2016

The Effects of Four Voice Qualities on the Perception of a Female Voice

Andrea G. Levitt
alevitt@wellesley.edu

Margery Lucas
mlucas@wellesley.edu

The Effects of Four Voice Qualities on the Perception of a Female Voice

Andrea G. Levitt and Margery Lucas

Wellesley College

Abstract

Stimuli produced by a female speaker with four different voice qualities, including modal, girlish, breathy and creaky, and three different pitches, were manipulated so that their formants were either more or less dispersed. Participants judged these stimuli on four scales (dominance, attractiveness, sexiness and youthfulness). Modal voice was highly rated for dominance. Creakiness was rated more highly for dominance than breathiness, and women rated creaky voice as more dominant than men. Breathiness was highly rated in voice judgments of attractiveness and sexiness, as was lowered pitch associated with breathiness. Girlish voice was highly rated for youthfulness, and women found girlish voices more attractive than men. Voices with a raised pitch and creaky voice quality also received relatively high youthfulness ratings.

Acknowledgment We are very grateful to Laura Dulude for her assistance in the research for this paper.

Correspondence concerning this article should be addressed to Andrea Levitt, Green Hall, Wellesley College, 106 Central Street, Wellesley, MA 02481. Email: alevitt@wellesley.edu

Features of the human voice, such as pitch and formant spacing, are sexually dimorphic, suggesting that the voice conveys the sex and aspects of the physical characteristics of the speaker that can be relevant in mate choice and determination of status (see Puts, Doll & Hill 2014 for a review). Although many studies have identified the vocal characteristics that lead listeners to identify qualities such as dominance, attractiveness or sexiness in men's voices, less is known about how women's vocal characteristics affect listeners' judgments for similar qualities. Men's judgments of the sexiness and attractiveness of female voices are often based on features that are associated with youthfulness, such as higher pitch (Feinberg, DeBruine, Jones & Perrett 2008; Puts, Barndt, Welling, Dawood & Burriss 2011) or greater formant dispersion (Puts et al. 2011), both of which are associated with young female voices. Whether there are vocal features that distinguish between voices that are considered sexy or attractive and those that are considered merely youthful has been little studied. In addition, other voice qualities such as breathiness, when the vocal cords vibrate loosely, and creakiness, also known as vocal fry, when the vocal cords are held tightly together and only vibrate at one end, have been little examined. Some attention has been paid to how listeners interpret breathiness, a voice quality associated with women (Van Borsel, Janssens & De Bodt 2009), with respect to voice judgments of sexiness (Batstone & Tuomi 1981), but the effect of creakiness, a voice quality on the rise in young American women (Mendoza-Denton 2007; Yuasa 2010) on judgments about the dominance, attractiveness, sexiness or youthfulness of the female voice has been largely unexplored.

For men's voices, judgments of dominance have been extensively studied, and in general, researchers have found that they are associated with low pitch (Apicella, Feinberg

& Marlowe 2007; Feinberg, Jones, Law Smith, Moore, DeBruine, Cornwell, Hillier & Perretta 2006; Fraccaro, O'Connor, Reb, Jones, DeBruine & Feinberg 2013; Jones, Feinberg, DeBruine, Little & Vukovica 2010) and with closely spaced formants; the latter correlates with greater physical size. Indeed Puts, Hodges, Cárdenas, & Gaulin (2007) found that closer formant spacing contributed more to listeners' dominance judgments than the fundamental frequency of the male voice. Judgments of vocal confidence, which is likely related to dominance, have also been associated with voices exhibiting greater intensity or loudness (Kimble & Seidel 1991).

In terms of attractiveness, as with judgments of dominance, listeners perceived men's voices to be attractive or pleasant when they had a low fundamental frequency (Bruckert, Liénard, Lacroix, Kreutzer & Leboucher 2006; Collins 2000; Feinberg, Jones, Little, Burt & Perrett 2005), although some researchers found that both normal pitched and lower pitched men's voices were attractive (Fraccaro, Jones, Vukovic, Smith, Watkins, Feinberg, Little & Debruine 2011; Riding, Lonsdale & Brown 2006). Again, voices with less dispersed formants were also judged to be more attractive (Collins 2000). Furthermore, some researchers have noted that male vocal attractiveness is associated with perceived male dominance (Berry 1990) and greater reproductive success (Apicella et al. 2007; Hughes, Dispenza & Gallup 2004; Puts 2005). Judgments of the sexiness of men's voices are also generally associated with voices with lower fundamental frequencies (Daniel & McCabe 1992). Men who were attempting to sound sexy (Tuomi & Fisher 1979) or who were judged to be successful seducers (Anolli & Cicero 2002) spoke with more pitch excursion and lowered their pitch over time, although Daniel and McCabe (1992) found that mid-pitched male voices were rated the sexiest. In contrast, women rated men's voices

with high fundamental frequency and greater formant spacing as less attractive than those with low pitch and less formant spacing, possibly because these voices are associated with youthfulness in male voices (Feinberg et al. 2005).

In the research on women's voices, listeners' perception of vocal dominance, as in the case of men's voices, is associated with lower pitch (Borkowska & Pawlowski 2011; Fraccaro et al. 2013; Jones et al. 2010) and generally masculinized voices, manipulated to have both lower pitch and less formant dispersion (Feinberg et al. 2006). The relationship between attractiveness and pitch in women's voices, however, is different from that in men's voices. In several studies, male listeners' perception of vocal attractiveness in women is associated with higher-pitched voices (Collins & Missing 2003), as long as the pitch is not too high (Borkowska & Pawlowski 2011; Fraccaro et al. 2013). But the findings overall are mixed, since other work has found that female listeners' judgments of vocal attractiveness in women are associated with voices with normal rather than high or low pitch (Fraccaro et al. 2013). As is the case with men's voices, judgments of female vocal attractiveness were more influenced by formant spacing than by fundamental frequency (Puts et al. 2011); however, in women, formant spacing was greater than in men, which correlates with their smaller physical size. Wider formant spacing in women is also correlated with youthfulness. We will describe stimuli with increased formant dispersion as younger and those with reduced formant dispersion as older.

The findings on sexiness and pitch in women's voices are also mixed. Some studies indicate that both men and women think that low pitch is sexy, because both lower the fundamental frequency of their voices when asked to sound seductive (Tuomi & Fisher 1979) and speak in a lower pitch to attractive opposite-sex listeners (Hughes, Farley &

Rhodes 2010). Farley, Hughes & LaFayette (2013) found that women lowered voice pitch when speaking to their current romantic partners, and such voices were also judged to be sexier than those used when talking to friends. But men in that study actually raised their pitch when speaking to partners. This gender difference might have been due to vocal accommodation to one's interaction partner, possibly to signal affection and connection. Daniel & McCabe (1992) found that men and women label mid-pitched voices as most sexy, and men but not women found high-pitched women's voices to be sexy as well. Low-pitched women's voices were rated as least sexy. However, their values for women's fundamental frequency in the low range (175.2-183.5 Hz) were considerably lower than the low values for Farley et al. (2013) (approximately 225 Hz). The latter were, in fact, closer to Daniel and McCabe's high values (226.6 to 229.2 Hz), so the difference in the results may be due how specific pitch ranges are labeled by researchers.

In general, women's voices are more attractive to men when they are perceived to be younger, i.e., when they have a higher pitch and greater formant spacing (Puts et al. 2011). As noted above, however, women tend to prefer female voices with a normal pitch, so youthful voice characteristics like high pitch are not uniformly associated with attractiveness judgments across genders. Also, youthfulness in the female voice is associated less with dominance than is youthfulness in the male voice (Berry 1992).

In summary, most of the literature looking at perceptions of vocal characteristics of men's and women's voices have focused mainly on the effects of fundamental frequency and formant spacing. But other vocal characteristics can also contribute to perceptions of dominance, attractiveness, sexiness or youthfulness. The voice quality normally used by a speaker is known as modal voice, but speakers can modify their voice quality in several

ways to change how listeners perceive them. For example, they can make adjustments to their vocal cords to make their voices sound breathy or creaky (Crystal & Quirk 1964). Also, women can make their voices sound more girlish and men can make theirs sound more boyish by raising their second formants through the use of tongue fronting (Ladefoged & Harshman 1979). Such effects are enhanced when speakers raise their pitch as well. Although changes in voice quality can be used to signal linguistic differences such as differences in meaning (e.g., Gordon & Ladefoged 2001), voice qualities such as creaky, breathy, and girlish or boyish voice may also serve to signal social or physical characteristics, such as dominance, attractiveness, sexiness or youthfulness.

The use of creaky voice, also known as glottal fry, in young American women has received a great deal of attention recently in the popular press, with articles appearing in the New York Times (Quenada 2012) and the Chicago Tribune (Hageman 2013), among others. Wolk, Abdelli-Beruh and Slavin (2012) found it to be a common voice quality in young English-speaking women, and Yuasa (2010) found that the use of creaky voice in young Californian women was greater than that found in the speech of young Japanese women or young Californian men. Abdelli-Beruh, Wolk and Slavin (2014) found that it was also less common in American males.

The recent increase in the use of creaky voice by young women may be due to its social signaling functions. Researchers have attributed several different social functions to the use of creaky voice. In early studies, investigators noticed that it was used by upper-class men in the United Kingdom to indicate their superior social status (Hendon & Bladon 1988; Esling 1978). Pittam (1987) also found it to be characteristic of Australian males rather than females. These early studies of the use of creaky voice generally interpreted it

as indicating masculinity or authority. The more recent use by young women of creaky voice has been described by Yuasa as "hesitant, nonaggressive and informal, but also educated, urban-oriented and upwardly mobile" (2010, p.315). Other studies have suggested that its use in women is linked to toughness (Mendoza-Denton 2007). To the extent that it is heard as masculine, creaky voice is likely to lower the ratings of the attractiveness of women's voices. On the other hand, it may also contribute to judgments of dominance in women's voices, given its association with authority. Whether the presence of creaky voice among young women is widespread enough to have become associated with youthfulness in the female voice is another open question.

Breathy voice can communicate intimacy (Laver 1980) or sexiness (Crystal, 1975) Indeed, Henton and Blandon (1985) suggest that breathiness signals intimacy because it is associated with the sound of sexual arousal. Kreiman and Sidtis (2013) argue that breathiness is one of several features that contribute to a stereotypically sexy voice, which include, in addition, extreme intonation contours and slow speaking rate. They note that sexy women's voices can either be high pitched, a characteristic associated with youthfulness, or low pitched, associated with maturity. Not many studies have directly examined the effects of breathiness on judgments of women's voices, however. Although breathy women's voices have been rated more feminine (Van Borsel et al. 2009), at least one study did not find that breathiness in women's voices was associated with judgments of sexiness (Batstone & Tuomi 1981), although Hughes, Mogilski and Harrison (2014) did find that women lowered the pitch of their voices and made them hoarser sounding when trying to sound sexy or attractive.

The present study was designed to explore the effect of these less studied voice qualities, in conjunction with components that have received more attention in the literature on vocal judgments. While manipulations exist to increase or decrease the formant dispersion or to raise or lower the fundamental frequency of natural utterances, to our knowledge, there is no technique that would allow natural voice stimuli to be manipulated so as to sound more breathy or more creaky than a speaker's original voice. To that end, we sought to develop a set of stimuli from a single speaker that would show the full range of voice qualities and several different pitches, but would allow us to keep the formant pattern relatively stable, so as to avoid differences in judgment based on different vocal tract sizes. Rather than manipulate the fundamental frequency of the utterances, we asked the speaker to produce her vowels with normal, raised or lowered pitch, so that the variability in pitch was natural. Our only manipulation of these natural utterances was to modify the formants acoustically to make the voice sound either younger or older than the original. We then elicited judgments from female and male listeners of the dominance, attractiveness, sexiness and youthfulness of the various voice types of the selected female speaker.

We expected to replicate generally the effects of pitch and formant spacing on judgments of dominance, attractiveness, sexiness and youthfulness. In terms of voice qualities, we predicted that girlish voices would be rated primarily as youthful and attractive, but not as dominant. Whether girlish voices would be perceived as sexy was less certain. We predicted that creakiness would contribute positively to judgments of dominance and negatively to judgments of attractiveness and sexiness. Whether it has begun to be associated with youthfulness was not clear. Finally, we predicted that breathiness would contribute to judgments of sexiness and perhaps also attractiveness,

since it has been found to make voices sound more feminine. Since dominance judgments are often based on vocal characteristics that are opposite of those used to make judgments of attractiveness, sexiness and youthfulness, we predicted that the dominance scale would elicit very different ratings than the other three. We also anticipated finding differences in the ways in which listeners judge voices on the attractiveness, sexiness and youthfulness scales, since we predicted different effects of voice qualities on these more closely related scales.

Method

Participants

Participants were recruited from an announcement on the Mechanical Turk service from Amazon.com. Only participants from the United States who successfully completed 95% or better of their previous tasks were included in the pool. One hundred and one participants completed the study. Since it is not clear that those with a non-heterosexual orientation would respond in the same way to male and female voices as those with a heterosexual orientation, only heterosexual subjects were included, which eliminated 6 participants, four women and two men. Data were analyzed from the remaining 95 participants (Males = 56), who were paid for their participation. Twenty-four percent of the subjects were between 18 and 25 years of age, 46% were between 26 and 35, 16% were between 36 and 45, 7% were between 46 and 55 and 8% were over 56 years old. Participants were 84% Caucasian, 10% African-American and 6% Asian. The protocol for this study was approved by the institutional review board.

Stimuli

In order to develop our natural stimuli from a single speaker, we began by recruiting four college-age women to record a series of vowels (/a/, /i/ and /u/), words, sentences, and a short passage. They were paid $20 for their participation. They produced each utterance multiple times in one of four voice types: modal, breathy, creaky and girlish. Clips of breathy, creaky and girlish female voices were provided as models the first time they were asked to record stimuli with those types of phonation. They were also asked to produce each utterance several times at their normal pitch as well as with a raised or lowered pitch. The utterances were recorded directly into PRAAT (Boersma & Weenink, 2013) at a 44.1 KHz using a Logitech USB Desktop Microphone. Only tokens of the vowel /a/ were used in this experiment.

The recordings of one woman were discarded because she failed to consistently pronounce the vowel sounds correctly. From each of the other three women, we selected one /a/ vowel from each of the 12 categories (modal, breathy, creaky and girlish at normal, raised and lowered pitches), such that the vowel pronunciation was consistent and each of the categories was easy to identify perceptually. We then did acoustic analyses and pilot testing of the stimuli. Each vowel was analyzed for duration, mean amplitude, mean F0, the frequency of the first three formants, and the amplitude of the first harmonic minus the amplitude of the first formant peak (H1-A1). The latter measure can be used as a marker of breathiness (positive values) and creakiness (negative values) with modal (normal) phonation values in between the other two (Gordon & Ladefoged 2001) Our pilot testing indicated that the utterances from one of the three women received significantly different ratings than the other two speaker's voices from both men and women for each of the four dimensions we were investigating, so we excluded her utterances from further

consideration. For the utterances of the remaining two speakers, the acoustic measurements indicating differences in breathiness and creakiness were more consistent in one set than in the other, so we chose to use that speaker's /a/ syllables for this study. The mean value for breathiness (8.67) was in the predicted positive direction and significantly different from the mean values for modal voice (-4.73), $t(4)=2.16$, $p=.049$, one tailed. The mean value for creakiness (-9.53) was in the predicted negative direction and also significantly different from that of modal voice, $t(4)=2.86$, $p=.023$, one tailed. See Table 1.

Pilot testing also indicated that participants responded strongly to differences in the intensity of the syllables, so we normalized all the syllables for mean intensity in PRAAT. We also determined that the utterances produced when our speakers were asked to speak in a girlish voice did have a higher second formant than their normal voices, but since we also wanted to examine the effect of formant spacing for all three formants on participants' judgments of all types of voices, we made two adjustments to each of the 12 /a/ syllables using the formant shift ratio function in PRAAT. In one case, the formants were generally raised slightly to give the voice a more youthful quality, and in the other, they were lowered, in order to make the speaker sound older. On average, for the stimuli manipulated to sound younger, the first and second formants were raised 10% and the third formant was lowered 25% compared to the original values. For the stimuli manipulated to sound older, the first and second formants were lowered 8% and the third was lowered 10% compared to the original values. None of the formant manipulations created voices that sounded masculine as opposed to feminine, since research has shown that both the fundamental frequency and the formant spacing need to be adjusted in order to make voices sound as

though they had changed gender (Hillenbrand & Clark 2009). See Table 1 for the mean F0, formant values for F1, F2 and F3, and H1-A1 values for the twelve original syllables.

Procedure

After completing the consent form, participants were asked to answer a series of demographic questions, including their gender, age, and relationship status. They were also asked to indicate whether they were using headphones to listen to the stimuli. Approximately 90% of the participants used headphones. In the main part of the study, for each trial, they were instructed to click on the play button to listen to the sound clip and then to answer the question for that trial. For each trial, they heard one of the 36 /a/s and were asked to rate it on a scale of 1-to-9 for dominance, attractiveness, sexiness or youthfulness. The 1 represented an absence of the quality (e.g., not dominant) and the 9 represented an unconditional degree of that quality (e.g., dominant). Participants were instructed to pay close attention to the voice and to use the whole range of the scale in their responses. The study consisted of a total of 144 trials; each participant heard all 36 vowels and rated each on all 4 scales. The order of presentation of the questions was randomized. At four points in the session, participants were asked to identify the scale that had been used in the question before last. These attention questions were inserted in order to make sure participants were not just clicking through the questions without thinking about them. Subjects had to get 3 out of 4 attention questions correct in order to be included in the sample. On average, participants took about 30 minutes to complete the test.

Results

We first conducted a MANOVA on the data from our 95 participants, with two between subject variables, gender and relationship status (single or in a relationship) and four within subject variables, scale (dominance, attractiveness, sexiness, youthfulness) voice quality (modal, girlish, breathy, creaky), pitch (lowered, normal, raised) and formant spacing (original, young, old). Since scale was a significant effect, $F(3,89)=22.83$, $p<.0001$, ??need partial eta here? we chose to analyze the data for each scale separately. Relationship status was not a significant main effect, nor did it enter into any significant interactions in the MANOVA, so it was eliminated as a factor in the separate ANOVAs. Only main effects and interactions that were significant in the MANOVA are reported in the individual ANOVAs for each scale.

For each ANOVA there was thus one between subject variable, gender, and three within subject variables: four levels of voice quality (modal, girlish, breathy and creaky); three levels of pitch (lowered, normal, raised); and three levels of formant spacing (original, young and old).

Mauchly's test indicated that the assumption of sphericity was violated in nearly all the univariate tests. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Dominance

For the dominance ratings, there were significant main effects of voice quality, pitch, and formant spacing. For voice quality, modal voices were rated the highest ($M = 4.32$, $SE = .12$) followed by creaky ($M = 3.80$, $SE = .15$), girlish ($M = 3.57$, $SE = .11$) and breathy ($M = 3.46$, $SE = .11$), $F(2.29, 213.33) = 21.81$, $p < .0001$, partial $\eta 2 = .19$. Planned

comparisons indicated that modal voices were significantly different from all others, $p<.0001$ and that breathy and creaky voices were also significantly different from one another, $p=.023$. For pitch, /a/s with lowered pitch were rated higher for dominance ($M=4.58$, $SE=.13$) than those with normal ($M=3.69$, $SE=.12$) or raised ($M=3.09$, $SE=.11$) pitches, $F(1.39,129.16)=107.01$, $p<.0001$, partial $\eta 2=.54$. Planned comparisons indicated that all of these values were significantly different from one another, $p<.0001$. For formant spacing, the /a/s with formants changed to make the voice sound older were rated more dominant ($M=4.07$,$SE=.12$) followed by /a/s with the original formants ($M=3.80$, $SE=.11$) and then those with formants adjusted to make the voice sound younger ($M=3.49$, $SE=.10$), $F(1.56, 145.48)=27.17$, $p<.0001$, partial $\eta 2=.23$. Planned comparisons indicated that each of these values were significantly different from the other, $p<.0001$. Since all four scales showed the same three main effects, Figure 1 provides a graphic representation of these effects for dominance, as well as for attractiveness, youthfulness and sexiness.

There were four significant two-way interactions that were also significant in the original MANOVA, one of which included gender as a factor, voice quality by gender, $F(2.29, 213.33)=6.55$, $p=.001$, partial $\eta 2=.07$. Post-hoc t-tests comparing male and female responses to each voice quality revealed that women tended to rate creaky voices ($M=4.11$) as more dominant than men did ($M=3.50$), $93)=-2.04$, $p=.05$, Cohen's $d=-0.42$. All other two-way comparisons for gender were not significant.

The remaining two-way interactions included voice quality by formant spacing, $F(5.50, 511.54)=3.52$, $p=.003$, partial $\eta 2=.04$. Modal voice was rated most dominant followed by creaky voice for each of the formant types - original, old and young - but the scores for all four voice qualities were highest for the stimuli with the formants adjusted to

make the voice sound older, followed by those with the original formants and then those with the formants adjusted to make the voice sound younger. Figure S1 in the supplementary materials provides a graphic representation of this two-way interaction for dominance and the other three scales, since all four scales showed the same three two-way interactions that did not involve gender.

There was a two-way interaction of pitch by formant spacing, $F(3.73, 347.11)=2.80$, $p=.029$, partial $\eta 2=.03$. Older voices with lowered pitch were rated most dominant and young voices with a high pitch were rated the least dominant. Figure S2 in the supplementary materials provides a graphic representation of this two-way interaction for dominance and the other three scales, since all four scales showed the same three two-way interactions.

Finally, there was a two-way interaction of voice quality by pitch, $F(4.71,438.33)=25.95$, $p<.0001$, partial $\eta 2=.22$. The effect of lowered pitch on modal, girlish and breathy voice /a/s was pronounced, but it was less pronounced with creaky voice syllables, so that creakiness made a similar contribution to ratings of dominance across all three pitches. Figure S3 in the supplementary materials provides a graphic representation of this two-way interaction for dominance and the other three scales, since all four scales showed the same three two-way interactions.

The three-way interaction between voice quality, pitch and formant spacing was also significant, $F(9.83,914.32)=5.51$, $p<.0001$, partial $\eta 2=.06$. In general, whether the formants had their original values or were modified to sound older or younger, across all four voice quality types, the lower the pitch, the more dominant the rating; however, when the formants were modified to make the voice sound younger, a raised pitch decreased

dominance ratings more dramatically for voices with a girlish or breathy voice quality as compared to those with a modal or creaky quality. Similarly, breathy voices made to sound older that had a raised pitch showed a dramatic decrease in dominance ratings. Figure S4 in the supplementary materials provides a graphic representation of this effect for dominance and the other three scales, since all four scales showed the same significant three-way interaction.

Attractiveness

For the attractiveness ratings, there were three significant main effects, voice quality, pitch and formant adjustment. For voice quality, breathy voices were rated the highest (M = 6.11, SE =.14) followed by girlish (M = 5.01, SE =.12), modal (M = 3.96, SE =.10) and creaky (M = 2.54, SE =.09), $F(2.00,185.68)=346.92$, $p<.0001$, partial $\eta 2=.79$. Planned comparisons indicated that all voice quality types were significantly different from one another, $p<.0001$. For pitch, /a/s with raised pitch were rated higher for attractiveness (M=5.19, SE=.13) than those with normal (M=4.13, SE=.10) or lowered (M=3.90, SE=.10) pitches, $F (1.29, 119.75)=88.78$, $p<.0001$, partial $\eta 2=.49$. Planned comparisons indicated that lowered and normal pitch were significantly different from one another, $p<.001$, and that raised pitch was significantly different from both low and normal, $p<.0001$. For formant spacing, the voices made to sound younger were rated most attractive (M=4.69,SE=.11) followed by those with the original formants (M=4.66, SE=.09) and then voices made to sound older (M=3.87, SE=.11), $F(1.47, 136.22)=52.55$, $p<.0001$, partial $\eta 2=.36$. Planned comparisons indicated that the /a/s with the original formants and those made to sound younger were rated significantly higher for attractiveness than those made to sound older, $p<.0001$, but there was no difference in the attractiveness ratings between the

original formants and /a/s made to sound younger. Figure 1 provides a graphic representation of these three main effects for attractiveness and the other three scales.

There were four significant two-way interactions, one of which involved gender, voice quality by gender, $F(2.00, 185.68)=4.38$, $p=.014$, partial $\eta 2=.05$. Post-hoc t-tests comparing male and female responses to each voice quality revealed that women tended to rate girlish voices as more attractive ($M=5.26$) than the men did ($M=4.76$), $t(93)=-2.15$, $p=.033$, Cohen's $d=-0.44$. There were no other significant effects involving gender.

The remaining two-way interactions included voice quality by formant spacing, $F(5.56, 518.60)=34.27$, $p<.0001$, partial $\eta 2=.27$. Although in general, breathy and girlish voices were rated higher for attractiveness than modal and creaky voices, when the formants were modified to make the voice sound older, attractiveness scores went down and the gap between the highest rated voice quality (breathy) and the lowest (creaky) was considerably narrowed, compared to /a/s with the original formants or those modified to sound younger. See Figure S1 in the supplementary materials.

There was a two-way interaction of pitch by formant spacing, $F(3.56, 331.58)=17.03$, $p<.0001$, partial $\eta 2=.16$. Voices with raised pitch were rated highest for attractiveness for younger, original and older voices, with normal pitch rated next for attractiveness in the case of original and old voices, whereas lowered pitch was rated next for young voices. The effect of pitch changes on the ratings was less pronounced with young voices as compared to both original voices and those made to sound older. See Figure S2 in the supplementary materials.

Finally, there was a two-way interaction of voice quality by pitch, $F(5.178,481.60)=24.60$, $p<.0001$, partial $\eta 2=.21$. Raised pitch voices were rated highest for

modal, girlish, breathy and creaky voices, but the next highest rating for breathy voices was for the lowered pitch voice, whereas it was for the normal pitched voice in the three other cases. The effect of pitch differences on attractiveness ratings was more pronounced in the case of modal and creaky voices and less pronounced for girlish and breathy voices. See Figure S3 in the supplementary materials.

The three-way interaction between voice quality, pitch and formant spacing was also significant, $F(9.58,890.69)=13.91$, $p<.0001$, partial $\eta 2=.13$. In most cases, breathiness ratings were highest of the four voice qualities and whether the formants had their original values or were modified to sound older or younger, across all four voice quality types, the ratings were highest for /a/s with raised pitches followed by those with normal pitch; however, when the formants were modified to make the voice sound younger, the highest attractiveness ratings were given to /a/s with lowered pitches and breathy voice quality. The voice rated the most attractive of all was one made to sound younger, with a breathy voice quality and lowered pitch. See Figure S4 in the supplementary materials.

Sexiness

For sexiness ratings, there three significant main effects, voice quality, pitch and formant adjustment. For voice quality, breathy voices were rated the highest (M = 5.91, SE =.15) followed by girlish (M = 4.33, SE =.13), modal (M = 3.61, SE =.11) and creaky (M = 2.43, SE =.10), $F(2.04,189.57)=303.46$, $p<.0001$, partial $\eta 2=.77$. Planned comparisons indicated that each of the voice qualities was significantly different from each other, $p<.0001$. For pitch, /a/s with raised pitch were rated higher for sexiness (M=4.39, SE=.15) than those with normal (M=3.99, SE=.10) or lowered (M=3.83, SE=.11) pitches, $F(1.26, 117.60)=11.56$, $p<.0001$, partial $\eta 2=.11$. Planned comparisons indicated that voices with

raised pitch were significantly different from those with normal or lowered pitch, $p<.001$, and that lowered voices were significantly different from normal ones, $p=.032$. For formant spacing, the stimuli with formants modified to make the voice sound younger were rated more sexy ($M=4.44$,$SE=.13$) followed by those with the original formants ($M=4.36$, $SE=.10$) and then those with the formants modified to make the voice sound older ($M=3.42$, $SE=.11$), $F(1.28, 118.74)=58.69$, $p<.0001$, partial $\eta 2=.39$. Planned comparisons indicated that older voices were significantly different from original and younger voices, $p<.0001$, but original and younger voices did not differ. See Figure 1.

There were three significant two-way interactions for sexiness. The first was voice quality by formant spacing, $F(5.25, 488.29)=34.27$, $p<.0001$, partial $\eta 2=.20$. Although breathy voices were rated highest for sexiness followed by girlish, modal and creaky across all three formant types – original, younger and older –breathiness greatly increased the sexiness ratings for syllables both with the original formants and those with the formant manipulated to make the speaker sound younger. As was the case for this significant interaction for attractiveness ratings, sexiness scores went down when the formants were adjusted to make the voice sound older, and the gap between the highest rated voice quality (breathy) and the lowest (creaky) was considerably narrowed, compared to /a/s with the original formants or those modified to sound younger. See Figure S1 in the supplementary materials.

There was a two-way interaction of pitch by formant spacing, $F(3.57, 497.34)=16.83$, $p<.0001$, partial $\eta 2=.15$. For syllables with the original formants and those that were modified to make the voice sound older, syllables with raised pitches were rated most sexy, followed by normal then lowered pitches; however, for formant manipulations

that made the voices sound younger, lowered pitches were rated second in sexiness after raised pitches. Finally, for syllables where the formants were changed to make the voices sound older, the scores for sexiness were generally lower overall and the gap between raised and lowered pitches was the greatest. See Figure S2 in the supplementary materials.

Finally, there was a two-way interaction of voice quality by pitch, $F(5.35,497.35)=14.44$, $p<.0001$, partial $\eta 2=.13$. Breathy voices were rated very highly for sexiness across all three pitch types and creaky voices were rated very low across all three pitch types. For both girlish and breathy voices, the normal pitch voices were rated highest, then raised and then lowered. For creaky voices, raised pitches were rated most sexy, followed by normal and lowered pitches. For modal voices, raised pitches were also rated most sexy, but then lowered pitch was rated next most sexy, followed by normal pitch. See Figure S3 in the supplementary materials.

The three-way interaction between voice quality, pitch and formant spacing was also significant for sexiness, $F(12.00,878.80)=10.80$, $p<.0001$, partial $\eta 2=.10$. In eight out of twelve cases, across the three types of formant spacing and four voice qualities, high-pitched voices tended to get the highest sexiness ratings, followed by normal and lowered pitch. In most cases, breathiness ratings were highest of the four voices; however, in the case of the original voice or the one manipulated to sound younger, breathy /a/s with lowered pitches were rated very highly as well. See Figure S4 in the supplementary materials.

Youthfulness

For the youthfulness ratings, there were three main effects, voice quality, pitch and formant adjustment For voice quality, girlish voices were rated the highest for youthfulness

(M = 5.93, SE =.10) followed by breathy (M = 5.28, SE =.11), modal (M = 4.92, SE =.10) and creaky (M = 3.93, SE =.13), $F(2.05,190.89)=99.04$, $p<.0001$, partial $\eta 2=.52$. Planned comparisons indicated that modal voice differed significantly from breathy, $p=.003$, and that all the other two-way comparisons among qualities were significant at $p<.0001$. For pitch, /a/s with raised pitch were rated higher for youthfulness (M=6.43, SE=.10) than those with normal (M=5.06, SE=.09) or lowered (M=4.29, SE=.11) pitches, F (1.57, 149.56)=387.73, $p<.0001$, partial $\eta 2=.81$ Planned comparisons indicated that each pitch was significantly different from the other, $p<.0001$. For formant spacing, the stimuli with formants adjusted to make them sound younger were rated more youthful (M=5.69,SE=.13) followed by /a/s with the original formants (M=4.36, SE=.10) and then those with the formants adjusted to make them sound older (M=3.42, SE=.11), $F(1.45, 145.90)=122.99$, $p<.0001$, partial $\eta 2=.57$. Planned comparisons indicated that each of the different formant patterns was different from the other, $p<.0001$. See Figure 1.

There were three significant two-way interactions for youthfulness. The first was voice quality by formant spacing, $F(3.32, 495.88)=10.67$, $p<.0001$, partial $\eta 2=.10$, revealed that whereas girlish voices were rated highest for youthfulness across all three formant types, followed by breathy, modal and then creaky voice, the gap between the highest and lowest rated four voice quality types was greatest for syllables whose formants had been manipulated to make the voice sound young. See Figure S1 in the supplementary materials.

There was also a two-way interaction of pitch by formant spacing for youthfulness, $F(3.57, 497.34)= 3.67$, $p<.008$, partial $\eta 2=.04$. Again, the pattern across all three formant types was the same: Voices with raised pitch were rated highest for youthfulness followed by those with normal and then those with lowered pitch. Although the voice rated most

youthful had both raised pitch and formants adjusted to make it sound younger, the voice with the original formants showed the greatest difference between the ratings for raised and lowered pitches. See Figure S2 in the supplementary materials.

Finally, there was a two-way interaction of voice quality by pitch, $F(5.13, 558.00)=29.97$, $p<.0001$, partial $\eta 2=.23$. For all four voice quality types, raised pitches gave the highest youthfulness ratings; however, creaky voice in the context of raised pitch was rated quite highly for youthfulness. See Figure S3 in the supplementary materials.

The three-way interaction between voice quality, pitch and formant spacing was also significant for youthfulness, $F(9.11, 846.79)=5.66$, $p<.0001$, partial $\eta 2=.06$. In nine out of twelve cases for the different voice qualities and formant spacing types, syllables with raised pitch were rated highest for youthfulness, followed by those with normal pitch and then those with lowered pitch. The effect of raised pitch was particularly strong in the case of creaky voice, which across all three formant spacing types – original, young and old – made strong contributions to judgments of youthfulness. See Figure S4 in the supplementary materials.

Discussion

Our results provide some new insights as to how voice qualities such as modal voice, breathiness, creakiness and girlishness affect listeners' judgments of dominance, attractiveness, sexiness and youthfulness in the female voice. They also largely corroborate findings in the literature for the effects of formant spacing and pitch differences on such judgments. We also found that the patterns for judgments of sexiness and attractiveness tended to be similar for our stimuli and that they differed somewhat from judgments of

youthfulness. High ratings for dominance were generally based on different values for formant spacing, pitch and voice quality than the other three scales.

The effects of voice quality differences on judgments of dominance were largely as predicted: Girlish voices were rated low for dominance, as were breathy voices. Indeed, voices made to sound younger with raised pitch and a girlish or breathy voice quality were rated the least dominant. Although modal voice was rated most dominant, creaky voice quality was rated more dominant than breathy voice. Women rated creaky voice significantly higher for dominance than men. Since the use of creaky voice appears to be more on the rise among young American women as compared to young men (Abdelli-Beruh et al. 2014; Yuasa 2010), using it to signal authority and the ability to recognize that signal may be gender linked. We also found that the contribution that creakiness made to judgments of dominance was less affected by whether the pitch was raised or lowered than were other voice qualities, which is not surprising, given that the mean fundamental frequency for our creaky stimuli was lower than those for stimuli with the other three voice qualities (See Table 1).

In terms of the effects of pitch and formant spacing, our results corroborate those in the literature. For both male and female listeners, our female voice sounded relatively dominant when it was low pitched with closer formant spacing adjusted to make the voice sound as though it was coming from an older speaker.

In terms of attractiveness judgments, each of the four voice qualities had a significantly different effect from the others, with breathiness rated the highest followed by girlish, modal and then creaky voice. Since breathiness has been associated with feminine

voices (Van Borsel et al. 2009) and with intimacy (Laver 1980), it makes sense that listeners would find the quality attractive, as we predicted. Although attractiveness is usually associated with raised pitch (Collins & Missing 2003; Feinberg et al. 2008; Jones et al. 2008) or normal pitch (Fraccaro et al. 2013), in our study attractiveness ratings for breathy voice remained high even when breathiness was associated with a lowered pitch. Indeed, the young, breathy, lowered pitch /a/ received the highest attractiveness rating of all the syllables. This pattern of results suggests that breathy voices are heard as very attractive, unless they are associated with formant features that suggest greater age, and that they also provide a context in which lowered, rather than raised pitch, is perceived to be very attractive. Other studies have found an association between women's use of lower pitch and their desire to sound attractive or sexy. Hughes et al. (2010) found that women used a lower pitch when speaking to an attractive opposite-sex partner, and listeners found the lower pitched voices more pleasant than those produced to an unattractive target. Tuomi and Fisher (1979) found that women lowered their voices when asked to sound sexy. Some actresses are noted for their husky, breathy, low voices, such as Scarlett Johansson and Kathleen Turner. Indeed, this type of voice is considered stereotypically sexy in our society (Karpf 2006). Hughes et al. (2010) suggest that the use of a deeper voice "may be a learned behavior based on sexual voice stereotypes rather than actual vocal characteristics of attractiveness" (p. 164).

We found that women rated girlish voices significantly higher for attractiveness than did men. It may be that men implicitly link sexiness and attractiveness in women's voices and women do not, so a very girlish voice would be less attractive to men because it would sound too young. We also found that when the voice was modified to sound older,

there were smaller differences in how the four voice qualities were rated for attractiveness and that creakiness in general gave rise to low attractiveness ratings, as predicted.

Our findings also largely corroborate results in the literature for effects of pitch and formant spacing for attractiveness judgments. In general, a higher pitch leads to higher attractiveness ratings (Feinberg, DeBruine, Jones & Perrett 2008; Puts, Barndt, Welling, Dawood & Burriss 2011). Even though in our results lowered pitch was considered attractive in the context of breathy voice, overall, for both male and female listeners, pitch differences made a significant contribution to judgments of attractiveness, with raised pitches rated more attractive than normal pitches and normal pitches rated more attractive than lowered pitches. Greater formant dispersion as is found with youthful voices has also been associated with higher attractiveness ratings (Puts et al. 2011). In our data, stimuli with formants adjusted to make the speaker sound younger were rated as more attractive than those with the original formants or those changed to make the voice sound older, although there was no difference in the ratings between the latter two types of voices. Although in general, raised pitch was rated most attractive, followed by normal and then lowered pitch, when the voices were modified to sound younger, lowered and normal pitch voices were rated equally attractive, although still not as attractive as voices with raised pitch. As noted above, when the voice was modified to sound younger and also had a breathy voice quality, lower pitch was rated more attractive than other pitches.

The results for sexiness were quite similar to those we found for attractiveness. Despite the similarity in the patterns of attractiveness and sexiness responses, the ratings for the latter tended overall to be lower. In terms of the voice qualities, each had a significantly different effect from the others on sexiness judgments, with breathiness

getting the highest raitngs followed by girlish, modal and then creaky voice. We also found that when the voice was modified to sound older, there were smaller differences in how the four voice qualities were rated for sexiness and that creakiness was in general associated with low sexiness ratings, as predicted. As was the case for the attractiveness scale, lowered pitch also enhanced sexiness ratings when the voice was original or modified to sound younger and also had a breathy voice quality. In that case, lowered pitch led to those /a/s receiving the high sexiness ratings, although the other breathy /a/s in the younger and original voices at all pitches were also highly rated, which suggests, as in the case of attractiveness, that breathy voices are heard as sexy, unless they are associated with formant patterns that suggest greater age.

Again, as with attractiveness, aspects of our results support the findings in the literature for pitch and formant spacing. Even though lowered pitch was considered sexy in the context of breathy voice, overall, for both male and female listeners, pitch differences made a significant contribution to judgments of sexiness, with raised pitches rated higher than normal and normal rated higher than lowered, and stimuli with formants adjusted to make the speaker sound younger and those with the original formants were rated as more attractive than those changed to make the voice sound younger, although there was no difference in the ratings between the first two types of voices.

Finally, the results for youthfulness patterned slightly differently than those for attractiveness and sexiness. In particular, in terms of voice qualities, girlish voices were rated most highly for youthfulness, as predicted, followed by breathy, then modal, then creaky, all of which were significantly different from one another. Creaky voice was rated relatively highly for youthfulness when it was found in the context of a raised pitch, unlike

the case for attractiveness and sexiness ratings, which dropped considerably in the presence of creaky voice. This finding suggests that the recent increased use of creaky voice or vocal fry by young American women (Wolk et al. 2011; Yuasa 2010) has had an effect on how listeners in general interpret that particular voice quality in young women's voices, at least in the context of a high fundamental frequency. Again, unlike the case of for attractiveness and sexiness ratings, breathy voice did not provide a context in which lowered pitch was highly rated. For youthfulness judgments, such voices received relatively low youthfulness ratings.

The results for youthfulness in terms of pitch and formant spacing were similar to those for attractiveness and sexiness. Raised pitch voices were rated highest for youthfulness followed by normal and then lowered pitch. Formant spacing also significantly affected youthfulness judgments, but unlike the case for attractiveness and sexiness, all three voices – younger, original and older – were significantly different from one another.

Our analyses also indicated that the four scales were significantly different from one another. Overall, the pitch and formant characteristics that contributed to perception of dominance in the voice were distinct from those that contributed to the other three scales of attractiveness, sexiness and youthfulness. Ratings of dominance were greatest for older voices with lowered pitch and modal or creaky voice. In comparison, in the other three scales, pitch and formant spacing patterned in the opposite way, so that, in general, younger voices with higher pitch were generally rated more attractive, sexy and youthful than those with normal pitch which were in turn rated more highly than those with lowered pitch. Dominance ratings were also lower overall than were the ratings for the other three

scales, which is perhaps to be expected since our speaker was a young woman and previous research has found that youthfulness in the female voice is less associated with dominance than in the male voice (Berry 1992).

The contribution made by the four voice qualities was very similar for attractiveness and sexiness, with breathiness making a substantial contribution to such judgments in the majority of cases. Although the response patterns were similar, the scores for sexiness were generally lower overall than those for attractiveness. One possible explanation for this finding is that ratings of sexiness might be greater for more extended speech samples, where effects of speaking rate and greater differences in intonation contours (Kreiman & Sidtis 2013) can be detected. Finally, we only found differences in the responses of our male and female participants in the case of the dominance scale (where women rated creaky voices more dominant than men) and the case of the attractiveness scale (where women rated girlish voices more attractive than men).

Formant spacing in vocalizations is a cue that animals use to assess the size of potential rivals (e.g., Charlton, Whisson & Reby 2008), and although it is a less reliable cue to size, F0 may also be used in this way, even by humans assessing animal size (Taylor, Reby & McComb 2011). Both cues are very old in evolutionary terms and formant spacing appears to be a very reliable source of information. Fundamental frequency is a less reliable cue to size than formant spacing because it is more under voluntary control. Women in different countries can choose to speak with a relatively high F0 (e.g., in Japan) where such a choice is preferred by listeners or a with relatively low F0 (e.g., The Netherlands) where medium or low voice pitches are preferred (van Bezooijen 1995). It is

likely that the effects of specific pitch values on voice judgments would be quite different in those countries.

Voice quality cues, however, for the most part, appear to be learned associations, used for either linguistic or social signaling purposes, which suggests that the ways in which they influenced listener judgments of how dominant, attractive, sexy or youthful a voice sounded might be quite different for members of other linguistic groups. It would be interesting to see how participants from other countries, where the use of breathy voice or vocal fry may be less prevalent among women, assess the stimuli used in this study

The current study has a number of limitations that could be eliminated in future research. Although producing acoustically consistent breathy and creaky utterances can be difficult for speakers, testing our findings with such stimuli would be a good first step. In addition, it would be useful to develop similar stimuli from male speakers to see to how these voice quality differences are interpreted as signs of dominance, youthfulness, attractiveness or sexiness in the context of men's speech.

References

Abdelli-Beruh, N. B., Wolk, L., & Slavin, D. (2014). Prevalence of vocal fry in young adult male American English speakers. *Journal of Voice*, *28*, 185-190.

Anolli, L. & Cicero, R. (2002). Analysis of the vocal profiles of male seduction: From exhibition to self-disclosure. *Journal of General Psychology*, *129*, 149-169.

Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, *3*, 682-684.

Batstone, S. & Tuomi, S. K. (1981). Perceptual characteristics of female voices. *Language and Speech*, *24*, 111-123.

Berry, D. S. (1992). Vocal types and stereotypes: Joint effects of vocal attractiveness and vocal maturity on person perception. *Journal of Nonverbal Behavior*, *16*, 41-54.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer [Computer program]. Version 5.3. 03. 2011.

Borkowska, B. & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*, 55-59.

Bruckert, L., Liénard, J., Lacroix, A., Kreutzer, M., & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society*, *273*, 83-89.

Charlton, B. D., Whisson, D. A., & Reby, D. (2013). Free-ranging male koalas use size-related variation in formant frequencies to assess rival males. *PLOS ONE*, *8*, e70279.

Collins, S. & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour, 65,* 997-1004.

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour, 60,* 773-780.

Crystal, D., & Quirk, R. (1964). *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton.

Crystal, D. (1975). *The English tone of voice*. London: Edward Arnold.

Daniel, H. J. III & McCabe, R. B. (1992). Gender differences in the perception of vocal sexiness. In J. v. d. Dennen (Ed.), *The nature of the sexes: the sociobiology of sex differences and the 'battle of the sexes'.* Groningen: Origin Press.

Esling, J. (1978). The identification of features of voice quality in social groups. *Journal of the International Phonetic Association, 8,* 18-23.

Farley, S. D., Hughes, S. M., & LaFayette, J. N. (2013). People will know we are in love: Evidence of differences between vocal samples directed toward lovers and friends. *Journal of Nonverbal Behavior, 37,* 123-138.

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception, 37,* 615-623.

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., & Perretta, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior, 49,* 215-222.

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D.M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*, 561-568.

Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., Little, A. C., & 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.

Fraccaro, P. J., O'Connor, J. J. M., Reb, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour, 85*, 127-136.

Gordon, M., & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics, 29*, 383-406.

Hageman, W. (2013, July 24) Stop! You're hurting my ears! Retrieved from http://www.chicagotribune.com/lifestyles/sc-fam-0723-voice-control-20130723-story.html

Henton, C.G. & Blandon, R.A.W. (1985) Breathiness in normal female speech: Inefficiency versus desirability. *Language & Communication*, *5*, 221-227.

Henton, C.G. & Blandon, R.A.W. (1988) Creak as a sociophonetic marker. In L. Hyman & C. Li (eds.) *Language, speech and mind: studies in honour of Victoria A. Fromkin.* New York: Routledge.

Hillenbrand, J. M., & Clark, M. J. (2009). The role of f 0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics, 71*, 1150-1166.

Hughes, S. M., Dispenza, F., & Gallup, G. G., Jr. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behav*ior, *25*, 295-304.

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.

Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior,38*, 107-127.

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., Vukovica, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behavior, 79*, 57-62.

Karpf, A. (2006). *The human voice: How this extraordinary instrument reveals essential clues about who we are*. New York: Bloomsbury Publishing..

Kreiman, J., & Sidtis, D. (2013). *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. West Sussex, UK: John Wiley & Sons.

Kimble, C. E. & Seidel, S. D. (1991). Vocal signs of confidence. *Journal of Nonverbal Behavior, 14*, 99-105.

Ladefoged, P. & Harshman, R. (1979). Formant frequencies and movements of the tongue. In Björn Lindblom and Sven Öhman, (eds.) *Frontiers of Speech Communication Research*. NewYork: Academic Press, 25–34.

Laver, John. (1980). *The phonetic description of voice quality*. Cambridge Studies in Linguistics. London: Cambridge University Press.

Mendoza-Denton (2007). Homegirls remembered: Memorializing practices linking language and materiality among California latina/o gang-involved youth. In P. Hodkinson and P. Deicke, *Youth Cultures: Scenes, Subcultures, and Tribes*, 123-147, NY: Routledge.

Pittam, J. (1987). Listeners' evaluations of voice quality in Australian English speakers. *Language and Speech, 30*, 99-113.

Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior, 26*, 388-397.

Puts, D. A., Barndt, J. L., Welling, L. L. M., 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., Doll, L.M. & Hill, A.K. (2014) Sexual selection on human voices. In V.A. Weekes-Shackelford and T.K. Shackelford (eds.), *Evolutionary Perspectives on Human Sexual Psychology and Behavior*. New York: Springer.

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.

Quenqua, D. (2012, February 27). They're, like, way ahead of the linguistic currrrve. *The New York Times, 28*.

Riding, D., Lonsdale, D., & Brown, B. (2006). The effects of average fundamental frequency and variance of fundamental frequency on male vocal attractiveness to women. *Journal of Nonverbal Behavior, 30*, 55-61.

Taylor, A. M., Reby, D., & McComb, K. (2008). Human listeners attend to size information in domestic dog growls. *The Journal of the Acoustical Society of America*, *123*, 2903-2909.

Tuomi, S. & Fisher, J. (1979). Characteristics of simulated sexy voice. *Folia Phonist*, 31, 242-249.

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and speech*, *38*, 253-265.

Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice, 23*, 291-294.

Wolk, L., Abdelli-Beruh, N. B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, *26*, e111-e116.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, *85*, 315-337.

Table 1

Acoustic Measures for the Twelve Original Syllables

Voice Type and Pitch	F0 Mean	F1	F2	F3	H1-A1
Modal lowered	180	685	1159	3121	-2.8
Modal normal	221	748	1253	3280	-7.2
Modal raised	286	789	1356	3198	- 4.2
Girlish lowered	189	834	1396	3279	-2.6
Girlish normal	217	888	1476	3371	-1.5
Girlish raised	285	944	1464	3091	-7.9
Breathy lowered	183	824	1463	3153	17.2
Breathy normal	216	871	1404	3227	11.9
Breathy raised	283	855	1513	3340	-3.1
Creaky lowered	132	826	1386	3070	-7.4
Creaky normal	179	756	1367	3081	-10.6
Creaky raised	179	756	1367	3081	-10.6

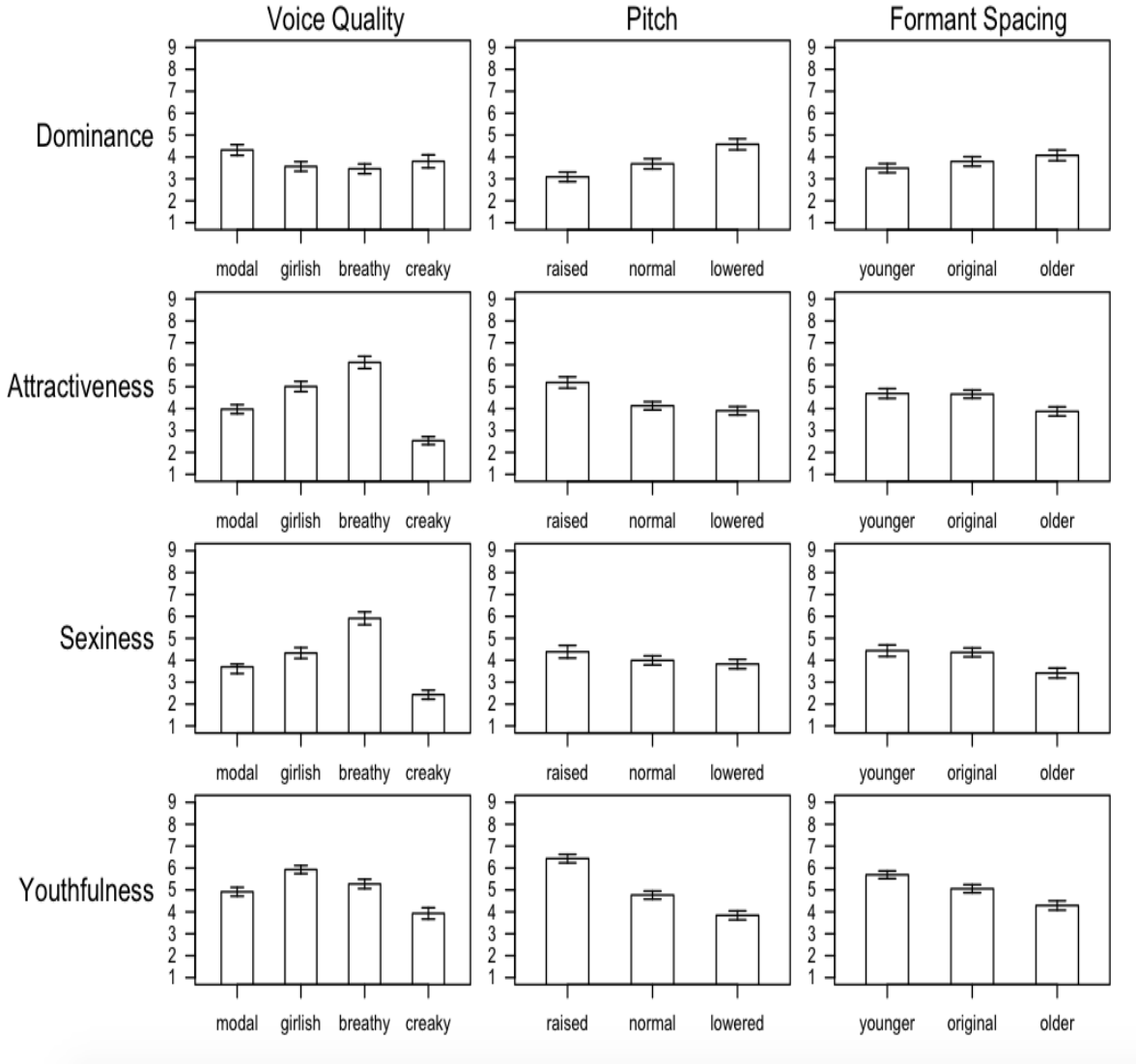

Figure 1 Main effects of voice quality, pitch and formant spacing for dominance, attractiveness, sexiness and youthfulness judgments. Error bars represent 95% confidence intervals.

Supplementary materials

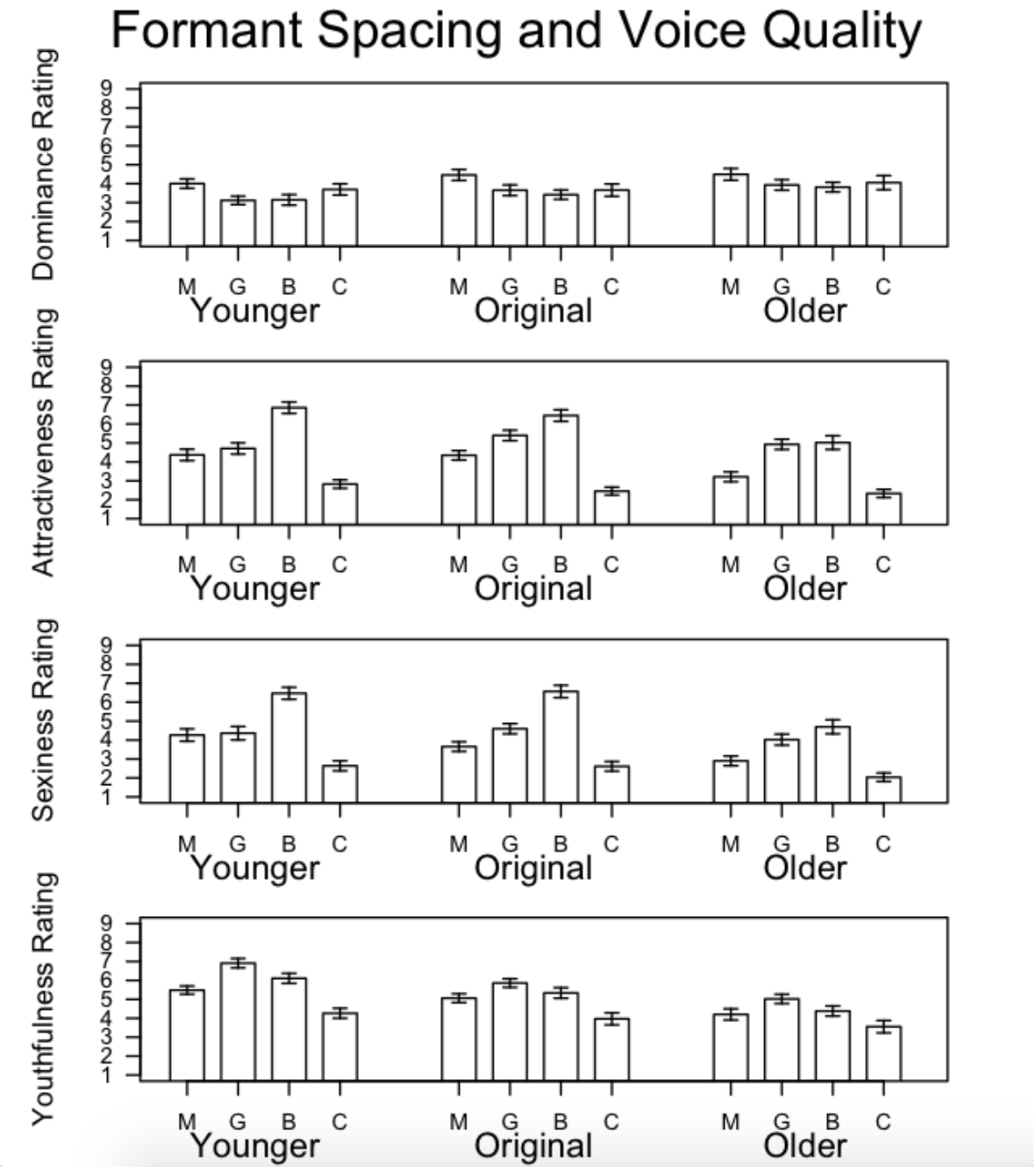

Figure S1 Interactions of formant spacing and voice quality for dominance, attractiveness, sexiness and youthfulness judgments. Error bars represent 95% confidence intervals.

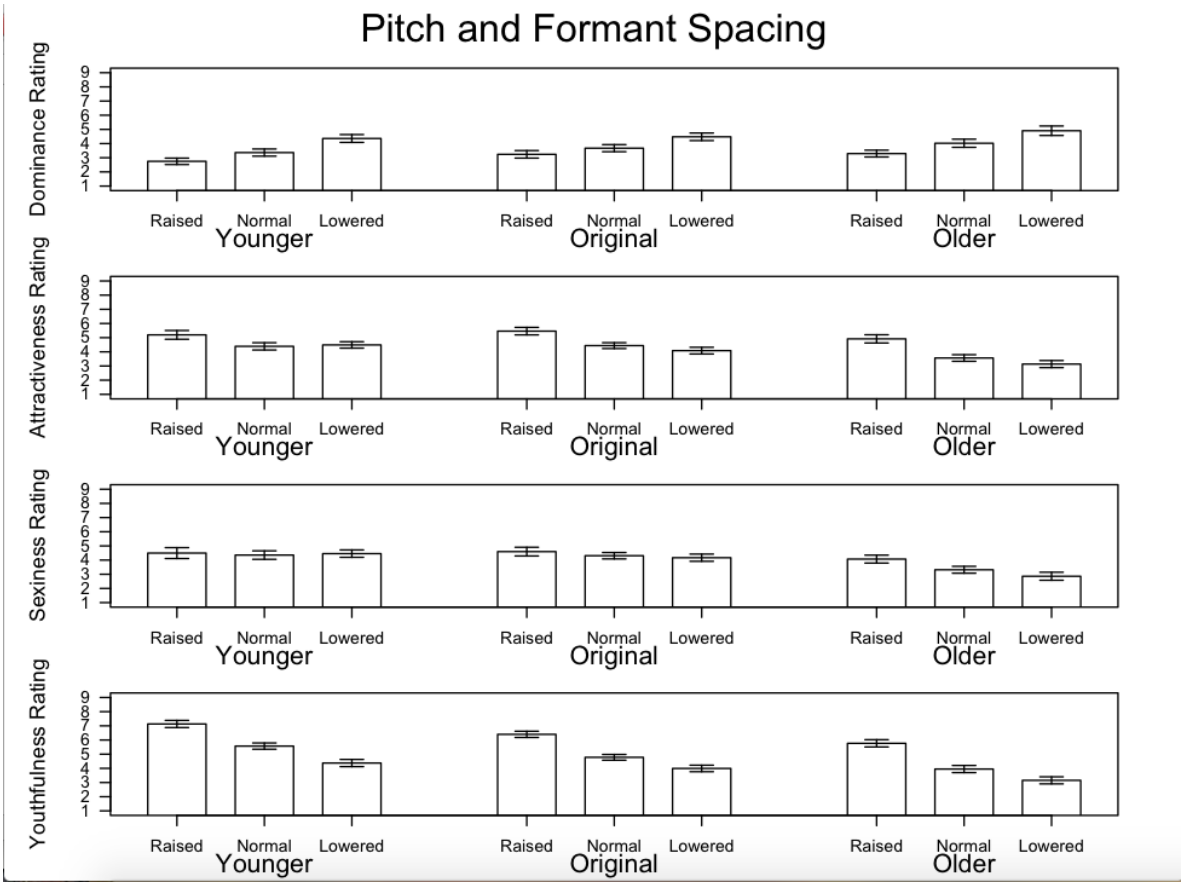

Figure S2 Interactions of pitch and formant spacing for dominance, attractiveness, sexiness and youthfulness judgments. Error bars represent 95% confidence intervals.

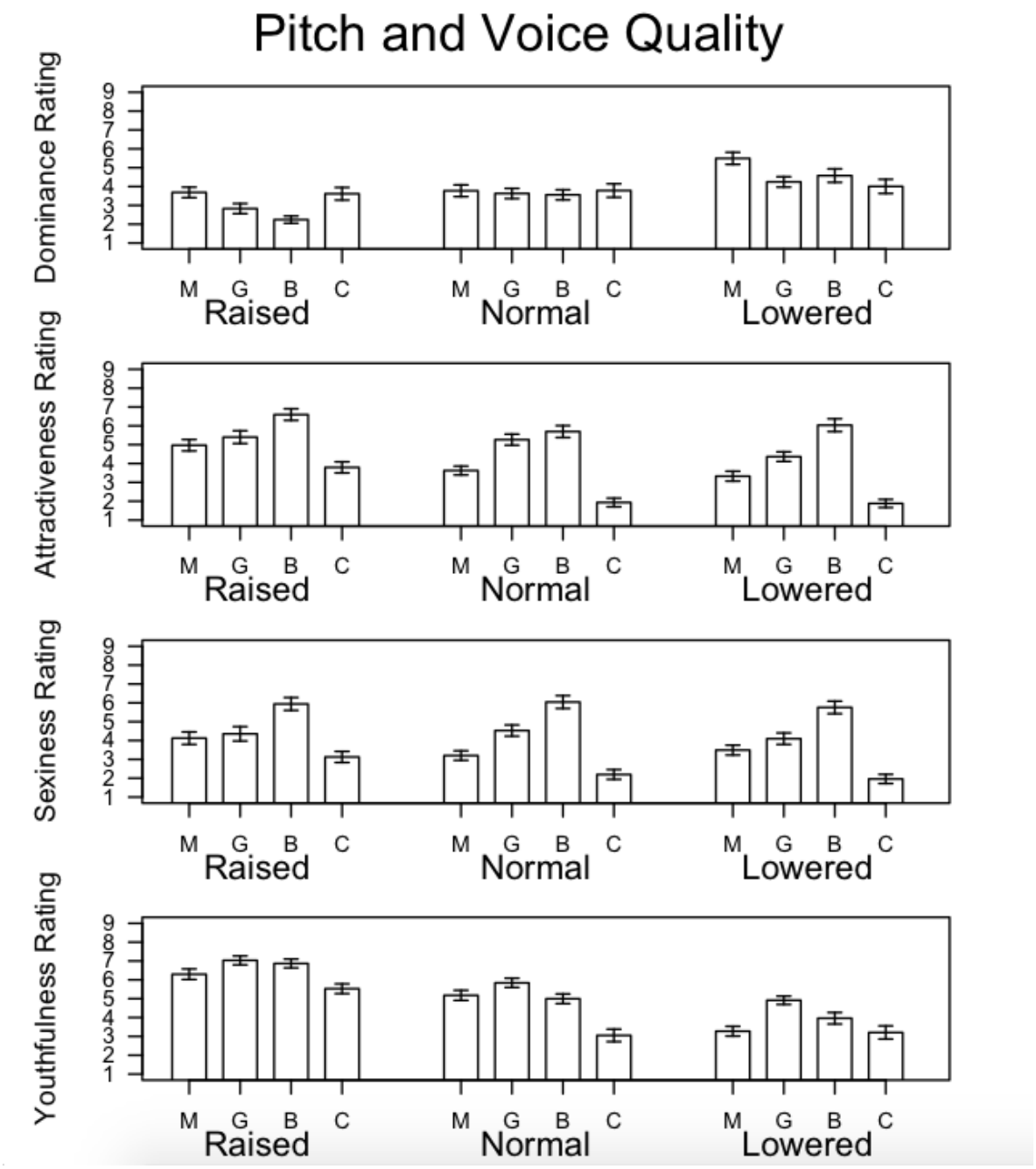

Figure S3 Interactions of pitch and voice quality for dominance, attractiveness, sexiness and youthfulness judgments. Error bars represent 95% confidence intervals.

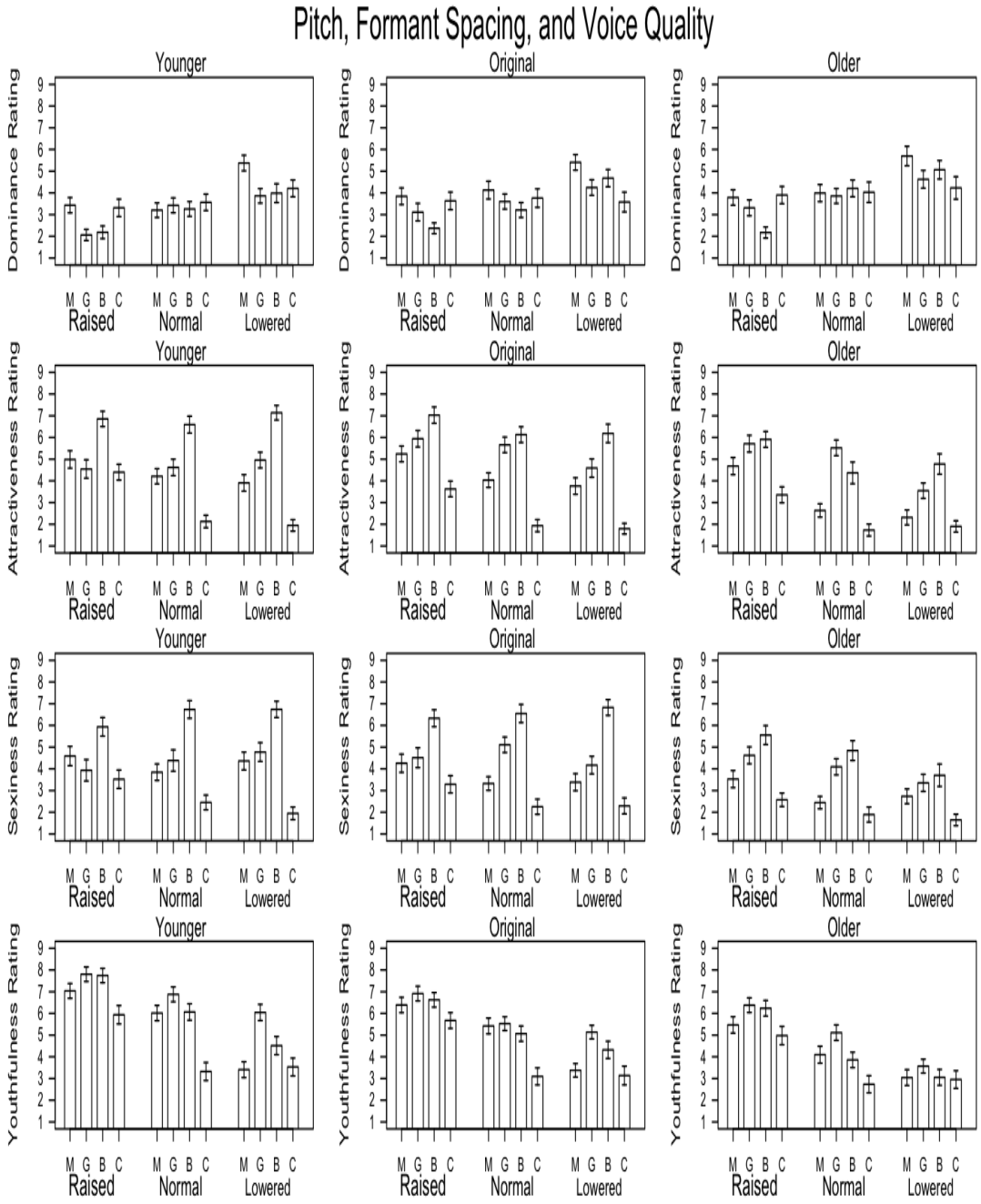
Pitch, Formant Spacing, and Voice Quality
Younger
Original
Older
Dominance Rating
Attractiveness Rating
Sexiness Rating
Youthfulness Rating
M G B C
Raised
Normal
Lowered

Figure S4 Interactions of pitch, formant spacing and voice quality for dominance, attractiveness, sexiness and youthfulness judgments. Error bars represent 95% confidence intervals.