocal fry at phonation onset on selected vowel-initial words (see the Appendix). Song samples were recorded in a WhisperRoom Noise Reduction booth (WhisperRoom, Inc., Knoxville, TN, USA), using a Shure SM48 dynamic microphone (Shure Incorporated, Niles, IL, USA), a KAY CSL 4500 pre-amplifier (Pentax Medical, Montvale, NJ, USA), and a Dell Optiplex GX520 computer (Dell Corp, Round Rock, TX, USA). The recordings were made at a sampling rate of 44.1K per second as 16-bit mono files using Goldwave 5.70 recording and editing software (GoldWave, Inc., Saint John's, NF, Canada). For the popular and country songs, the singers sang along to karaoke accompaniment tracks (iTunes, Apple, Inc., Cupertino, CA, USA) that were played on a MacBook Air laptop computer (Apple Inc., Cupertino, CA, USA) and listened to through Beats Audio 190003-00 noise cancelling headphones (Beats Electronics LLC, Cupertino, CA, USA). The two traditional songs were recorded *a cappella*.

Post recording, two versions of each song selection were created, one with the vocal fry intact, and the other with the vocal fry edited out, using the Goldwave 5.70 editing software. The pop and country vocal recordings were then synchronized with the karaoke accompaniment tracks using Audacity 1.2.4b (open-source software; Carnegie-Mellon University, Pittsburgh, PA, USA; www.audacityteam.org). The unaccompanied selections went through the same editing procedure regarding the vocal fry.

Listeners ($n = 141$) were randomly assigned to one of four surveys, each of which contained specific recordings with and without vocal fry. More specifically, Pilot Surveys A and B included the pop and country selections with accompaniment, while Pilot Surveys C and D only included *a cappella* traditional selections. After listening to each recording, participants were asked to evaluate the singer along an extensive list of social and stylistic properties, including how feminine or masculine, assertive, sexy, sincere, mature, sophisticated, emotionally committed, cool, educated, natural, playful, pained, lazy, and vulnerable each singer sounded. Listeners were also asked to identify the singer's perceived age and race.

The pilot study resulted in two important methodological findings. First, comments left by respondents to the surveys indicated that the accompaniment tracks for the popular and country selections may have distracted the listeners from making accurate ratings of the singers' performances. In order to remove other variables and distractions from the primary listener task, for the final version of the survey only unaccompanied performances would be used. Second, the authors were able to limit the number of descriptors used in the main experiment based on participants' reactions in the pilot study. A factor analysis[27] showed that several properties were evaluated similarly in the pilot study, e.g., singers rated as less mature were also perceived as younger. As a result, the authors removed overlapping properties or factors that showed very little variation across singers like perceived education level. While listeners seem to be highly attuned to a speaker's status in speech,[10] the participants' responses in the pilot study showed that status-based factors were not as readily perceived in song.

Main study experimental design

After the pilot study helped reduce the song selection and descriptors to be included in the main study, two songs were chosen for the final experiment: "Amazing Grace" and "The Star-Spangled Banner," both of which were recorded in the pilot study without accompaniment. As in the pilot study, these baseline recordings with vocal fry were used to create the two guises: (i) an unmodified recording with onset vocal fry and (ii) a recording in which the onset fry had been

removed, employing the editing procedure used in the pilot study. As a result, eight baseline audio files and 16 total guises were created to be evaluated by listeners.

The final recordings were uploaded to the online survey platform SurveyGizmo.[26] To be eligible for the experiment, participants had to be 18 or older, native speakers of American English, born and raised in the United States, wearing headphones, and without any hearing loss. Participants confirmed that they met these criteria and consented to participate in the research study (UTSA IRB#16-075E). Once consent was given, participants answered basic demographic and musical training questions, including their sex, age, sexual orientation, ethnicity, musical genres listened to (e.g., classical, pop, jazz, country, and rap/hip-hop), and musical training (no formal training, participation in ensembles, or individual instruction). For those who had participated in ensembles or had received individual instruction, however, no questions were asked regarding the type of instrument(s) played or vocal parts sung, nor were there questions about the length of participation/training.

Next, participants listened to two practice audio files to familiarize themselves with the task at handbefore listening to the target audio files. Along with each audio file, which played automatically, participants were asked to evaluate how the singer sounded given a six-point scale (1 to 6) of social and stylistic attributes, chosen given their significance and independence in the pilot study. These scales included less feminine/less masculine to more feminine/more masculine, less confident to more confident, less sexy to more sexy, less sincere to more sincere, less mature to more mature, less emotionally committed to more emotionally committed, less cool to more cool, less sophisticated to more sophisticated, less natural sounding to more natural sounding, and less vulnerable to more vulnerable. Listeners were also asked about perceived singer ethnicity (Hispanic, African American, Caucasian, or Other/Not sure), and an optional comment box was provided for additional thoughts on the singer and his/her performance.

To reduce the duration of the experiment with the goal of minimizing participant attrition, participants were randomly assigned to one of two branches, Branch A or Branch B, each of which contained half of the target stimuli. A pseudorandom presentation order of the audio files was employed to maximally separate individual singers, songs, male and female voices, and guises (fry and non-fry). The full list of stimuli in each branch is shown in the Appendix. Figure 1 below presents a screenshot of the evaluation matrix completed by the participants for each audio file.

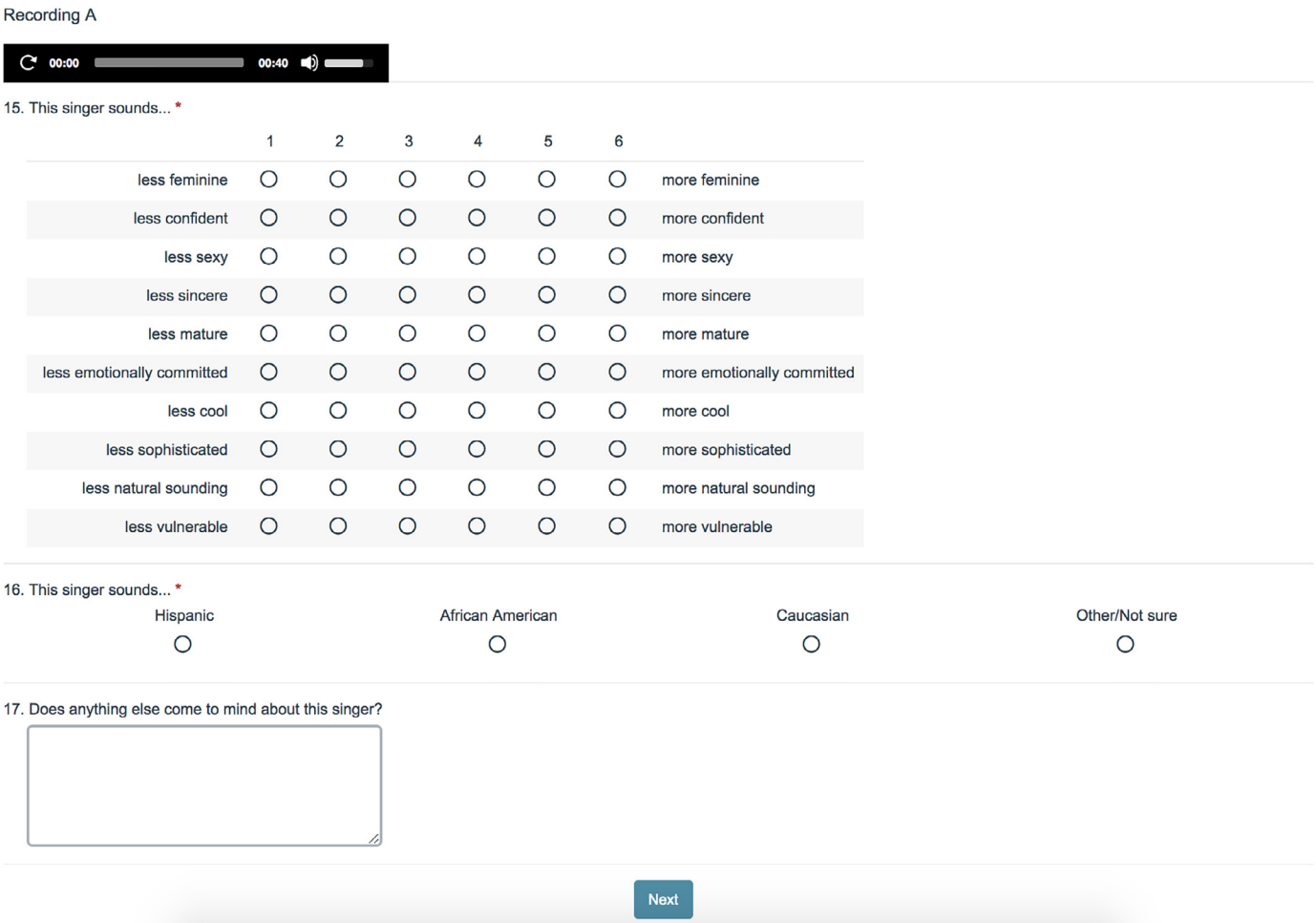

FIGURE 1. Evaluations matrix screenshot for Recording A.

TABLE 1.
Demographic Information About the Participants

Mean/median age (years)	29/23
Age range (years)	18-77
Female : Male : Other (*n*)	169 : 82 : 2
Branch (A : B)	127 : 126
Musical training (*n*)	Individual instruction (150)
	Ensembles (43)
	No musical training (60)
Race (*n*)	Caucasian (129)
	Hispanic (79)
	African American (17)
	Asian (16)
	Biracial (4)
	Other (8)
State of provenance (*n*)	Texas (176)
	Illinois (8)
	Utah (7)
	California (6)
	New York (6)
	Washington (5)
	Other (45)
Total participants (*n*)	253

Participants

Participants were recruited both through social media platforms and in music appreciation and dance classes at a public university in Texas. In total, 253 listeners participated in the study. More detailed demographic information is provided about the participants in Table 1.

Statistical analysis

For ease of interpretation, listeners' ratings of singers' social and stylistic attributes were centered, making 0 the new midpoint. In other words, in the figures in the Results section, positive numbers represent listener's evaluations that are above the midpoint of the scale, and negative numbers represent listener's evaluations that are below the midpoint. Before mixed effects regression models were fitted to the attributes evaluated by listeners, a factor analysis using R's factanal function[27] was conducted to determine if any correlations existed between the attributes. For example, if singers who are evaluated as more sincere are also evaluated as more committed, these two scales can be combined for the statistical analysis. Using the Kaiser rule and a scree plot as a visual aid, three factors were found to account for the majority of variation in the dataset: a joint factor loading for both sincerity and commitment, a lone factor for coolness, and a joint factor for sophistication/maturity. As a result, evaluations of (i) how sincere and committed and (ii) how sophisticated and mature the singers sounded were conflated in the analysis below, and all other factors were analyzed independently.

To determine what independent variables significantly impacted listeners' evaluations of each social or stylistic property, mixed effects linear models were then fitted to each social and stylistic attribute using the lmer function in the package lme4.[28] Originally, models were constructed using all listener evaluations, but a more detailed exploration of the data showed that listeners exhibited a wider range of evaluations for the "Star-Spangled Banner" than "Amazing Grace", perhaps due to American listeners' stronger opinions about the appropriate performance practice of this piece. As a result, only listener evaluations of the "Star-Spangled Banner" performance will be analyzed below. In each model, evaluations of a joint (e.g., sincere/committed) or lone (e.g., cool) factor serve as the dependent variable, and the independent variables were added using a stepwise procedure. After each independent variable was added, the analysis of variance function[27] was used to establish if the variable significantly improved the model. Fixed effects and interactions between fixed effects were tested and kept in the best-fit model if they significantly improved the overall fit.

Treatment contrasts were applied to the categorical fixed effects, which means that the levels of a categorical independent variable were all compared to a single base or reference level. These independent variables include singer sex (reference level = female), listener sex (reference level = female), musical training (reference level = none), guise heard (reference level = without vocal fry), and whether the participant listed to (i) classical, (ii) pop, (iii) jazz, (iv) country, (v) rap/hip-hop, (vi) oldies, (vii) rock, (viii) R&B, (ix) Latin, and (x) other (reference level for all genres = no). Listener age was included as a continuous independent variable, and presentation order (reference level = without vocal fry heard first) was also tested to determine if the order in which the guises were heard influenced listener evaluations. Finally, by-listener and by-participant random intercepts and by-vocal fry random slopes were included as random effects, in accordance with the maximal random effects structure justified by the experiment's design.[29]

RESULTS

Vocal fry emerged as a significant predictor of listener evaluations in four models: sincerity/commitment, maturity/sophistication, naturalness, and confidence. In other words, listeners' evaluations of the aforementioned criteria changed significantly when they heard the same audio files with and without fry. The results of the best-fit models for these factors are presented in Tables 2 to 5 and include the estimate, standard error (SE), *t*-value, and *P*-value. The alpha value for each independent variable is 0.05, and lower values are considered significant. It should be noted that the models in which vocal fry was not a significant predictor of listener evaluations will not be discussed further in this paper, as the research question revolves around the interpretation of vocal fry in song.

For the categorical independent variables in Tables 2 to 5, e.g., musical training, a positive estimate indicates higher

evaluations for the given level than the reference level to which it is compared, and a negative estimate indicates the opposite, that is, that the given level received lower evaluations than the reference level. For continuous variables like age, a negative estimate indicates that evaluations decrease as age increases, and a positive estimate suggests that listener evaluations increase as age increases. An interaction means two independent variables work together to condition listener evaluations; the relationship between the variables is clarified in the text and in graphical format in Figures 2 to 5. When interactions occur, the lower-level main effects are kept in the model but should be interpreted on their own with caution, which is why they are presented between parentheses.

Sincerity/Commitment

In the best-fit model for the joint factor of singer sincerity/commitment, both listener age and the interaction between vocal fry and singer sex significantly affected listener evaluations. Of primary importance, Table 2 and Figure 2 show that the presence of vocal fry made the male singers seem less sincere and committed to the listeners, while the female singers were perceived as more sincere and committed. Additionally, as listener age increased, evaluations of singers' sincerity/commitment decreased. The interaction plot in Figure 2 shows the mean evaluation of sincerity/commitment with and without fry for the male and female singers, with standard error bars provided.

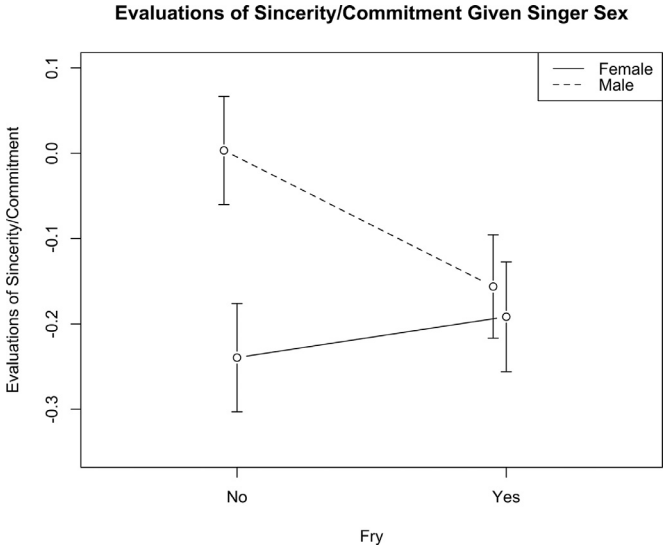

FIGURE 2. Evaluations of sincerity/commitment given singer sex.

TABLE 3.
Best-fit Mixed Effects Model for Evaluations of Mature/Sophisticated (Significant Predictors in Bold, *n* = 1,012)

	Estimate	*SE*	*t*-value	*P*-value
(Intercept)	4.459	0.161	27.712	<0.001
Age	**−0.013**	**0.003**	**−4.048**	**<0.001**
Vocal fry = yes	−0.211	0.056	−3.766	<0.001
Musical training				
Training = individual instruction	**−0.224**	**0.11**	**−2.032**	**<0.044**
Training = ensembles	−0.044	0.141	−0.312	0.755
	AIC: 2918			

Maturity/Sophistication

The second best-fit model, this time fitted to evaluations of maturity/sophistication, showed that singers were perceived as being less mature/sophisticated when they produced fry, and the most critical listeners were the most trained musicians, who rated the singers as less mature and sophisticated than other listeners (Table 3). Similarly, older listeners found the singers to be less mature and sophisticated overall than the younger listeners. Figure 3 provides a conditional inference tree, a classification technique based on binary recursive partitioning that first establishes which independent variable has the strongest association to the dependent variable and then divides the data in subsets, showing the significant predictors within those subsets. The boxplots below the conditional inference tree present the middle 50% of listener evaluations, and the horizontal line through each box indicates the median. The conditional inference tree shows a significant split in evaluations for listeners below 22 and those over 22. For the youngest listeners, vocal fry was associated with less maturity/sophistication, while for the

TABLE 2.
Best-fit Mixed Effects Model for Evaluations of Sincere/Committed (Significant Predictors in Bold, *n* = 1,012)

	Estimate	*SE*	*t*-value	*P*-value
(Intercept)	4.208	0.208	20.185	<0.001
Age	**−0.007**	**0.003**	**−2.237**	**0.026**
Interaction between vocal fry and singer sex				
Vocal fry = yes: Singer sex = male	**−0.232**	**0.113**	**−2.048**	**0.041**
(Vocal fry = yes)	0.052	0.083	0.628	0.53
(Singer sex = male)	0.271	0.254	1.067	0.39
	AIC: 3002			

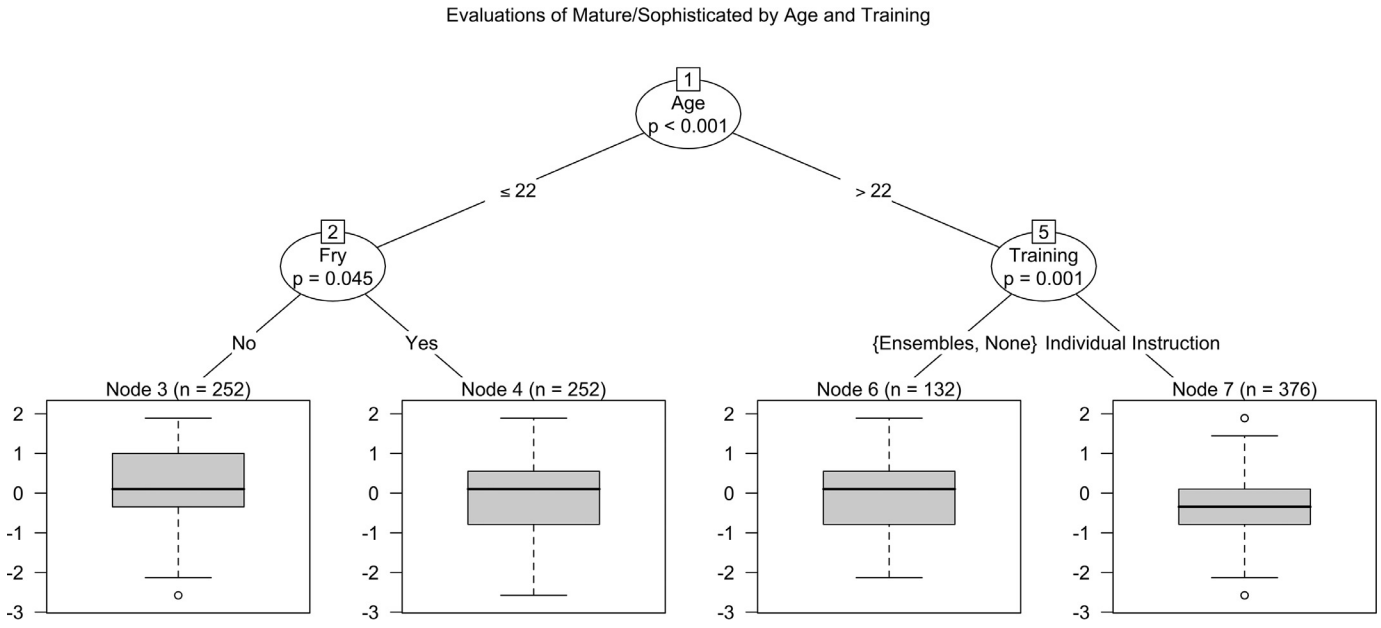

FIGURE 3. Evaluations of mature/sophisticated by age and training.

older listeners, the most trained listeners gave the lowest evaluations overall, but vocal fry did not significantly alter their evaluations.

Naturalness

The third best-fit model for naturalness shows that singer sex, vocal fry, and age all significantly conditioned listener evaluations, as is shown in Table 4 and Figure 4. Male singers were rated as more natural, older listeners found the singers to be less natural than younger listeners, and the vocal fry guises were rated as less natural than the non-vocal fry guises. The boxplot in Figure 4 shows that older listeners provided lower evaluations of naturalness overall, while younger listeners in particular found fry to be indicative of less natural performances, particularly in women's voices.

Confidence

Finally, the best-fit model for confidence shows an interaction between vocal fry and training. While listeners with no musical training provided the same evaluations of confidence for the guises with and without vocal fry, listeners with individual instruction decreased evaluations of confidence dramatically when presented with the vocal fry guises. Listeners with ensemble experience behaved similarly to the most trained listeners (Table 5 and Figure 5). Again, as age increased, evaluations of confidence decreased overall.

TABLE 4.
Best-fit Mixed Effects Model for Evaluations of Natural (Significant Predictors in Bold, *n* = 1,012)

	Estimate	*SE*	*t*-value	*P*-value
(Intercept)	0.113	0.102	1.106	0.27
Singer sex = male	**0.174**	**0.051**	**3.421**	**<0.001**
Vocal fry = yes	**−0.143**	**0.052**	**−2.759**	**0.006**
Age	**−0.008**	**0.003**	**−2.567**	**<0.011**
	AIC: 2742			

DISCUSSION

The results of this study show a consistent trend. Overall, vocal fry in singing is perceived negatively by the participants, a consensus reflected in listeners' lower evaluations of the singers' sincerity/commitment, maturity/sophistication, naturalness, and confidence in the guises with vocal fry. However, this broad generalization disguises more nuanced trends in the dataset, which suggest that vocal fry seems to mean two different things to listeners, depending on the age and musical training of the audience.

First, older and more musically trained listeners tend to evaluate singers differently than younger and less musically trained listeners, providing lower and more critical evaluations across guises. Unlike the older listeners, younger listeners perceive fry guises as indicative of lower levels of maturity and sophistication as compared to non-fry guises, which may signal an association with youthfulness or playfulness in the minds of younger listeners to which older listeners are less attuned. These conclusions support the qualitative observations of Thompson,[23] who proposed that vocal fry could add color and intimacy to vocal performances. Importantly, the present study demonstrates that more nuanced and positive interpretations of vocal fry in music are only accessible to younger and less musically trained listeners, creating a performative quality and stylism tailored exclusively to a younger audience.

Secondly, the perception of vocal fry appears to hinge on singer gender, as vocal fry made female singers sound

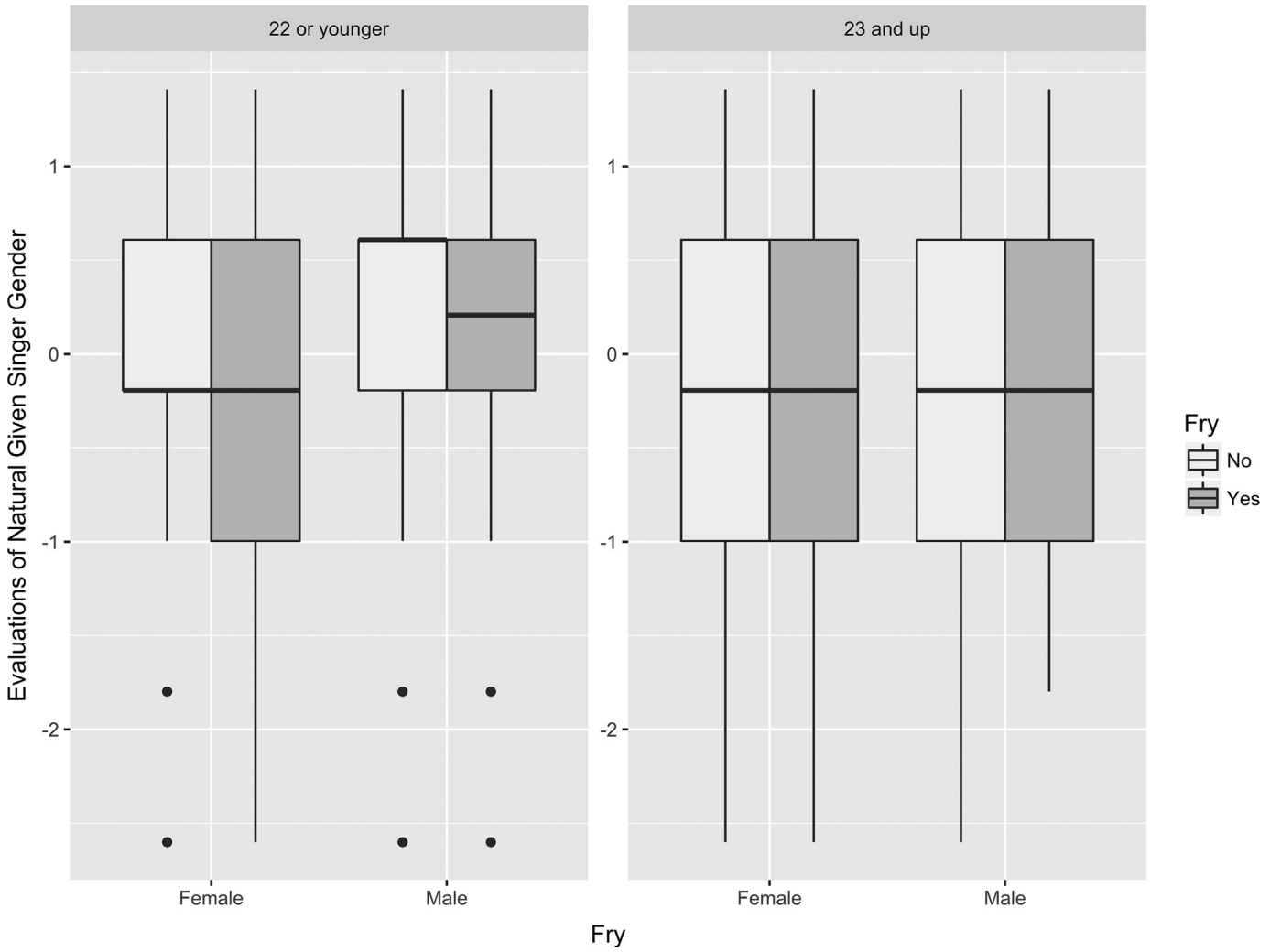

FIGURE 4. Evaluations of naturalness with respect to singer gender and listener age.

more sincere and committed, while male singers were viewed as less sincere and committed with the vocal fry guise. This result supports the observations of Pecknold,[24] who argued that vocal fry in music tends to be associated with teenage girls' bodies and identities. In other words, onset vocal fry as a stylism indicative of emotional commitment seems to be better received in female voices, perhaps because it is more expected in young women's speech.[6-7,10-11]

While future studies are needed to determine how singers, in particular, perceive vocal fry, it is likely that vocalists are capitalizing on these listener responses in their performances

TABLE 5.
Best-fit Mixed Effects Model for Evaluations of Confident (Significant Predictors in Bold, *n* = 1,012)

	Estimate	*SE*	*t*-value	*P*-value
(Intercept)	0.26	0.174	1.49	0.176
Age	**−0.006**	**0.003**	**−2.302**	**0.022**
Interaction between vocal fry and training				
Vocal fry = yes*training = individual instruction	**−0.311**	**0.134**	**−2.328**	**<0.021**
Vocal fry = yes*training = ensembles	−0.267	0.175	−1.528	0.128
(Vocal fry = yes)	0.035	0.113	0.311	0.756
(Training = individual instruction)	0.058	0.107	0.545	0.586
(Training = ensembles)	0.076	0.138	0.551	0.582
	AIC: 2766			

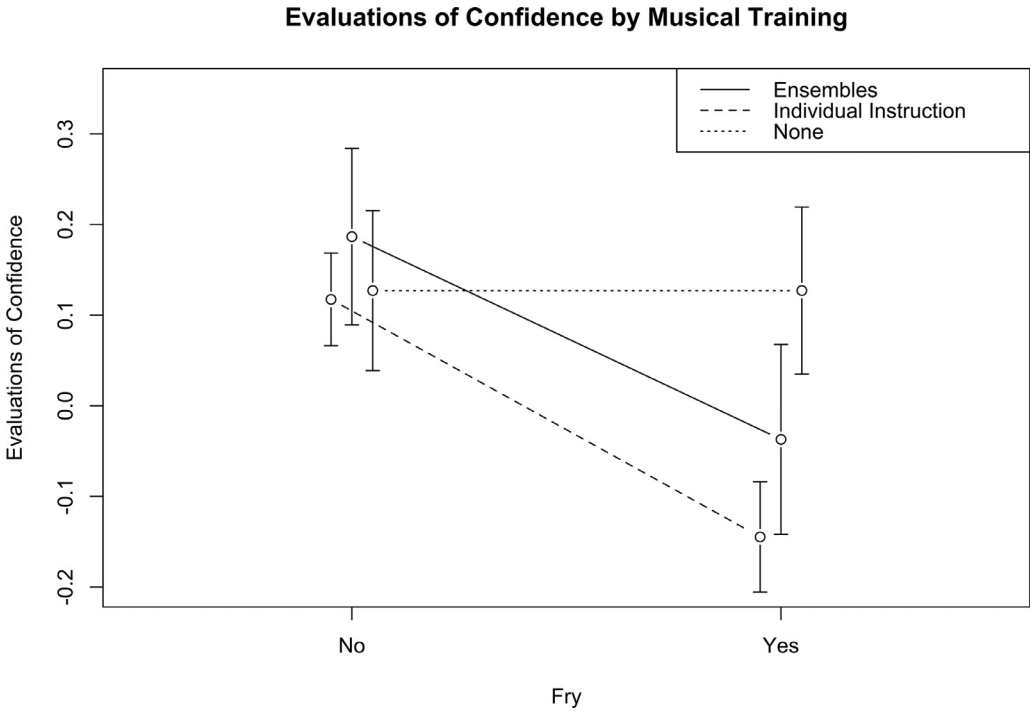

FIGURE 5. Evaluations of confidence with respect to listener musical training.

to more successfully engage their target audience. Especially for female pop singers, if vocal fry can increase listener perceptions of sincerity and commitment, on the one hand, and signal a youthful playfulness on the other, the performance is likely to resonate with the intended audience: young listeners without much musical training. In addition, female pop singers may choose to employ vocal fry to exclude certain listener groups who have more negative perceptions of fry, namely older and more musically trained listeners. As a result, vocal fry may simultaneously build in-group solidarity among younger listeners and disinvite parents and professionals from participating in the genre. This is likely why vocal fry is prevalent in popular music and among female singers.[24]

Before concluding, two potential caveats should be pointed out: first, the surveys' demographic questions did not ask the listeners who had participated in ensembles or had received individual instruction in music whether their musical training had been classical or contemporary in emphasis, nor did the survey ask listeners what type of instrument(s) they had played (e.g., voice, percussion, brass, woodwind, strings, keyboard). One could speculate that listeners who are professional voice teachers and speech pathologists whose training emphasized the classical music domain might rate a stylism like vocal fry, which is commonly associated with popular styles, in a less favorable light. This is certainly a follow-up question the researchers will ask in future studies. Second, it is possible that those with an instrumental (as opposed to a vocal) background might have had less exposure to "Amazing Grace" than "The Star-Spangled Banner." A difference in exposure to the pieces might have had an effect on the ratings. In future studies, the investigators will be sure to include a question about the listeners' familiarity with the songs.

CONCLUSION

This study has found that listeners are highly attuned to vocal fry in music but respond to it differently based on their age, musical training, and the singer's sex. Vocal fry is evaluated more positively among younger, less musically trained listeners and is better received in women's voices, suggesting that the use of fry strategically targets a specific audience—younger and less trained listeners—who interpret fry as a marker of youth and emotional earnestness. We conclude that vocal fry as a stylism in *a cappella* performances is imbued with a multiplicity of meanings that depend on both the singer and audience.

APPENDIX

Singer information.

Female Singer 1: 27 years old, Caucasian, Master's Degree in Music

Male Singer 1: 25 years old, African American, Bachelor of Arts student in Theater

Female Singer 2: 28 years old, Hispanic, Bachelor of Arts Degree in Voice

Male Singer 2: 24 years old, Caucasian, Bachelor of Music Education Degree

Organization of the online survey. Target recordings bolded.

Training Activity (all participants heard the same practice recordings)

1. Practice Female Singer, "The Star-Spangled Banner", without fry
2. Practice Male Singer, "Amazing Grace", without fry

Branch A (50% of participants directed to Branch A after the training)

1. Female Singer 1, "Amazing Grace", without fry
2. **Male Singer 1, "The Star-Spangled Banner", with fry**
3. Female Singer 2, "Amazing Grace", with fry
4. **Male Singer 2, "The Star-Spangled Banner", without fry**
5. **Female Singer 1, "The Star-Spangled Banner", with fry**
6. Male Singer 1, "Amazing Grace", without fry
7. **Female Singer 2, "The Star-Spangled Banner", without fry**
8. Male Singer 2, "Amazing Grace", with fry

Branch B (50% of participants directed to Branch B after the training)

1. Female Singer 1, "Amazing Grace", with fry
2. **Male Singer 1, "The Star-Spangled Banner", without fry**
3. Female Singer 2, "Amazing Grace", without fry
4. **Male Singer 2, "The Star-Spangled Banner", with fry**
5. **Female Singer 1, "The Star-Spangled Banner", without fry**
6. Male Singer 1, "Amazing Grace", with fry
7. **Female Singer 2, "The Star-Spangled Banner", with fry**
8. Male Singer 2, "Amazing Grace", without fry

Places where onset vocal fry was used for each song are indicated in bold.

"Amazing Grace"
Amazing Grace, how sweet the sound,
That saved **a** wretch like me!
I once was lost, but now am found,
Was blind, but now **I** see.

"The Star-Spangled Banner"
Oh say can you see,
By the dawn's early light,
What so proudly we hailed
At the twilight's last gleaming.

REFERENCES

1. Hollien H, Michel JF. Vocal fry as a phonational register. *J Speech Hear Res*. 1968;11:600–604.
2. Hollien H, Wendahl RW. Perceptual study of vocal fry. *J. Acoust. Soc. Amer*. 1968;43:509.
3. Titze IR. *Principles of Voice Production*1994 Prentice-Hall, Englewood Cliffs, NJ; 1994.
4. Wolk L, Abdelli-Beruh NB, Slavin D. Habitual use of vocal fry in young adult female speakers. *J. Voice*. 2011;26:e111–e116.
5. Gibson T. The role of lexical stress on the use of vocal fry in young adult female speakers. *J. Voice*. 2016;31:62–66.
6. Abdelli-Beruh NB, Wolk L, Slavin D. Prevalence of vocal fry in young adult male American English speakers. *J. Voice*. 2013;28:185–190.
7. Podesva RJ. Gender and the social meaning of non-modal phonation types. In: *Proceedings of the 37th Annual Meeting of the Berkeley Linguistic Society*. 427–448.
8. Mendoza-Denton N. The semiotic hitchhiker's guide to creaky voice: circulation and gendered hardcore in a Chicana/o gang persona. *J. Linguist. Anthropol*. 2011;21:261–280.
9. Lefkowitz D, Sicoli M. Creaky voice: constructions of gender and authority in American English conversation. *Paper Presented at 106th Meeting of the American Anthropology Association*. Washington DC.
10. Yuasa IP. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women. *Am. Speech*. 2010;85:315–337.
11. Anderson RC, Klofstad CA, Mayew WJ, et al. Vocal fry may undermine the success of young women in the labor market. *PLoS ONE*. 2014;9:e97506.
12. Vennard W. *Singing: The Mechanism and The Technic*. rev. ed. Carl Fischer, New York; 1967.
13. Miller R. *The Structure of Singing*. Schirmer, New York; 1996:125–126.
14. Brown O. *Discover Your Voice: How To Develop Healthy Voice Habits*. Singular, San Diego; 1996.
15. Brown O. *Discover Your Voice: Vocal Exercises by Oren L.* (Audio CD). Singular, San Diego; 1996: Track 9.
16. Riggs S. Singing for the Stars: A Complete Program for Training Your Voice. In: J.D. Carratello. Alfred Publishing, Van Nuys, CA; 1998:58–63.
17. Riggs S. *Singing For the Stars (Audio CD 2)*. Alfred Publishing, Van Nuys, CA; 1998. Tracks 7–22.
18. Nix J, Emerich K, Titze IR. Application of vocal fry to the training of singers. *J. Singing*. 2005;62:53–59.
19. Chapman JL. *Singing and Teaching: A Holistic Approach to Classical Voice*. Plural, San Diego; 2011. 284.
20. Titze IR, Verdolini-Abbott K. *Vocology: The Science and Practice of Voice Habilitation*. NCVS, Salt Lake City, UT; 2012. 268–269, 273.
21. Lunte R. Vocal fry and onsets for singing. http://www.themodernvocalistworld.com/articles.html/article/vocal-fry-onsets-for-singing-r1154/. Accessed February 22, 2018.
22. Parrott ML, Nix JP. Listener ratings of singer expressivity in musical performance; *Proc. Meet. Acoust*. 2016;26:035003 https://doi.org/10.1121/2.0000237.
23. Thompson M. Creaking, growling: feminine noisiness and vocal fry in the music of Joan La Barbara and Runhild Gammelsñter. *N. Paradoxa*. 2016;37:5–11.
24. Pecknold D. "These stupid little sounds in her voice": valuing and vilifying the new girl voice. In: Warwick J, Adrian A, (Eds.), *Voicing Girlhood in Popular Music: Performance, Authority, Authenticity*. Routledge, New York; 2016, pp. 77–98.
25. Lambert WE, Hodgson RC, Gardner RC. Evaluational reactions to spoken language. *J. Abnorm. Soc. Psych*. 1960;60:44–51.
26. Vanek C, McDaniel S. *Surveygizmo*. Widgix, Boulder CO; 2006.
27. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna; 2018.
28. Bates D, Maechler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. *J. Stat. Softw*. 2015;67:1–48.
29. Barr DJ, Levy R, Scheepers C. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang*. 2013;68:255–278.