

Judgments of Intelligence and Likability of Young Adult Female Speakers of American English: The Influence of Vocal Fry and the Surrounding Acoustic-Prosodic Context

Michelle A. Parker and Stephanie A. Borrie, Logan, Utah

Summary: Objective. Vocal fry is a prevalent speech feature in college-aged American women living in the United States. However, there is currently little consensus about how its use influences listener judgments of the speaker. This study investigated how vocal fry influences judgments of intelligence and the likability of young adult female speakers of American English while taking into account the surrounding acoustic-prosodic context, specifically voice pitch and speech rate.

Method. Speech samples were obtained from eight American English-speaking females who presented with different combinations of voice pitch (low or high), speech rate (slow or fast), and vocal fry (presence or absence). Listener judgments of ratings of intelligence and likability were collected from 463 adults via online crowdsourcing.

Results. Generalized estimating equation models revealed significant three-way interactions between the voice pitch, speech rate, and vocal fry for listener judgments of intelligence and likability. While vocal fry had favorable effects in some contexts (eg, high pitch, fast rate) it had unfavorable effects in others (eg, low pitch, fast rate).

Conclusion. Listener judgments of young American women based on information afforded in their speech are not solely based on the presence or absence of vocal fry, but rather a combination of features that interact with one another in unique ways. Thus, whether or not the use of vocal fry in this population projects a favorable impression depends on the acoustic-prosodic context in which it is produced.

Key Words: Vocal fry–Speech rate–Voice pitch–Listener judgments–Speaker characteristics.

INTRODUCTION

Vocal fry describes a salient speech feature that is linked to the quality of a speaker's voice. Perceptually, vocal fry has been depicted as creaky sounding or popping^{1,2} and is likened to the imitated sound of a "motor boat engine."³ During the production of vocal fry, the vocal folds vibrate in a unique dicrotic pattern, whereby relatively short glottal pulses (with low amplitude) are ensued by longer phases of vocal fold adduction.⁴ Simply, the vocal folds vibrate at a slow enough rate that single vibrations are discernable.⁵ Acoustically, vocal fry is characterized by very low frequencies, (around 20–70 Hz) regardless of speaker gender, and increased measures of jitter and shimmer.³ Although once regarded as a voice pathology, today vocal fry is accepted as the lowest end of the normal pitch range⁶ and one of three normal phonational registers.⁷

A number of studies have reported the prevalence of vocal fry in young adult female speakers of American English and speculated that the use of this speech feature is increasing (eg, Wolk et al,⁸ Yuasa,⁹ and Anderson et al¹⁰). Wolk et al⁸ examined speech samples from 34 female American English speakers (18–25 years) and identified the use of vocal fry in over two thirds of the participants. A follow-up study with young adult male speakers of

American English reported that vocal fry was four times less prevalent in male speakers when compared with the previously reported female data.¹¹ In an investigation of the influence of gender and nationality on vocal fry use in college-aged students engaged in conversation, young American female speakers were observed to use significantly more vocal fry than male and Japanese speakers.⁹ Additionally, Yuasa⁹ surveyed 175 college students (18–34 years) in the United States (Northern California and Iowa) and reported that approximately 80% of the students agreed with the statement that they heard vocal fry frequently used by women in their residential area. As Borrie and Delfino¹² have previously pointed out, there is currently no empirical evidence that the use of vocal fry in young American women is increasing; however, there is certainly robust evidence that it is prevalent in the speech behaviors of this population.

Given that initial impressions, based on the speaker's speech features, have been shown to influence social interaction, including mate selection, leader selection, and consumer choices,^{13–15} the high prevalence of vocal fry in speech behaviors of young adult female speakers of American English has raised enquiries as to how its presence may influence listener impressions. To date, the evidence in this area is largely equivocal. Some studies report that vocal fry is associated with favorable impressions of the female speaker. In the survey of 175 college students in Northern California and Iowa, Yuasa⁹ also asked the listeners to judge the speech of a young American woman according to a set of preselected speaker attributes. When substantial vocal fry, relative to no vocal fry, was present in the woman's utterances, listeners identified her as sounding fundamentally more educated, professional, genuine, and nonaggressive. Other studies have also reported favorable listener impressions of speakers using

Accepted for publication August 1, 2017.

From the Department of Communicative Disorders and Deaf Education, Utah State University, Logan, Utah.

Address correspondence and reprint requests to Stephanie Borrie, Department of Communicative Disorders and Deaf Education, Utah State University, 1000 Old Main Hill, Logan, UT 84322-1000. E-mail: stephanie.borrie@usu.edu

Journal of Voice, Vol. 32, No. 5, pp. 538–545

0892-1997

© 2017 The Voice Foundation. Published by Elsevier Inc. All rights reserved.

<https://doi.org/10.1016/j.jvoice.2017.08.002>

vocal fry, linking the speech feature with higher ratings of trustworthiness, intelligence, and friendliness.¹⁶

However, unfavorable listener impressions associated with the use of vocal fry in the speech of American women have also been reported. Anderson and colleagues¹⁰ examined the effect of vocal fry on listener judgments (18–65 years) of speakers (19–27 years) on attributes considered important for success in the job market. Audio recordings of the phrase “*thank you for considering me for this opportunity*” were elicited from seven males and seven females. Each speaker produced one phrase production with vocal fry and one without. Using an online survey tool, listeners were presented with pairs of phrase productions (vocal fry and no fry) and asked to select which one sounded more educated, competent, trustworthy, and attractive, and which speaker they would choose to hire. Data on listener judgments of the female speakers revealed that the use of vocal fry may negatively influence a woman’s job prospects—productions without vocal fry were judged as more educated, competent, trustworthy, and hireable compared with productions with vocal fry. The authors concluded that the presence of vocal fry in the speech of young women may undermine their success in the labor market. Other negative associations have been previously reported with the use of vocal fry linked with impressions of speaker boredom and sadness.¹⁷

In sum, while vocal fry is clearly prevalent in the speech of young American women, there is currently little consensus about how its presence influences listener judgments of the speaker. While methodological differences undoubtedly contribute to discrepancies between existing studies (see Borrie and Delfino¹²), it may also be that the impressions of speakers with vocal fry are influenced by accompanying speech features. As Gobl and Ní Chasaide¹⁷ have stated, a single speech feature cannot be produced in isolation. That is, spoken utterances are produced with a set of speech features including perceptually salient features such as speech rate and voice pitch—both of which have been shown to independently influence listener judgments of the speaker (eg, Aronovitch¹⁸; Breitenstein et al¹⁹). Thus, we postulate that the way in which vocal fry influences listener judgments of speaker attributes is shaped by the surrounding acoustic-prosodic context.

In contrast to the vocal fry literature, there is a relatively strong consensus regarding the influence of speech rate on listener judgments of a speaker. Overall, the studies in the area of speech rate afford evidence that a fast speech rate is perceived more favorably than a slow speech rate regardless of the attribute being examined (eg, Brown²⁰; Rosenberg and Hirschberg²¹). Speakers who use a faster speech rate are consistently judged as more competent, charismatic, and socially attractive than their slower-speaking peers (eg, Rosenberg and Hirschberg²²; Smith et al²³; and Street and Brady²⁴). Additionally, a faster speech rate has been linked with enhanced speaker persuasiveness and credibility,^{25–27} whereas a slower speech rate has been linked with speaker sadness and inexperience.^{19,28}

Voice pitch has also been reported to influence listener judgments of speaker attributes; however the effect is more variable. Studies have reported that lower-pitched voices are associated with speaker dominance and strength^{29,30} and that higher-

pitched voices are judged as less competent and less trustworthy than lower-pitched voices.^{13,31} Yet, other studies have observed that higher-pitched voices are perceived as more charismatic than lower-pitched voices,^{22,32} suggesting that the way in which voice pitch influences listener judgments may be differentially influenced by the attribute under judgment (eg, charisma versus competence). Speaker gender may also differentially influence the pattern of results. When making judgments of speaker attractiveness, listeners judged higher-pitched female voices as more attractive than lower-pitched female voices, but lower-pitched male voices as more attractive than higher-pitched male voices.^{33,34} Conversely, for ratings of agreeableness, listeners judged lower-pitched female voices as more agreeable than higher-pitched female voices, but higher-pitched male voices as more agreeable than lower-pitched male voices.³⁵

The purpose of this study was to offer a more comprehensive understanding of how the use of vocal fry by young adult female speakers of American English influences listener judgments of speaker attributes by accounting for the surrounding acoustic-prosodic context. Given that spoken utterances are always produced with a speech rate and voice pitch, and that these vocal features have been shown to independently influence listener judgments, these speech features were accounted for in the current study design. We also selected two speaker attributes, intelligence and likability, allowing us to relate our work to previous studies, namely those by Yuasa⁹ and Anderson et al,¹⁰ and examine if different attributes yield a different pattern of results. Thus, our research questions addressed the following: (1) Does the influence of vocal fry on listener judgments of young adult female speakers of American English depend on the accompanying voice pitch and speech rate? (2) Are listener judgments based on the speech features of young adult female speakers of American English influenced by the particular attribute being rated (ie, intelligence versus likability)? In our efforts to answer these questions, we addressed limitations of previous studies by using spoken utterances with naturally produced vocal fry and no fry, carefully controlling for other vocal features that may impact judgments (eg, nasality, breathiness) and drawing from a large, geographically distributed sample of listener participants.

METHODS

Listener participants

A total of 463 adults (262 males and 201 females) 20–70 years of age ($M = 35.84$, $SD = 9.78$) participated as listeners in this study. All listener participants were native speakers of American English currently residing in the United States. Demographic information regarding age, geographic region, and level of education of the participants is available in Table 1.

Participants were recruited using the crowdsourcing website Amazon Mechanical Turk (MTurk; <http://www.mturk.com>). All participants were volunteers and protected through MTurk’s participation agreement and privacy notice. We used the MTurk setup option to limit participation to high-performing workers with an earned “master’s” status.³⁶ We also applied a location restriction, permitting only individuals confirmed as current residents

TABLE 1.
Demographic Distribution Data Expressed in Percentage Scores for Listener Participants

	Percentage
Gender	
Males	57
Females	43
Age	
≥ 50	12
40–49	17
30–39	46
≤ 29	25
Education	
Doctorate	<1
Master's	11
Bachelor's	51
Attending college	6
High school graduate	29
GED	<1
Haven't graduated high school	2
Region	
Midwest	17
Northeast	23
Pacific	18
Rocky Mountain	3
Southeast	25
Southwest	14

in the United States to participate in this study. Participants received remuneration (\$1) for taking part in the study.

Speech stimuli

Speech stimuli used in the current study consisted of eight audio-recorded phrase productions obtained from female college students ($n = 8$) aged 18–28 years. The speakers who provided the recordings were all native speakers of Standard American English with no self-reported history of voice, speech, language, or cognitive impairment. Speakers were enrolled in undergraduate and graduate classes at Utah State University (USU), and they signed an Institutional Review Board (IRB)-approved consent form stating that they agreed for their recording to be used in an online crowdsourcing study.

The speakers were explicitly selected based on their habitual vocal features, with different combinations of voice pitch (low or high), speech rate (slow or fast), and vocal fry (presence or absence). As an example, there was one speaker with a low pitch, a slow rate, and the presence of vocal fry; another with a high pitch, a slow rate, and the presence of vocal fry; and another with a low pitch, a fast rate, and the presence of vocal fry, so that our speech stimuli set consisted of eight files representing each of the possible eight vocal feature combinations (see Table 2 for combination details).

On separate occasions, the speakers were brought into a sound-proof audiology booth in the Department of Communicative Disorders at the USU and recorded using a portable digital audio recorder (Zoom H4N, Ronkonkoma, New York). The speakers

TABLE 2.
Vocal Feature Combinations of the Female Speakers Who Produced the Speech Stimuli

Speaker	Voice Pitch	Speech Rate	Vocal Fry
1	Low	Slow	Present
2	Low	Slow	Absent
3	Low	Fast	Present
4	Low	Fast	Absent
5	High	Slow	Present
6	High	Slow	Absent
7	High	Fast	Present
8	High	Fast	Absent

TABLE 3.
Mean and Standard Deviation for Fundamental Frequency and Speech Rate of the Speech Stimuli Produced by the Female Speakers

	Fundamental Frequency (Hz)		Speech Rate (sps)	
	Low	High	Slow	Fast
Mean	172.49	248.79	3.84	5.39
SD	15.01	13.82	0.09	0.24

Notes: Four speakers per category. Hz and sps refer to hertz and syllables per second, respectively. SD refers to standard deviation.

were told that their task was to read aloud a phrase using their typical speaking voice. The phrase (“*when a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow*”), taken from the standard passage reading, the Rainbow Passage,³⁷ was selected to provide approximately 5–8 seconds^a of audio per speaker.

The audio file of each speaker's phrase production was analyzed for a measure of fundamental frequency (F0) (in Hertz) and speech rate (in syllables per second [sps]) using acoustic analysis software, *Praat* (Amsterdam, The Netherlands).³⁹ Acoustic analysis confirmed that relative to the average F0 range for adult female voices,⁴⁰ the speakers classified as having low or high pitch were indeed low (≤ 195 Hz) and high (≥ 225 Hz), respectively. The speakers classified as having a fast or slow speech rate were indeed fast (≥ 5 sps) and slow (≤ 4 sps), respectively. Means and standard deviations for objective voice pitch and speech rate data are reported in Table 3. The presence of vocal fry was verified visually using waveforms in *Praat*, with vocal fry identified as having more irregular pulses relative to modal voice in the waveforms and low-frequency vibration.⁴¹ Two certified speech-language pathologists not associated with the study also independently identified that the four audio files classified as having the presence of vocal fry were indeed characterized as having substantial vocal fry. The four audio files classified as being absent of vocal fry were also characterized by no occurrences of vocal fry.

^aInitial impressions of speakers, based on their voice, can be formed after listening to only subsecond extracts of speech.³⁸

Experimental program

The eight audio files were calibrated to a 65-dB sound pressure level and programmed into a Web-based listener-perception application hosted on a secure university-based Web server. The core nature of the application (see Procedure for specific details) was to present the listener participants with the eight files and have them rate each one on their agreement with statements pertaining to judgment of our two speaker attributes of interest: *intelligence* (ie, this person is intelligent) and *likability* (ie, this person is likable). To reduce the dimensionality of the data, two additional statements representing latent variables for each attribute of interest were also included in the application. See [Appendix A](#) for all six statements used in the experimental program.

Procedure

A brief description of the study task (ie, listening to voices of different speakers and rating them on how intelligent and likable they sound), requirements (use of headphones), time commitment (approximately 10 minutes), and remuneration amount (\$1) was posted on MTurk. Interested workers were directed to a Web page that was loaded with the experimental procedure. Before beginning the study, MTurk workers/listener participants were required to read through the IRB-approved consent form and indicate acceptance of the terms and conditions in the document by clicking agree. Participants were then required to complete a brief questionnaire regarding age, gender, and basic demographics (see Participants section for details). After this, the experimental portion of the procedure began.

Participants were told that they would be presented with eight audio clips of different female speakers saying the same phrase. They were informed that after each audio clip was played, their task would be to rate the clip on a series of six statements. They were told to listen carefully to the clip as they would only hear it once.^b The clips were then presented one at a time, and following each clip, participants were presented each statement accompanied by a seven-point Likert-type rating scale. They were asked to use the scale to indicate to what extent they agreed with each statement (see [Appendix B](#) for a sample statement with a rating scale). After responding to each statement, the participants pressed submit to move on to the next statement. Once responses for all six statements had been completed for the audio file in question, the next file was presented. Once the eight files had been completed (yielding a total of 48 data points), the participant was thanked and given a confirmation code to use to receive remuneration in MTurk. The self-paced experimental portion of the procedure took, on average, 10 minutes to complete. Audio files and their accompanying statements were both randomized across all participants to eliminate any potential order effects.

Statistical analysis

Pairwise comparisons between the substatements and the key statements of interest for intelligence (ie, this person is intelligent)

^bA single trial ensured that all participants had the same amount of exposure to form their initial impression. Initial impressions, based on a single trial, have been shown to be consistent across large numbers of listeners.³⁸

and likability (ie, this person is likable) revealed significant correlations (<0.47–0.69). Accordingly, rating data from our key statements of interest were used in the analysis. To analyze our data, we used generalized estimating equations (GEE) modeling—a method built on linear regression that accounts for the repeated measures, producing unbiased estimates and valid standard errors. The GEE models simply control for the within-subject correlation, allowing for an accurate estimation of the between-subject effects. The GEE model has been shown to be a useful tool for analyzing longitudinal data and repeated measures.⁴² In our study, we perform two GEE models, one for listener ratings of speaker intelligence and one for listener ratings of speaker likability.

RESULTS AND DISCUSSION

Intelligence

As illustrated in [Figure 1](#) (top panels), results for the GEE model for listener ratings of speaker intelligence revealed a significant three-way interaction between voice pitch, speech rate, and vocal fry ($\beta = 1.48$, $SEM = .13$, $P < 0.001$). The results revealed that ratings of intelligence were highest for the speaker who presented with a low pitch, a fast rate, and no vocal fry. When vocal fry was present in exactly the same acoustic-prosodic context (ie, low pitch, fast rate), ratings of intelligence were significantly reduced. If considered in isolation, this finding that the presence of vocal fry is less favorable than its absence can be taken as support for the Anderson et al¹⁰ study and their results advancing that the use of vocal fry by young American women may undermine their ability to be successful in the labor market.^c This, however, is not the full story. When the speaker presented with a low pitch and a slow rate, the influence of vocal fry was reversed. In this acoustic-prosodic context, intelligence ratings for the speaker with vocal fry were significantly greater than ratings for the speaker with no vocal fry. Again, if this finding was considered in isolation, it could be taken as support for the Yuasa⁹ study in which listeners identified a woman with vocal fry as sounding more educated, professional, genuine, and nonaggressive.^d Taken together, however, our results demonstrate that the influence of vocal fry is largely dependent on the surrounding acoustic-prosodic context, specifically voice pitch and speech rate.

Interestingly, vocal fry did not influence listener judgments in all acoustic-prosodic contexts. The top-right panel of [Figure 1](#) reveals that when the speaker presented with a high pitch, the presence/absence of vocal fry had no significant impact on ratings of intelligence. Rather, in this particular context, speech rate appears to drive ratings of intelligence. Whether vocal fry was present or not, speakers with a high pitch and a fast rate achieved significantly greater ratings of intelligence than speakers with

^cWe carried out an acoustic analysis of the speech samples used in the Anderson et al¹⁰ study, accessed at <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0097506>. The analysis revealed that on average, and according to the high/low and fast/slow parameters used in the current study, the female speakers using vocal fry presented with a low pitch and a fast speech rate.

^dSpeech samples in the Yuasa⁹ study are not publically available. Without an acoustic analysis of the voice pitch and speech rate of the samples, comparisons with the current study are purely speculative.

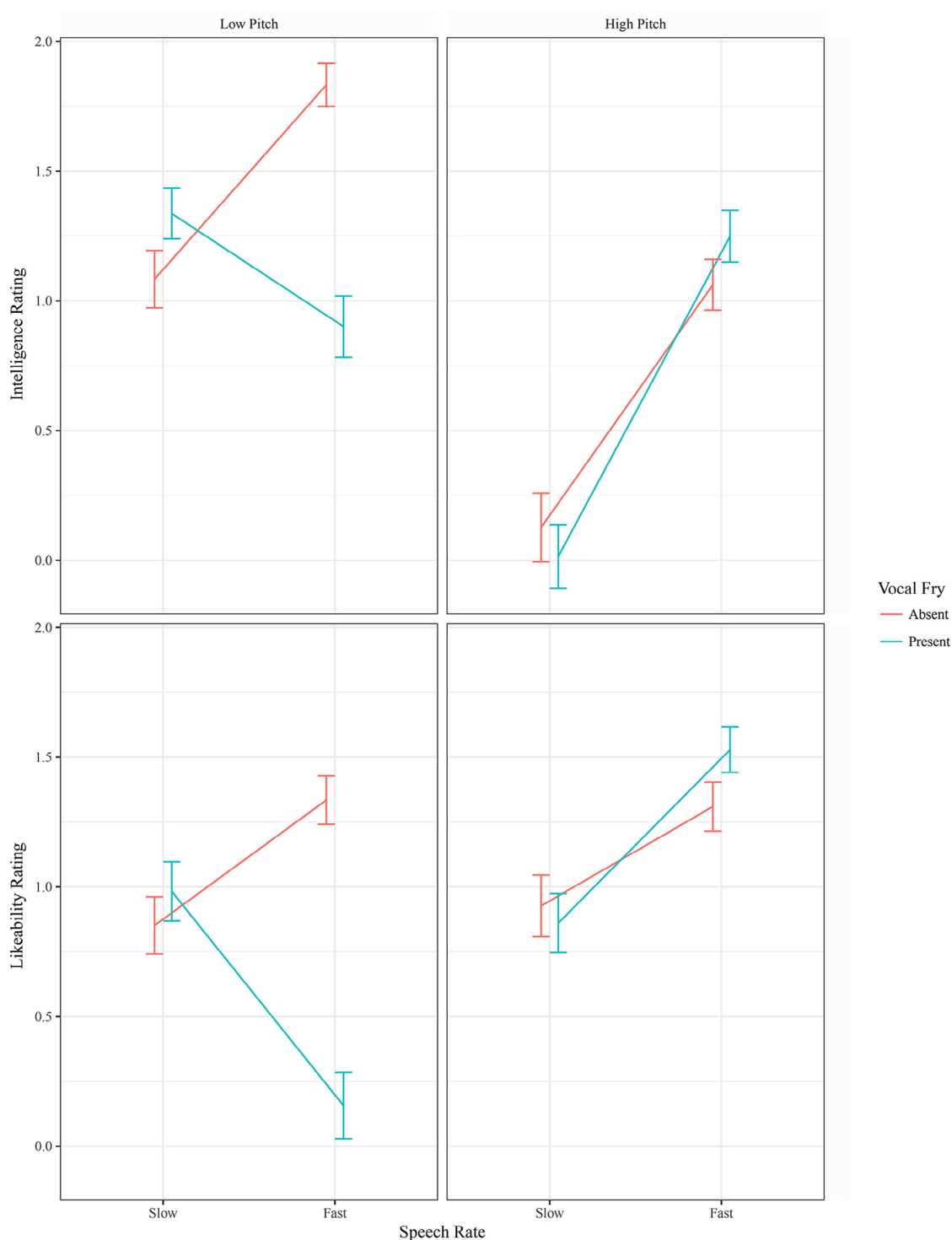


FIGURE 1. Listener judgments of speaker attributes. From top to bottom, the panels reflect three-way interactions for the voice features of voice pitch, speech rate, and vocal fry for attributes of intelligence and likability, respectively.

a high pitch and a slow rate. The positive influence of a fast speech rate is perhaps not surprising given the large body of literature linking this speech feature with high listener ratings on many desired attributes, including speaker intelligence (eg, Rosenberg and Hirschberg^{21,22}). Investigators have postulated that the link between a fast speech rate and favorable impressions of a speaker may arise from cognitive-linguistic development during

adolescence, whereby a fast speech rate correlates with increased cognitive processing, speech motor control, and linguistic processing.⁴³ As such, speakers who habitually use a fast rate in adulthood may project a high level of intelligence. This effect of speech rate, however, was not consistent across contexts. In the context of low pitch and vocal fry, fast rate was not favorable. These results highlight that the influence of speech behaviors

on listener impressions should not be studied in isolation. The surrounding acoustic-prosodic context matters.

Likability

Results for the GEE model for listener ratings of likability also revealed a significant three-way interaction between voice pitch, speech rate, and vocal fry ($\beta = 1.60$, $SEM = .13$, $P < 0.001$) and are illustrated in Figure 1 (bottom panels). Overall, the results for judgments of likability revealed global similarities with the patterns revealed for judgments of intelligence. Consistent with intelligibility judgments, in the context of low pitch and fast rate, the presence of vocal fry had an unfavorable influence on listener ratings. Given that low pitch and fast rate were found to characterize the female speakers with vocal fry in the Anderson et al¹⁰ study, these results can partially validate the conclusions of that study: employers are looking for employees who are not only intelligent but also likable, the caveat being that the unfavorable influence of vocal fry and potential implications for navigating the job market is only valid in the context of low pitch and fast rate.

In the context of high pitch and fast rate, the presence of vocal fry had a favorable influence on listener ratings. Indeed, ratings of likability were highest for the speaker who presented with this specific combination of speech features (ie, high pitch, fast rate, vocal fry). This finding was not evident in the intelligibility judgments. We surmise that the link between likability and the presence of vocal fry in this particular context may arise from the widespread attention that vocal fry has received in regard to its use by famous actresses and influential socialites.⁴⁴ Anecdotaly, a number of the prolific social role models who use vocal fry also employ a high voice pitch and a fast speech rate. Thus, listeners may associate this particular combination of vocal features with popularity and high levels of likability. An acoustic analysis of the speech rate and voice pitch of these women is required to validate this speculation.

General discussion

The primary purpose of our study was to examine the hypothesis that listener judgments of vocal fry in the speech of young adult female speakers of American English are differentially influenced by the surrounding acoustic-prosodic context, specifically speech rate and voice pitch. We also hypothesized that listener judgments are differentially influenced by the attribute being rated. The results of this study confirmed both hypotheses, demonstrating that unique combinations of voice pitch (high or low), speech rate (fast or slow), and vocal fry (presence or absence) influenced listener judgments of intelligence and likability in varying ways. To briefly summarize, in the context of low pitch and fast rate, the presence of vocal fry reduced listener ratings of both attributes. Conversely, in the context of high pitch and fast rate, the presence of vocal fry elevated listener ratings (significant for likability and trending for intelligence). In all other pitch and rate combinations, the presence of vocal fry was relatively inconsequential to initial impressions of female speakers of American English. These findings afford a novel explanation for the discrepancies reported in existing literature.^{9,10} To highlight explicitly, the female speakers with vocal fry in Anderson

et al¹⁰ presented with a low pitch and a fast rate, rendering negative listener impressions when vocal fry was present. Conversely, the speaker in Yuasa⁹ may have presented with a high pitch and a fast rate, resulting in positive impressions of her productions with vocal fry relative to her productions without vocal fry.

To address the key purpose of this study, our results were based on a large sample of listener participants that included both males and females, ages ranging from 20 to 70 years, and residents in geographic regions across the United States. Ratings were collectively analyzed to achieve a high level of power regarding how different combinations of vocal features shape listener judgments of speaker attributes. Previous studies have revealed that listener age can influence impressions of speakers. Anderson et al,¹⁰ for example, performed a fine-grained analysis of their data and reported that while all listeners judged vocal fry as negative for ratings of speaker competence, older listeners were significantly more negative in their judgments (see also Blume⁴⁵; Saunders⁴⁶). Studies have also revealed that listener judgments may be differentially influenced by listener gender.^{33,34} and by geographic location.⁴⁷ To explore this with our data, we carried out two *post hoc* analyses—the first comparing data from younger (below 40 years; $n = 329$) and older (above 40 years; $n = 134$) listeners and the second comparing data from female ($n = 201$) and male listeners ($n = 262$). For both analyses, the patterns of results were virtually identical to the collectively analyzed data. While the age comparison was not evenly distributed and targeted comparisons across geographic locations are also warranted, the *post hoc* analyses of this study indicates that the pattern of results may be relatively robust across adult listeners in the United States.

Speakers have long been interested in how to use their voice to create positive impressions and/or connect with those they are interacting with. The adoption of vocal features can be subconscious, as in mothers using infant-directed speech to convey emotionally rich communication to their infant (eg, Trainor et al⁴⁸) or modifying speech patterns, including the use of vocal fry, to entrain to the patterns of one's communication partner (eg, Borrie and Delfino¹²; Borrie and Liss⁴⁹). The adoption of vocal features can also be conscious. Professional voice users, such as actors, engage in vocal training in order to convey a confident and dynamic character for their audience.⁵⁰ Another example of a conscious adoption of vocal features can be seen in transgender individuals who elect to receive voice therapy to assist with projecting a more feminine or masculine voice.^{51,52} While issues surrounding subconscious versus conscious deployment of vocal features require evaluation, our results suggest that if a young American woman desires to be perceived as intelligent, adopting a low pitch, a fast rate, and no vocal fry may be beneficial. Conversely, if likability is desired, a high pitch, a fast rate, and vocal fry might be optimal.

To control for potential accent and dialectal influences on listener impressions, the speakers who provided the speech stimuli for the present study were recruited from the same region: the state of Utah. This does, however, reduce the generalizability of the findings, particularly to speakers from other geographic regions and those with different dialects (eg, southern accent). Boucher et al,⁵³ for example, have reported that speakers with

southern accents are judged less competent than speakers with more standard American accents. Accordingly, a study replication using speech stimuli elicited from speakers with different dialects offers an important future direction for this work. Further, the current study could be carried out using speech stimuli elicited from male speakers. While the large majority of the literature on speakers with vocal fry, particularly the studies that have targeted listener judgments, have focused on females, males also use vocal fry (eg, Abdelli-Beruh et al¹¹). Additionally, features such as low pitch have been evidenced as preferable to high pitch in male speakers (eg, Tsantani et al⁵⁴; Puts et al⁵⁵). Thus, different patterns regarding the way in which vocal fry and the surrounding acoustic-prosodic context shapes listener judgments may be revealed for males. Future studies could also expand the acoustic-prosodic context to include other speech features such as nasality, tenseness, and breathiness (see Pittam¹⁶ for an earlier study on how such features may contribute to listener impressions).

CONCLUSION

In conclusion, listener judgments of young American women based on information afforded in their speech behaviors are not solely based on the presence or absence of vocal fry, but rather a combination of speech features that interact with one another in unique ways. For example, when the speaker had a low pitch and a fast rate, the presence of vocal fry reduced listener ratings, but when the speaker had a high pitch and a fast rate, the presence of vocal fry elevated listener ratings. Thus, whether or not vocal fry is associated with favorable judgments is largely dependent on the surrounding acoustic-prosodic context. Furthermore, the contribution of vocal fry to listener judgments also depends on the speaker attribute being rated. The highest ratings of intelligence were achieved when the speaker presented with a low pitch, a fast rate, and no vocal fry, whereas the highest ratings of likability were achieved with a high pitch, a fast rate, and vocal fry. Such findings shed light on a possible reason for the discrepancies in the existing literature regarding listener impressions of vocal fry and advance the idea that listener judgments of speaker attributes are based on a combination of speech features.

Acknowledgments

We gratefully acknowledge Paul Vicioso Osoria for the development of the Web-based application and Tyson Barrett for consultation on statistical methods.

APPENDIX A

Rating Questions

To what extent do you to agree with this statement?

1. This person is intelligent.
2. You would take advice from this person.
3. This person holds a graduate level college degree.
4. This person is likeable.
5. You could be friends with this person.
6. People enjoy spending time with this person.

APPENDIX B

Sample Question and Rating Scale

To what extent do you to agree with this statement?

1. This person is intelligent.

Strongly Agree	Agree	Slightly Agree	Neutral	Slightly Disagree	Disagree	Strongly Disagree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="button" value="Submit"/>						

REFERENCES

1. Hollien H, Michel J. Vocal fry as a phonational register. *J Speech Hear Res.* 1968;11:600–604.
2. McGlone RE, Schipp T. Some physiologic correlates of vocal-fry phonation. *J Speech Hear Res.* 1971;14:769–775.
3. Blomgren M, Chen Y, Ng M, et al. Acoustic, aerodynamic, and perceptual characteristics of modal and vocal fry registers. *J Acoust Soc Am.* 1998; 103:2649–2658.
4. Childers DG, Lee CK. Vocal quality factors: analysis, synthesis, and perception. *J Acoust Soc Am.* 1991;90:2394–2410.
5. Colton RH, Casper JK, Leonard R. *Compreendendo os problemas de voz: uma perspectiva fisiológica ao diagnóstico e ao tratamento.* 1996;3:215.
6. Zemlin WR. Phonation. In: *Speech and Hearing Science: Anatomy and Physiology.* 1998:100–196.
7. Hollien H. On vocal registers. *J Phon.* 1974;2:125–143.
8. Wolk L, Abdelli-Beruh NB, Slavin D. Habitual use of vocal fry in young adult female speakers. *J Voice.* 2012;26:111e–116e.
9. Yuasa IP. Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women? *Am Speech.* 2010;85:315–337.
10. 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.
11. Abdelli-Beruh NB, Wolk L, Slavin D. Prevalence of vocal fry in young adult male American English speakers. *J Voice.* 2014;28:111–116.
12. Borrie SA, Delfino CR. Conversational entrainment of vocal fry in young adult female American English speakers. *J Voice.* 2017;31:513, e25–e32.
13. Klofstad CA, Anderson RC, Peters S. Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proc R Soc B.* 2012;297:2698–2704.
14. Nass C, Lee KM. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *J Exp Psychol Appl.* 2001;7:171.
15. Scherer KR. *Personality Markers in Speech.* Cambridge University Press; 1979.
16. Pittam J. Listeners evaluations of voice quality in Australian English-speakers. *Lang Speech.* 1987;30:99–113.
17. Gobl C, Ní Chasaide A. The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* 2003;40:189–212.
18. Aronovitch CD. The voice of personality: stereotyped judgments and their relation to voice quality and sex of speaker. *J Soc Psychol.* 1976;99:207–220.
19. Breitenstein C, Van Lancker D, Daum I. The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cogn Emot.* 2001;15:57–79.
20. Brown BL. Effects of speech rate on personality attributions and competency evaluations. In: Giles H, Robinson WP, Smith P, eds. *Language: Social Psychological Perspectives.* 1980:293–300.
21. Rosenberg A, Hirschberg J. Acoustic/prosodic and lexical correlates of charismatic speech. *INTERSPEECH.* 2005;513–516.

22. Rosenberg A, Hirschberg J. Detecting pitch accents at the word, syllable and vowel level. In: *Proceedings of the Human Language Technologies: Association of Computational Linguistics, Short Papers*. 2009:81–84.
23. Smith BL, Brown BL, Strong WJ, et al. Effects of speech rate on personality perception. *Lang Speech*. 1975;18:145–152.
24. Street RL Jr, Brady RM. Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Commun Monogr*. 1982;49:290–308.
25. Mehrabian A, Williams M. Nonverbal concomitants of perceived and intended persuasiveness. *J Pers Soc Psychol*. 1969;13:37–58.
26. Miller GA, Johnson-Laird PN. *Lang Percept*. Cambridge, MA: Belknap Press.; 1976:760.
27. Perloff RM. *The Dynamics of Persuasion. Communication and Attitude in the 21st Century*. Hillsdale, NJ: Lawrence Erlbaum; 2003.
28. Yokoyama H, Daibo I. Effects of gaze and speech rate on receivers' evaluations of persuasive speech. *Psychol Rep*. 2012;110:663–676.
29. Borkowska B, Pawlowski B. Female voice frequency in the context of dominance and attractiveness perception. *Anim Behav*. 2011;82:55–59.
30. Puts DA, Hodges CR, Cárdenas RA, et al. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol Hum Behav*. 2007;28:340–344.
31. Apple W, Streeter LA, Krauss RM. Effects of pitch and speech rate on personal attributions. *J Pers Soc Psychol*. 1979;37:715–727.
32. Biadys F, Rosenberg A, Carlson R, et al. A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech. *Speech Pros*. 2008:579–582. Available at: <http://www.isca-speech.org/archive>.
33. Re DE, O'Connor JJM, Bennett PJ, et al. Preferences for very low and very high voice pitch in humans. *PLoS ONE*. 2012;7:e32719.
34. Hodges-Simeon CR, Gaulin SJC, Puts DA. Different vocal parameters predict perceptions of dominance and attractiveness. *Hum Nat*. 2010;21:406–427.
35. Imhof M. Listening to voices and judging people. *Int J Listen*. 2010;24:19–33.
36. Peer E, Vosgerau J, Acquisti A. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behav Res Methods*. 2014;46:1023–1031.
37. Fairbanks G. *Voice and Articulation Drillbook*. 2nd ed. New York: Harper & Row.; 1960:124–139.
38. McAleer P, Todorov A, Belin P. How do you say "Hello"? Personality impressions from brief novel voices. *PLoS ONE*. 2014;9:e90779.
39. Boersma P, Weenink D. Praat: doing phonetics by computer (Version 6.0.09) [computer program]. 2013: Available at: <http://www.praat.org/>.
40. Stemple JC, Roy N, Klaben BK. *Clinical Voice Pathology: Theory and Management*. 5th ed. San Diego, CA: Plural Publishing.; 2013.
41. Ladefoged P, Maddieson I. The sounds of the world's languages. *Language*. 1998;74:374–376.
42. Diggle P, Liang K-Y, Zeger SL. *Longitudinal Data Analysis*. New York, NY: Oxford University Press; 1994;5:13.
43. Nip ISB, Green JR. Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age. *Child Dev*. 2013;84:1324–1337.
44. Wolf N. Young women, give up the vocal fry and reclaim your strong female voice. *The Guardian*. 2015.
45. Blume H. Totally fried. *ASHA Leader*. 2016;21:50–56.
46. Saunders M. Use of vocal fry may damage professional image of young employees. Available at: <http://www.k-state.edu/media/newsreleases/feb13/vocalfry22513.html>.
47. Alford RL, Strother JB. A southern opinion of regional accents. *Fla Comm J* 1992;20:49–59.
48. Trainor LJ, Austin CM, Desjardins RN. Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol Sci*. 2000;11:188–195.
49. Borrie SA, Liss JM. Rhythm as a coordinating device: entrainment with disordered speech. *J Speech Lang Hear Res*. 2014;57:815–824.
50. Hardison DM, Sonchaeng C. Theater voice training and technology in teaching oral skills: integrating the components of a speech event. *System*. 2005;33:593–608.
51. Azul D, Nygren U, Södersten M, et al. Transmasculine people's voice function: a review of the currently available evidence. *J Voice*. 2017;31:261, e9–e23.
52. Hancock AB, Krissing J, Owen K. Voice perceptions and quality of life in transgender people. *J Voice*. 2011;25:553–558.
53. Boucher CJ, Hannock GS, McLaughlin SD, et al. Perceptions of competency as a function of accent. *Psi Chi J Psychol Res*. 2013;18:27–32.
54. Tsantani MS, Belin P, Paterson HM, et al. Low vocal pitch preference drives first impressions irrespective of context in male voices but not female voices. *Perception*. 2016;0:1–18.
55. Puts DA, Jones BC, DeBruine LM. Sexual selection on human faces and voices. *J Sex Res*. 2012;49:227–243.