Fertility-Dependent Acoustic Variation in Women’s Voices Previously Shown to Affect Listener Physiology and Perception

Melanie L. Shoup-Knox, Grant M. Ostrander, Gabrielle E. Reimann, and R. Nathan Pipitone

Abstract
Previous research demonstrates that listeners perceive women’s voices as more attractive when recorded at high compared to low fertility phases of the menstrual cycle. This effect has been repeated with multiple voice recording samples, but one stimuli set has shown particularly robust replications. First collected by Pipitone and Gallup (2008), women were recorded counting from 1–10 on approximately the same day and time once a week for 4 weeks. Repeatedly, studies using these recordings have shown that naturally cycling women recorded at high fertility are rated as more attractive compared to voices of the same women at low fertility. Additionally, these stimuli have been shown to elicit autonomic nervous system arousal and precipitate a rise in testosterone levels among listeners. Although previous studies have examined the acoustic properties of voices across the menstrual cycle, they reach little consensus. The current study evaluates Pipitone and Gallup’s voice stimuli from an acoustic perspective, analyzing specific vocal characteristics of both naturally cycling women and women taking hormonal contraceptives. Results show that among naturally cycling women, variation in vocal amplitude (shimmer) was significantly lower in high fertility recordings compared to the women’s voices at low fertility. Harmonics-to-noise ratio and variation in voice pitch (jitter) also fluctuated systematically across voices sampled at different times during the menstrual cycle, though these effects were not statistically significant. It is possible that these acoustic changes could account for some of the replicated perceptual, hormonal, and physiological changes documented in prior literature using these voice stimuli.

Keywords
acoustic analysis, menstrual cycle, fertility, female voice attractiveness

Hormonal fluctuations across the menstrual cycle influence important behaviors related to mate assessment and attractiveness. For example, faces, body odors, and even gait of fertile-phase women who are naturally cycling have been judged as more attractive compared to the same women during non-fertile phases (Bobst & Lobmaier, 2012; Gildersleeve, Haselton, & Fales, 2014; Guéguen, 2012; Havlíček, Dvořáková, Bartoš, & Flegr, 2006; Rikowski & Grammer, 1999; Roberts et al., 2004). These subtle yet systematic changes in women’s physiology and appearance as a function of menstrual cycle phase have important evolutionary implications. As Thornhill and Gangestad (2008) originally propose, the evidence to date suggests that traits changing as a function of fertility are not signals but are “leaked” cues connected to the hormone changes necessary for the reproductive cycle. While women’s estrous behaviors may serve to obtain males with traits indicative of “good genes,” perceiving these cyclic changes poses an adaptive strategy for men, who could benefit from preferring or monitoring fertile women, or for women who might benefit from monitoring the fertility status of rivals (Gangestad & Thornhill, 2008; Hurst, Alquist, & Puts, 2017; Krems, Neelj, Neuberg, Puts, & Kenrick, 2014).

Corresponding Author:
Melanie Shoup-Knox, Psychology Department, James Madison University, 91 East Grace St., MSC 7704, Harrisonburg, VA 22807, USA. Email: shoupml@jmu.edu

1 Psychology Department, James Madison University, Harrisonburg, VA, USA
2 Department of Psychology, Florida Gulf Coast University, Fort Myers, FL, USA

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Vocal Acoustics Across the Menstrual Cycle

Previous work has identified similarities between the hormonal cytology of the larynx and the cervix, demonstrating that the larynx is a target organ for cyclical fluctuations of hormones (Abitbol et al., 1989; Abitbol, Abitbol, & Abitbol, 1999). The larynx is particularly sensitive to sex hormones: estrogen, progesterone, and androgens. These may promote or obstruct vocal clarity, stability, and production, depending on the level of each hormone (Abitbol et al., 1999). Some researchers have even argued that almost every drastic change in women’s hormonal profile has a concomitant effect on voice quality and/or production (Caruso et al., 2000; Firat et al., 2009).

If shifts in women’s hormonal profiles alter voice perception, there could be concomitant changes in vocal acoustics that compel listeners’ perceptual differences. Vocal acoustics commonly examined in the literature include (but are not limited to) fundamental frequency \( F_0 \), formant dispersion, amplitude, jitter, shimmer, and harmonics-to-noise ratio (HNR). Less commonly examined properties of vocal acoustics that may affect listeners’ perception of vocalizations include spectral slope and spectral centroid. All of these parameters can vary dramatically between individuals, and a large body of evidence shows that changes in vocal parameters affect perceptual judgments of the speaker (e.g., Pipitone et al., 2016; Puts, Barndt, Welling, Dawood, & Burriss, 2011; Puts, Gaulin, & Ver dolorini, 2006; Puts, Hodges, Cárdenas, & Gaulin, 2007; Puts, Jones, & DeBruine, 2012).

Fundamental frequency is the objective correlate of pitch. The standard deviation \( SD \) of \( F_0( F_0 - SD ) \) quantifies how much change in pitch there is across an entire utterance. Low values of \( F_0 - SD \) across an utterance may lead to perceptions of monotony, whereas high variation may influence perceptions of increased prosody. Jitter is a more refined measure of the cycle-to-cycle \( F_0 \) variability that quantifies rapid changes in \( F_0 \) as opposed to changes occurring throughout the course of an utterance. Amplitude is the objective correlate of loudness, and shimmer is a measure of the rapid variation in amplitude throughout a vocalization. HNR quantifies how much sound the vibration of the vocal folds produces relative to other sounds produced by the passage of air through the articulators or produced by the articulators themselves. Low HNR can be caused by a less abrupt or a lack of complete glottal closure and may contribute to perceptions of increased breathiness, roughness, or hoarseness in a voice (de Krom, 1995; Yumoto, Gould, & Baer, 1982). Spectral slope is the rate at which the amplitude of an utterance declines as a factor of increasing frequency and may also influence perceptions of breathiness (Hillenbrand, Cleveland, & Erickson, 1994; Hillenbrand & Houde, 1996). The spectral centroid of an utterance is the average of the frequencies produced within a vocalization, weighted by their corresponding amplitudes, conceptually representing the center of mass of an utterance. Spectral centroid position correlates with the perception of a sound’s brightness and sharpness (Schubert & Wolfe, 2006). Sounds at higher centroid positions are perceived as brighter and sharper than those at lower centroid position, which are perceived as warmer.

Using within-subjects designs, several researchers have explored how these acoustic parameters vary across the menstrual cycle, but the results are mixed. Bryant and Haselton (2009) showed increases in \( F_0 \) during higher fertility times. Pavela Banai (2017) found an increase in the minimum \( F_0 \) at high fertility compared to lower fertility, but no change in mean \( F_0 \) values across the cycle. In contrast, both Karthikeyan and Locke (2015) and Fischer et al. (2011) documented decreases in \( F_0 \) at high fertility. Fischer et al. (2011) also found lower HNR during menstruation than during ovulation. Studies examining jitter and shimmer have consistently found lower levels of both measures among women on hormonal contraceptives compared to those who were naturally cycling (Amir, Biron-Shental, Muchnik, & Kishon-Rabin, 2003; Amir & Kishon-Rabin, 2004; Amir, Kishon-Rabin, & Muchnik, 2002; Lã, Ledger, Davidson, Howard, & Jones, 2007). However, findings of systematic variation in levels of jitter and shimmer within naturally cycling women across the menstrual cycle have been inconsistent (Amir & Kishon-Rabin, 2004; Chae, Choi, Kang, Choi, & Jin, 2001; Higgins & Saxman, 1989). Furthermore, other researchers have been unable to identify any systematic acoustic changes across the menstrual cycle (Celi, et al., 2013; Kunduk, Vansant, Ikuma, & McWhorter, 2017; Puts et al., 2013).
Given the influence of hormones on the physiology of the vocal folds, researchers have investigated whether menstrual cycle phase affects ratings of voice attractiveness. Many studies have found increased attraction to women speaking at higher fertility times (Karthikeyan & Locke, 2015; Ostrander et al., 2018; Pipitone & Gallup, 2008, 2012; Pipitone et al., 2016; Puts et al., 2013; Shoup-Knox & Pipitone, 2015). One stimuli set, in particular, has produced multiple replications of perceptual effects. Using voices recorded from the same women counting from 1 to 10 at four different times across the menstrual cycle, Pipitone and Gallup (2008) found that both men and women rated recordings of women’s voices as more attractive as the probability of conception increased. Using the same vocal stimuli, Pipitone and Gallup (2012) also showed that men judge women’s voices recorded during menstruation to be of lower mood, quality, and attractiveness than the same women’s voices recorded during other phases. Shoup-Knox and Pipitone (2015) later used these recordings to replicate listeners’ preference for high fertility voices as well as demonstrate that high fertility voices produce a larger sympathetic nervous system response than low fertility voices from the same women. The voices of women using hormonal contraception produced no difference in listener physiology. Most recently, these voice stimuli were found to induce an increase in testosterone among women listeners (Ostrander et al., 2018). These same listeners also rated each voice for attractiveness once during ovulation and then again during their late luteal phase. Again, high fertility voices were rated as more attractive than low fertility voices, but interestingly, the effect was greater when the listener was ovulating compared to ratings of the voices during the late luteal phase. These findings suggest that although men might attend to women’s voice as a mating strategy, women might also attend to other women’s voice in the context of intrasexual competition.

It is intriguing that the same set of voice stimuli has repeatedly produced cycle-dependent differences in ratings of attractiveness, changes in physiology, and changes in hormones among listeners. Therefore, it is of particular importance to explore the acoustic parameters of this particular stimuli set, despite the lack of consensus in the literature regarding generalized acoustic changes across the menstrual cycle. The current article explores whether there are any within-subject systematic shifts in acoustics among this particular set of naturally cycling women’s voices and voices of women using hormonal contraceptives. Thus, our analysis will provide a better understanding of the proximate mechanisms driving the previously documented perceptual shifts in women’s voices across the menstrual cycle.

Method

Participants

Participants were recruited and screened as previously described by Pipitone and Gallup (2008). Women included in the sample were nonsmokers, reported regular menstrual cycles, and had not taken the morning-after pill or been pregnant within 3 months prior to the recording session. We assessed menstrual cycle regularity through participant report of whether they experience equal or varied number of days per cycle. Fifty-one women were originally recruited through a research participant pool at a northeastern United States university and compensated US $2.50 for each voice sample they provided. The study was approved by the university’s Institutional Review Board. Those included in the present analysis consists of 19 naturally cycling women ranging in age from 17 to 30 (M = 20.63, SD = 3.2) and 18 women on hormonal contraceptives ranging in age from 18 to 26 (M = 19.78, SD = 1.83). Each analyzed recording was used as stimuli in one or more of the following studies: Pipitone and Gallup (2008), Pipitone and Gallup (2012; recordings were shortened to counting from 1 to 5), Shoup-Knox and Pipitone (2015), and Ostrander, Pipitone, and Shoup-Knox (2018).

Fertility Assessment

In their first recording session, women reported the date of onset of their most recent menstruation, provided a voice sample, and returned once a week, at approximately the same day and time, for 3 additional weeks to provide a total of four voice recordings. Since each woman began her first recording session at a different day in her cycle (i.e., whenever she signed up for the experiment), the order of recordings across the cycle was treated as a random variable. Cycle phase was reassessed during each weekly meeting by having women re-report the date of their most recent menstruation. Most women experienced another menstrual onset at some point during the study, rendering this self-report measure more accurate than most one-time self-report methods. Continuous backward assessment of fertility with follow-up confirmation of the next menstruation onset is the most valid technique for assessing fertility outside of hormonal analysis (Gangestad et al., 2016). The experimenters calculated the probability of conception on each day of recording based on established probabilities described by Wilcox, Dunson, Weinberg, Trussell, and Baird (2001). Cycle days were then standardized to generate more systematic conception rates (Pipitone & Gallup, 2008). The voice recordings from each woman were then organized from lowest to highest conception likelihood. The present study examined two of the four voice recordings; recordings that corresponded to the lowest and highest fertility times, resulting in a total of 74 recordings analyzed, as these were the voices that produced replicated perceptual effects among listeners. Voice recordings from intermediate likelihood of conception do not map onto specific cycle phases and therefore do not represent unique and distinguishable hormonal profiles.

The average cycle day among the naturally cycling women at high fertility was 12.47 (SD = 1.86). Among voices recorded at low fertility in the same women, the average cycle day was 3.31 (SD = 1.75) for recordings taken during the early follicular phase and 25.17 (SD = 2.31) for recordings taken during the late luteal phase. Among the women using hormonal contraceptives, the average cycle day for voices recorded at high...
Proceeding to analyze for acoustic measures. All recordings were of participants counting from 1 to 10 in a slow, steady pace. This utterance was chosen by previous researchers (Pipitone & Gallup, 2008) such that the content would be neutral and standardized across each participant.

All voice samples provided were of participants counting from 1 to 10 in a slow, steady pace. This utterance was chosen by previous researchers (Pipitone & Gallup, 2008) such that the content would be neutral and standardized across each participant. Visual inspection of audio spectrograms revealed that some manipulations of the recordings were necessary before proceeding to analyze for acoustic measures. All recordings exhibited direct current (DC) offset and were normalized to exhibit 0% DC bias. Due to experimenter error, seven of the recordings were sampled at 22,050 Hz and a bit depth of 8. These recordings were converted to a sampling rate of 44.1 kHz and a bit depth of 16. Of these converted recordings, four were provided by naturally cycling women (one at high fertility and three at low fertility) and three were provided by women on hormonal contraceptives (one at high fertility and two at low fertility). Additionally, ambient background noise (i.e., air circulation, white noise caused by electronics, etc.) was eliminated using the noise reduction feature of Adobe Audition.

The noise print for each recording was captured from approximately the first second of each recording, prior to the onset of the utterance. Based on this noise print, we then reduced the ambient noise by 60 dB throughout the entirety of each recording to ensure that the background noise would not interfere with analyses of vocal acoustics. The result of this manipulation was to reduce (by a factor of 1,000) only the signal that exhibited direct current (DC) offset and were normalized to exhibit 0% DC bias. Due to experimenter error, seven of the recordings were sampled at 22,050 Hz and a bit depth of 8. These recordings were converted to a sampling rate of 44.1 kHz and a bit depth of 16. Of these converted recordings, four were provided by naturally cycling women (one at high fertility and three at low fertility) and three were provided by women on hormonal contraceptives (one at high fertility and two at low fertility). Additionally, ambient background noise (i.e., air circulation, white noise caused by electronics, etc.) was eliminated using the noise reduction feature of Adobe Audition.

The noise print for each recording was captured from approximately the first second of each recording, prior to the onset of the utterance. Based on this noise print, we then reduced the ambient noise by 60 dB throughout the entirety of each recording to ensure that the background noise would not interfere with analyses of vocal acoustics. The result of this manipulation was to reduce (by a factor of 1,000) only the signal that precisely matched the background signal recorded in the absence of the speaker making a vocalization.

After editing the samples for recording consistency and background noise, Praat software was used to generate voice reports including measures of median $F_0$, mean $F_0$, $F_0 - \sigma D$, jitter, shimmer, HNR, and formant dispersion (Df). Although mean $F_0$ is commonly used in vocal analyses, median $F_0$ was also retained as it is not affected as severely by outliers when frequency tracking software falsely identifies unnaturally high or low fundamental frequencies within an utterance. Praat software also provides five different measures of jitter and six measures of shimmer. We report individual acoustics in Table 1 and performed a principal component analysis (PCA) to collapse

### Materials

As previously described by Pipitone and Gallup (2008), voice samples were recorded using an Altec Lansing AHS515 headset (Altec Lansing, New York, NY) with the microphone placed 8 cm away from all participants’ mouths to control for variation in signal intensity and signal-to-noise ratio that could result as an artifact of distance between speaker and microphone. Samples were recorded onto a computer using Microsoft Sound Recorder 5.0, in mono, at an intended sampling rate of 44.1 kHz and bit depth of 16. Of these converted recordings, four were provided by naturally cycling women (one at high fertility and three at low fertility) and three were provided by women on hormonal contraceptives (one at high fertility and two at low fertility). Additionally, ambient background noise (i.e., air circulation, white noise caused by electronics, etc.) was eliminated using the noise reduction feature of Adobe Audition.

The noise print for each recording was captured from approximately the first second of each recording, prior to the onset of the utterance. Based on this noise print, we then reduced the ambient noise by 60 dB throughout the entirety of each recording to ensure that the background noise would not interfere with analyses of vocal acoustics. The result of this manipulation was to reduce (by a factor of 1,000) only the signal that precisely matched the background signal recorded in the absence of the speaker making a vocalization.

After editing the samples for recording consistency and background noise, Praat software was used to generate voice reports including measures of median $F_0$, mean $F_0$, $F_0 - \sigma D$, jitter, shimmer, HNR, and formant dispersion (Df). Although mean $F_0$ is commonly used in vocal analyses, median $F_0$ was also retained as it is not affected as severely by outliers when frequency tracking software falsely identifies unnaturally high or low fundamental frequencies within an utterance. Praat software also provides five different measures of jitter and six measures of shimmer. We report individual acoustics in Table 1 and performed a principal component analysis (PCA) to collapse

### Acoustic Analyses

All voice samples provided were of participants counting from 1 to 10 in a slow, steady pace. This utterance was chosen by previous researchers (Pipitone & Gallup, 2008) such that the content would be neutral and standardized across each participant. Visual inspection of audio spectrograms revealed that some manipulations of the recordings were necessary before proceeding to analyze for acoustic measures. All recordings

---

**Table 1.** Mean values for acoustics of recordings of naturally cycling women and women on hormonal contraceptives at high and low fertility (SD).

| Acoustic                        | Naturally Cycling (N = 19) | Hormonal Contraceptive Users (N = 18) |
|--------------------------------|---------------------------|---------------------------------------|
|                                | High Fertility | Low Fertility | $\eta^2_p$ | High Fertility | Low Fertility | $\eta^2_p$ |
| Median $F_0$                   | 197.27 (14.85) | 196.70 (15.12) | .003 | 203.63 (19.90) | 202.78 (22.85) | .005 |
| Mean $F_0$                     | 198.12 (16.46) | 194.58 (17.34) | .056 | 200.82 (21.02) | 197.98 (26.52) | .048 |
| $F_0 - \sigma D$               | 29.17 (10.10) | 29.42 (12.18) | <.001 | 30.94 (13.57) | 30.46 (18.88) | .001 |
| Jitter (local)                 | 1.33 (0.30) | 1.45 (0.37) | .142 | 1.44 (0.41) | 1.40 (0.32) | .199 |
| Jitter (local, absolute)       | 0.0001 (0.00002) | 0.0001 (0.00003) | .042 | 0.0001 (0.00002) | 0.0001 (0.00002) | <.001 |
| Jitter (rap)                   | 0.0063 (0.002) | 0.0096 (0.011) | .085 | 0.0070 (0.002) | 0.0068 (0.001) | .019 |
| Jitter (ppq5)                  | 0.0064 (0.002) | 0.0098 (0.012) | .077 | 0.0069 (0.002) | 0.0067 (0.001) | .014 |
| Jitter (ddp)                   | 0.0188 (0.005) | 0.0208 (0.007) | .117 | 0.0211 (0.006) | 0.0205 (0.004) | .019 |
| Shimmer (local)                | 6.81* (1.14) | 7.28* (1.34) | .225 | 7.15 (0.31) | 7.18 (1.24) | .001 |
| Shimmer (local, dB)            | 0.73* (0.10) | 0.77* (0.10) | .295 | 0.77 (0.11) | 0.78 (0.10) | .008 |
| Shimmer (apq3)                 | 0.0268 (0.007) | 0.0293 (0.007) | .187 | 0.0288 (0.007) | 0.0285 (0.006) | .002 |
| Shimmer (apq5)                 | 0.0341* (0.008) | 0.0375* (0.008) | .314 | 0.0357 (0.007) | 0.0365 (0.008) | .012 |
| Shimmer (apq11)                | 0.056 (0.01) | 0.060 (0.01) | .150 | 0.058 (0.01) | 0.057 (0.01) | <.001 |
| Shimmer (dds)                  | 0.081 (0.02) | 0.090 (0.02) | .187 | 0.087 (0.02) | 0.086 (0.02) | .002 |
| HNR                            | 15.22 (1.77) | 14.67 (1.85) | .155 | 14.02 (2.40) | 14.18 (2.03) | .009 |
| Df of first four formants      | 908.64 (35.67) | 915.99 (31.00) | .128 | 918.87 (61.89) | 910.64 (64.48) | .012 |
| Spectral centroid              | 1.435.97 (176.18) | 1.438.33 (138.16) | <.001 | 1.420.37 (163.18) | 1.438.29 (181.60) | .014 |
| Spectral slope                 | −0.0042 (0.0012) | −0.0042 (0.0011) | .002 | −0.0041 (0.0008) | −0.0041 (0.0008) | .003 |

Note. Bolded acoustics significantly different at *p < .05. HNR = harmonics-to-noise ratio.
Kreiman, 2016; Kreimann, Gerratt, & Antoñanzas-Barroso, 2007), regressing all measured amplitudes on this range of frequencies was chosen as it captures the entirety of the measured vocal spectrum.

**Statistical Analyses**

First, each acoustic measure was analyzed using a within-subjects design across high and low fertility times. This was done for both naturally cycling women, as well as those using hormonal contraceptives, despite these women lacking true cyclic fertility. If the perceptual preference for high fertility voices is due to changes in hormones affecting vocal acoustics, we expected to see differences between the high and low fertility voices. However, several of these measures (jitter: rap, ppq5, ddp; shimmer: apq3, apq5, apq11) are produced by averaging bins that span multiple points during the recording. Because these recordings are of continuous utterances rather than sustained sounds, these measures would capture increased variation due to variation in words and phonemes rather than vocal characteristics. Therefore, our discussion focuses only on the data from local measures of jitter and shimmer. Formant dispersion, the averaged difference between successive formant frequencies, should be interpreted with similar caution. Because formants are affected by the changing shape of the articulators, this measure is also most appropriate for sustained vowel utterances. The production of sustained vowel utterances involves minimal change in the shape of one’s articulators, and hence better ability to accurately assess formant dispersion, as opposed to continuous speech that will inherently require movement of and variation in the shape of the articulators. We used Adobe Audition (Version 5) software to conduct a frequency analysis used in the calculation of the spectral centroid of each recording. The frequency analysis used a Fast Fourier Transformation size of 4,096, yielding 2,048 bins. The 500 bins that comprised frequencies between 100 Hz and 5,500 Hz (frequencies that would be feasibly produced by human speech) were retained for the calculation of spectral centroids. Spectral slope was calculated by regressing the amplitudes of each bin on their respective frequencies. Although multiple ways of measuring spectral slope have been used (Garrellek, Samlan, Gerratt, & Kreiman, 2016; Kreimann, Gerratt, & Antoñanzas-Barroso, 2007), regressing all measured amplitudes on this range of frequencies was chosen as it captures the entirety of the measured vocal spectrum.

**Results**

**Assessment of Individual Acoustics**

Paired t-tests were used to assess whether median $F_0$, mean $F_0$, $F_0 - SD$, jitter, shimmer, HNR, spectral centroids, or

| Principle Component | High vs. Low Fertility | Contraception vs. Naturally Cycling | Fertility × Contraception Use |
|---------------------|------------------------|------------------------------------|-----------------------------|
| Component 1         |                        |                                    |                             |
| Jitter, jitDDP, JitAbs, Shimmer, Shimmerdb, ShimAPQ3, 5, and 11, ShimDDA, and HNR | $F(1, 35) = 1.29$ | $F(1, 35) = 0.25$ | $F(1, 35) = 2.34$ |
| $p = .19$ | $p = .62$ | $p = .14$ |
| $\eta^2_p = .05$ | $\eta^2_p = .01$ | $\eta^2_p = .06$ |
| Component 2         |                        |                                    |                             |
| JitRAP and JitPPQ5  | $F(1, 35) = 1.17$ | $F(1,35) = 0.46$ | $F(1,35) = 1.64$ |
| $p = .29$ | $p = .5$ | $p = .21$ |
| $\eta^2_p = .03$ | $\eta^2_p = .01$ | $\eta^2_p = .05$ |
| Component 3         |                        |                                    |                             |
| Median and mean $F_0$, JitAbs | $F(1, 35) = 0.35$ | $F(1, 35) = 0.34$ | $F(1, 35) = 0.01$ |
| $p = .56$ | $p = .57$ | $p = .94$ |
| $\eta^2_p = .01$ | $\eta^2_p = .01$ | $\eta^2_p = .006$ |
| Component 4         |                        |                                    |                             |
| $D_1$ and $D_2$, spectral slope, and $F_0 - SD$ | $F(1, 35) = 0.02$ | $F(1, 35) = 0.02$ | $F(1, 35) = 0.34$ |
| $p = .88$ | $p = .89$ | $p = .57$ |
| $\eta^2_p = .001$ | $\eta^2_p = .00$ | $\eta^2_p = .01$ |
| Component 5         |                        |                                    |                             |
| Spectral slope and spectral centroid | $F(1, 35) = 0.03$ | $F(1, 35) = 0.08$ | $F(1, 35) = 0.03$ |
| $p = .87$ | $p = .78$ | $p = .86$ |
| $\eta^2_p = .001$ | $\eta^2_p = .002$ | $\eta^2_p = .001$ |

Note. Individual acoustic variables comprising each component are listed below component number. HNR = harmonics-to-noise ratio; PCA = principal component analysis.
spectral slope varied as a factor of fertility status (see Table 1). Among naturally cycling women, fertility status accounted for 29.5% of the variance in average local shimmer values, \( t(18) = -2.75, p = 0.01, \eta^2_p = 0.3 \). Shimmer values were lower at high fertility \((M = 0.73, SD = 0.10)\) than at low fertility \((M = 0.77, SD = 0.10)\); see Figure 1). We found no other significant differences in acoustics between the high and low fertility recordings of naturally cycling women \((all \ t(s) < |1.85|, ps > 0.08)\). However, although not statistically significant, fertility status did account for 15.5% of the variance in HNR (see Figure 2) and 14.2% of the variance in jitter (see Figure 3) among naturally cycling women, with women at high fertility exhibiting higher HNR and lower jitter than when at low fertility. Among women taking hormonal contraceptives, there was no significant difference in any acoustic measure across the two recordings \((all \ t(s) < |1|, ps > 0.36)\).

Between-subjects analyses were also conducted to assess if any acoustics differed as a factor of hormonal contraceptive use. No differences in any acoustic were evident between naturally cycling women and women taking hormonal contraceptives at high fertility \((all \ t(s) < |1.73|, ps > 0.09)\) or at low fertility \((all \ t(s) < |1.07|, ps > 0.29)\).

**Assessment of Principal Components**

The PCA extraction yielded five components, with eigenvalues above 1.0 being considered important when compared to individual observed variables (Tabachnick & Fidell, 2007). The Kaiser–Meyer–Olkin measure of sampling adequacy was not significant \((p = 0.75)\).

Individual variable loadings were considered to load onto components if they reached the 0.4 threshold (Tabachnick & Fidell, 2007). Examination of the pattern matrix revealed that the first component was comprised of mostly jitter and shimmer variables as well as HNR (Jitter, JitDDP, JitAbs, Shimmer, Shimmerdb, ShimAPQ3, ShimAPQ5, ShimAPQ11, ShimDDA, and HNR). This “Jitter/Shimmer” component had the highest eigenvalue of 8.58 and explained 45.17% of the variance in factor loadings. The second component was comprised of two measures of jitter (JitRAP and JitPPQ5) and had an eigenvalue of 2.2, explaining 11% of the variance. Component 3 was comprised of two fundamental frequency variables (median \(F_0\) and mean \(F_0\)) and one jitter variable (JitAbs) and had an eigenvalue of 1.88, explaining 9.89% of the variance. Component 4 was comprised of two formant variables (\(Df_1\) and \(Df_2\)), spectral slope, and variability of fundamental frequency \((F_0 – SD)\) and had an eigenvalue of 1.72, explaining 9.06% of the variance. Component 5 was comprised of spectral slope and centroid acoustics and had an eigenvalue of 1.34, explaining 7.04% of the variance.

A 2 × 2 mixed model analysis of variance was performed to examine the effects of contraception use and fertility on each component identified in the PCA. None of the five components
significantly varied as a function of contraception use, fertility phase of recording, or the interaction between the two factors (see Table 2). The highest amount of variance explained was 6%, explained by the interaction of fertility and contraception use on the first component (jitter/shimmer). The data trend for this finding showed that naturally cycling women had lower scores on the jitter/shimmer component when at high fertility compared to the same women at low fertility or compared to women on hormonal contraception (which did not differ based on fertility). Testing \( t \)-tests to assess whether the components varied as a factor of fertility status within each group of women revealed that naturally cycling women had significantly lower values in the first component at high fertility compared to low fertility, \( t(18) = -2.4, p = 0.03 \). No other significant differences were found in the other components between the high and low fertility among naturally cycling women (all \( ts < |1.23|, ps > 0.23 \)). No significant differences were found in any components between the high and low fertility among women using hormonal contraception (all \( ts < |0.55|, ps > 0.59 \)).

**Discussion**

When a set of vocal stimuli produces multiple replicated effects among listeners, it is appropriate to examine the proximate mechanisms that could be responsible for the ultimate function of such changes (i.e., enabling the voice to serve as a fertility cue). Therefore, the present study examined voice recordings used in previous studies (Ostrander et al., 2018; Pipitone & Gallup, 2008, 2012; Shoup-Knox & Pipitone, 2015) for acoustic properties and compared those properties both within subjects across the menstrual cycle and between subjects based on hormonal contraceptive use. Results showed that voices recorded at high fertility exhibited significantly lower shimmer compared to voices recorded at low fertility, among naturally cycling women (see Figure 1). Lower shimmer represents less variability in amplitude, ostensibly producing a steadier sounding voice. This supports what others have found regarding shimmer at high compared to low fertility (Tatar et al., 2016) and no significant difference in shimmer at high compared to low fertility (Shoup-Knox et al., 2018; Pipitone & Gallup, 2008; Shoup-Knox & Pipitone, 2015), differences in shimmer between high and low fertility were not found in recordings of women taking hormonal contraceptives. Although not apparent in our sample, several studies have shown that when collapsing across cycle phase, women taking hormonal contraceptives exhibit lower shimmer and jitter values than naturally cycling women (Amir et al., 2002, 2003; Amir & Kishon-Rabin, 2004). Taking cycle phase into account, Läß and colleagues (2007) found naturally cycling women to have higher shimmer values than those on hormonal contraceptives during their menstrual and follicular phases but not their luteal phase.

We also found higher HNR and lower jitter values (see Figures 2 and 3) among naturally cycling women at high fertility, though these differences were not statistically significant. Among naturally cycling women, fertility status accounted for over 15% of the variance in HNR and 14% of the variance in jitter values. Differences of this magnitude were not found among women on hormonal contraception. Similarly, Fischer and colleagues (2011) found HNR of women’s voices to be highest at ovulation compared to 3 days prior to ovulation or during menstruation. Further, Tatar et al. (2016) found HNR to be lowest during the late luteal phase compared to other phases. However, neither Çelik et al. (2013) nor Bryant and Haselton (2009) found HNR to vary cyclically.

Similar to several studies by Amir, Kishon-Rabin, and Muchnik (2002), Amir, Biron-Shental, Muchnik, and Kishon-Rabin (2003), a study by Çelik et al. (2013), as well as Chae, Choi, Kang, Choi, and Jin (2001), we found a small effect size and no significant difference in \( F_0 \) as a function of fertility status. As with HNR, previous literature does not provide consistent evidence that \( F_0 \) varies systematically across the menstrual cycle. Bryant and Haselton (2009) found higher \( F_0 \) within 3 days of women’s surge of luteinizing hormone that typically triggers ovulation, and Pavela Banai (2017) found higher minimum \( F_0 \) values at high fertility (also using luteinizing hormone as a measure of ovulation). Other studies have reported that \( F_0 \) is lower at the time of ovulation (Fischer et al., 2011; Karthikeyan & Locke, 2015). Still others have documented variation in \( F_0 \) specific to other phases of the menstrual cycle (Tatar et al., 2016).

There are limitations within the present study. This analysis was exploratory in nature, referencing previous research for
assessment of acoustics. Similar to these previous studies which implemented no or lenient corrections, we did not control for experiment-wise error (Bryant & Haselton, 2009; Chae et al., 2001; Celik et al., 2013; Fischer et al., 2011; Kunduk et al., 2017; Pavela Banai, 2017; Puts et al., 2013; Tatar et al., 2016). Although the current findings yield no significant differences with correction, it bears relevance to highlight impactful effect sizes in research. A PCA was employed to reduce acoustic variables into meaningful components, thus minimizing the likelihood of Type I error. While the composite variables produced show interesting patterns of acoustic relationships, none significantly differed based on fertility or contraceptive use. However, the first “jitter/shimmer” component was significantly lower for naturally cycling women at high fertility but not for women on contraceptives, which comports with the findings from the original analysis of individual acoustic variables (see Table 1). It should be noted that our samples size was below that recommended for PCA (Osborne & Costello, 2004). Sample size may have also limited our ability to detect important fertility-based differences. Ganges-tad et al. (2016) recommend at least 34 women in our within-subjects groups to achieve adequate power, given the validity of our fertility assessment methods. These sample size limitations cannot be avoided by working with an archival set of recordings. However, given the importance of this set of voices, we maintain that our results are considerable and informative to the field in which it currently stands.

Conclusions

The present study found shimmer to be significantly lower during counting utterances at high fertility compared to counting utterances of the same women at low fertility, along with heightened HNR and lower jitter values at high fertility (though these latter results were not statistically significantly different). This comports with some previous work (Chae et al., 2001; Tatar et al., 2016), which suggests improved voice quality at high fertility (Pavela Banai, 2017). It is possible that these acoustic parameters are important components of vocalizations that are influenced by the effects of increased estrogen and lower progesterone levels around the time of ovulation. We suspect that variations in these parameters, as well as others perhaps too subtle to detect, are proximate mechanisms that contribute to listeners’ perceptual and physiological reactions to the high fertility voices. These mechanisms may serve as fertility cues and have reproductive implications.

When it comes to human mate assessment, it is important to remember that human interactions involve speech with heavily embedded content, emotion, intent (spoken or unspoken), accent, and so on, which are accompanied by gestures and adjustments in postural cues. It may not be possible to measure all acoustic parameters associated with these factors, yet humans are capable of evaluating them almost immediately. To put the present findings in a larger perspective, it is important to realize that, among humans, the voice is a medium for the transmission of a surprisingly rich array of biologically relevant information. The mere sound of a person’s voice has been shown to convey information about the speaker’s sex, age, deviations from bilateral symmetry, health, fertility, body configuration, facial attractiveness, grip strength, sexual behavior, propensity for infidelity, use of hormonal contraceptives, and fertility status (for review see: Gallup & Frederick, 2010).

Even with the human voice conveying important physiological and reproductive parameters, it is important to note that every voice is a unique ensemble of the acoustic parameters measured in this study (as well as others we are not currently able to measure or that are yet to be identified). Some of the changes necessary to make a voice more attractive to listeners will be unique to that individual depending on the “baseline” acoustics of that voice. Therefore, it is not surprising that studies have failed to produce robust evidence for any single acoustic parameter that varies systematically across the menstrual cycle for all women. Furthermore, the exact utterance and specific social context in which the vocalization is generated will change vocal characteristics. For example, Frac-caro et al. (2011) and Hughes, Farley, and Rhodes (2010) demonstrate changes in women’s voice pitch when speaking to an attractive man, but results were in opposite directions—higher pitch in the former and lower pitch in the latter. However, the human brain has seemingly evolved to respond to slight and subtle changes in features linked with reproductive potential, thus listeners do not need to explicitly know about these specific proximate mechanisms in order to make assessments of others. Our evolutionary history appears to have solved that problem for us.

Acknowledgments

The authors thank Dr. Michael Hall for his contributions to these analyses. His time, software, and expertise in psychoacoustics made this project possible.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Melanie L. Shoup-Knox https://orcid.org/0000-0002-3261-1370

References

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. Journal of Voice, 13, 424–446.

Abitbol, J., de Brux, J., Millot, G., Masson, M. F., Languille Mimoun, O., Pau, H., . . . Abitbol, B. (1989). Does a hormonal vocal cord cycle exist in women? Study of vocal premenstrual syndrome in voice performers by videostroboscopy-glottography and cytology on 38 women. Journal of Voice, 3, 157–162.
Amir, O., Biron-Shental, T., Muchnik, C., & Kishon-Rabin, L. (2003). Do oral contraceptives improve voice quality? Limited trial on low-dose formulations. *Obstetrics & Gynecology, 101*, 773–777.

Amir, O., & Kishon-Rabin, L. (2004). Association between birth control pills and voice quality. *Laryngoscope, 114*, 1021–1026.

Amir, O., Kishon-Rabin, L., & Muchnik, C. (2002). The effect of oral contraceptives on voice: Preliminary observations. *Journal of Voice, 16*, 267–273.

Ancillao, A., Galli, M., Mignano, M., Dellaballe, R., & Albertinin, G. (2013). Quantitative analysis of pathological female human voice by processing complete sentence recordings. *Journal of Laryngology and Voice, 3*, 46–51.

Bobst, C., & Lobmaier, J. S. (2012). Men’s preference for the ovulating female is triggered by subtle face shape differences. *Hormones and Behavior, 62*, 413–417.

Bryant, G. A., & Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biological Letters, 5*, 12–15.

Caruso, S., Roccasalva, L., Sapienza, G., Zappala, M., Nuciforo, G., & Biondi, S. (2000). Laryngeal cytological aspects in women with surgically induced menopause who were treated with transdermal estrogen replacement therapy. *Fertility and Sterility, 74*, 1073–1079.

Çelik, O., Çelik, A., Atespere, A., Boyaci, Z., Çelebi, S., Gündüz, T., . . . Yelken, K. (2013). Voice and speech changes in various phases of menstrual cycle. *Journal of Voice, 27*, 622–626.

Chae, S. W., Choi, G., Kang, H. J., Choi, J. O., & Jin, S. M. (2001). Clinical analysis of voice change as a parameter of premenstrual syndrome. *Journal of Voice, 15*, 278–283.

de Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *Journal of Speech and Hearing Research, 38*, 794–811.

Deal, R., & Emanuel, F. (1978). Some waveform and spectral features of vowel roughness. *Journal of Speech and Hearing Research, 21*, 250–264.

Feinberg, D. R. (2004). Fundamental frequency perturbation indicates perceived health and age in male and female speakers. *Acoustical Society of America Journal, 115*, 2609–2609.

Frat, Y., Engin-Ustun, Y., Kizilay, A., Ustun, Y., Akarcay, M., Selimoglu, E., . . . Kafkasli, A. (2009). Effect of intranasal estrogen on vocal quality. *Journal of Voice, 23*, 716–720.

Fischer, J., Semple, S., Fickenscher, G., Jürgens, R., Kruse, E., Heisterman, M., & Amir, O. (2011). Do women’s voices provide cues of the likelihood of ovulation? The importance of sampling regime. *PLoS One, 6*, e24490.

Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., . . . DeBruine, L. M. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology, 9*, 57–67.

Gallup, G. G. Jr., & Frederick, D. A. (2010). The science of sex appeal: An evolutionary perspective. *Review of General Psychology, 14*, 240–250.

Gangestad, S. W., Haselton, M. G., Welling, L. L., Gildersleeve, K., Pillsworth, E. G., Burriss, R. P., . . . Put, D. A. (2016). How valid are assessments of conception probability in ovulatory cycle research? Evaluations, recommendations, and theoretical implications. *Evolution and Human Behavior, 37*, 85–96.

Gangestad, S. W., & Thornhill, R. (2008). Human oestrous. *Proceedings of the Royal Society B, 275*, 991–1000.

Garellek, M., Samlan, R., Gerratt, B. R., & Kreiman, J. (2016). Modeling the voice source in terms of spectral slope. *Journal of the Acoustical Society of America, 139*, 1404–1410.

Gildersleeve, K., Haselton, M. G., & Fales, M. (2014). Do women’s mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin, 140*, 1205–1259.

Guéguen, N. (2012). Gait and menstrual cycle: Ovulating women use sexier gaits and walk slowly ahead of men. *Gait Posture, 35*, 621–624.

Havlicek, J., Dvorakova, R., Bartos, L., & Flegr, J. (2006). Non-advertized does not mean concealed: Body odour changes across the human menstrual cycle. *Ethology, 112*, 81–90.

Higgins, M. B., & Saxman, J. H. (1989). Variations in vocal frequency perturbation across the menstrual cycle. *Journal of Voice, 3*, 233–243.

Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research, 37*, 769–778.

Hillenbrand, J., & Houdé, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39*, 311–321.

Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and physiological changes in response to the physical attractiveness of conversational partners. *Journal of Nonverbal Behavior, 34*, 155–167.

Hurst, A. C., Alquist, J. L., & Puts, D. A. (2017). Women’s fertility status alters other women’s jealousy and mate guarding. *Personality and Social Psychology Bulletin, 43*, 191–203.

Karthekeyan, S., & Locke, J. L. (2015). Men’s evaluation of women’s speech in a simulated dating context: Effects of female fertility on vocal pitch and attractiveness. *Evolutionary Behavioral Sciences, 9*, 55–67.

Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of Speech, Language and Hearing Research, 50*, 595–610.

Krems, J. A., Neel, R., Neuberg, S. L., Puts, D. A., & Kentick, D. T. (2016). Women selectively guard their (desirable) mates from ovulating women. *Journal of Personality and Social Psychology, 110*, 551–573.

Kunduk, M., Vansant, M. B., Ikuma, T., & McWhorter, A. (2017). The effects of the menstrual cycle on vibratory characteristics of the vocal folds investigated with high-speed digital imaging. *Journal of Voice, 31*, 182–187.

Lå, F. M. B., Ledger, W. L., Davidson, J. W., Howard, D. M., & Jones, G. L. (2007). The effects of a third generation combined oral contraceptive pill on the classical singing voice. *Journal of Voice, 21*, 754–761.

Linville, S. E. (1996). The sound of senescence. *Journal of Voice, 10*, 190–200.

Necka, E. A., Faig, K. E., Van Hedger, K., Lyons, I. M., Dimitroff, S. J., Luhmann, M., . . . Norman, G. J. (2018). Women’s attention to and memory for fertile- and non-fertile phase women across the menstrual cycle. *Adaptive Human Behavior and Physiology, 4*, 283–305.
Orlikoff, R. (1990). The relationship of age and cardiovascular health to certain acoustic characteristics of male voices. *Journal of Speech and Hearing Research, 33*, 450–457.

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principle components analysis. *Practical Assessment, Research & Evaluation, 9*, 1–9.

Ostrander, G. M., Pipitone, R. N., & Shoup-Knox, M. L. (2018). Interactions between observers and stimuli fertility status: Endocrine and perceptual responses to intrasexual vocal fertility cues. *Hormones and Behavior, 98*, 191–197.

Pavela Banai, I. (2017). Voice in different phases of menstrual cycle among naturally cycling women and users of hormonal contraceptives. *PLoS One, 12*, e0183462.

Pipitone, R. N., & Gallup, G. G. (2008). Women’s voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior, 29*, 268–274.

Pipitone, R. N., & Gallup, G. G. (2012). The unique impact of menstruation on the female voice: Implications for the evolution of menstrual cycle cues. *Ethology, 118*, 281–291.

Pipitone, R. N., Gallup, G. G., & Bartels, A. (2016). Variation in men’s masculinity affects preferences for women’s voices at different points in the menstrual cycle. *Evolutionary Behavioral Sciences, 10*, 188–201.

Puts, D. A., Bailey, D. H., Cárdenas, R. A., Burriss, R. P., Welling, L. L., Wheatley, J. R., . . . Dawood, K. (2013). Women’s attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Hormones and Behavior, 63*, 13–19.

Puts, D. A., Barndt, J. L., Welling, L. M. D., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences, 50*, 111–115.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior, 27*, 283–296.

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men’s voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior, 28*, 340–344.

Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of Sex Research, 49*, 227–243.

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proceedings of the Royal Society London B*, 266, 869–874.

Roberts, S. C., Havlicek, J., Flegr, J, Hruskova, M., Little, A. C., & Petrie, M. (2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proceedings of the Royal Society B*, 271, S270–S272.

Schubert, E., & Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid. *Acta Acustica United with Acustica, 92*, 820–825.

Shoup-Knox, M. L., & Pipitone, R. N. (2015). Physiological changes in response to hearing female voices recorded at high fertility. *Physiology & Behavior, 139*, 386–392.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York, NY: Allyn and Bacon.

Tatar, E. C., Sahin, M., Demiral, D., Bayir, O., Saylam, G., Ozdek, A., . . . Korkmaz, M. H. (2016). Normative values of voice analysis parameters with respect to menstrual cycle in healthy adult Turkish women. *Journal of Voice, 30*, 322–328.

Thornhill, R., & Gangestad, S. W. (2008). *The evolutionary biology of human female sexuality*. Oxford, England: Oxford University Press.

Wilcox, A. J., Dunson, D. B., Weinberg, C. R., Trussell, J., & Baird, D. D. (2001). Likelihood of conception with a single act of intercourse: Providing benchmark rates for assessment of post-coital contraceptives. *Contraception, 63*, 211–215.

Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *Journal of the Acoustical Society of America, 71*, 1544–1550.