J Nonverbal Behav (2012) 36:79–93
DOI 10.1007/s10919-011-0123-4

ORIGINAL PAPER

Acting Gay: Male Actors Shift the Frequency Components of Their Voices Towards Female Values When Playing Homosexual Characters

Valentina Cartei · David Reby

Published online: 14 October 2011

Abstract The purpose of this study was to investigate whether actors playing homosexual male characters in North-American television shows speak with a feminized voice, thus following longstanding stereotypes that attribute feminine characteristics to male homosexuals. We predicted that when playing homosexual characters, actors would raise the frequency components of their voice towards more stereotypically feminine values. This study compares fundamental frequency ($F0$) and formant frequencies (F_i) parameters in the speech of fifteen actors playing homosexual and heterosexual characters in North-American television shows. Our results reveal that the voices of actors playing homosexual male characters are characterized by a raised $F0$ (corresponding to a higher pitch), and raised formant frequencies (corresponding to a less baritone timbre), approaching values typical of female voices. Besides providing further evidence of the existence of an "effeminacy" stereotype in portraying male homosexuals in the media, these results show that actors perform pitch and vocal tract length adjustments in order to alter their perceived sexual orientation, emphasizing the role of these frequency components in the behavioral expression of gender attributes in the human voice.

Keywords Voice · Formant frequency · Fundamental frequency · Gender · Sexual orientation

Introduction

The portrayal of male homosexuals in films and television often follows an "effeminacy stereotype" (Kite and Deaux 1987; Madon 1997), which attributes feminine connotations to adult male homosexuals. Whilst acknowledging that stereotypes affect multiple dimensions of behavior (Blashill and Powlishta 2009), previous literature has focused on the feminization of homosexual characters' mannerisms and lifestyles (Battles and Hilton-Morrow 2002; Blashill and Powlishta 2009; Chung 2007; Linneman 2008; Raley

V. Cartei (✉) · D. Reby
School of Psychology, University of Sussex, Falmer BN1 9QH, UK
e-mail: v.cartei@sussex.ac.uk

and Lucas 2006; Staats 1978), but has overlooked the characters' voices. The present study investigates whether the voices of homosexual characters are feminized, that is, whether actors playing such roles modify (either consciously or unconsciously) their habitual voice towards values characteristic of heterosexual female voices.

According to the source-filter theory (Fant 1960), the production of the human voice is characterized by two successive and independent stages. First the glottal wave is generated in the larynx (the "source") by periodic vibration of the vocal folds. This wave is a complex periodic signal with a fundamental frequency, or $F0$ (equal to the rate of glottal vibration, and responsible for the perceived "pitch" of the voice), and its integer multiple frequencies, the harmonics. As the glottal wave propagates from the larynx to the lips, the vocal tract acts as a filter and selectively amplifies or dampens frequencies, producing spectral peaks called formant frequencies (which affect the perceived "timbre" of the voice). While the source- and filter-related components can vary independently, they are constrained by the dimensions of the vocal apparatus (Fant 1960). Compared to women, men speak with a lower $F0$ (Hollien et al. 1994; Lee et al. 1999; Perry et al. 2001; Rendall et al. 2005; Wolfe et al. 1990), giving them a lower pitch, and also with lower formant frequencies (Busby and Plant 1995; Lee et al. 1999; Perry et al. 2001; Rendall et al. 2005), giving them a more resonant, baritone timbre (Fitch and Giedd 1999). This sexual dimorphism in $F0$ and formant frequencies largely results from testosterone-related changes in size and morphology that occur during male puberty (Busby and Plant 1995). The lengthening of the vocal folds associated with laryngeal growth causes a dramatic drop of $F0$ in adolescent males to one octave lower (100–120 Hz: Simpson 2009) than females' $F0$ (200–220 Hz: Simpson 2009). In parallel, the lengthening of the vocal tract that follows the overall differential body growth (Fitch and Giedd 1999), and that is accentuated by a secondary laryngeal descent in adolescent males, results in a 1.2 ratio of female to male formant frequencies (Hillenbrand et al. 1995).

However, static, biological factors do not explain the entirety of voice gender differences. Pre-pubertal children's voices are dimorphic despite the absence of substantial differences in the morphology or dimensions of the vocal apparatus between sexes (Fitch and Giedd 1999; Titze 1994; Vorperian et al. 2005): boys' voices are consistently characterized by lower formants than girls' (Lee et al. 1999; Vorperian and Kent 2007). This suggests that children acquire gender-specific articulatory behaviors that enable them to mimic the sexual dimorphism in formant frequencies present in adults. More specifically, it has been suggested that children make small adjustments to the length of their vocal tract, thereby altering formant frequency spacing and feminizing or masculinizing their voices (Mattingly 1966; Sachs et al. 1973; Vorperian and Kent 2007).

Similar vocal gestures may continue to play a role in adults, particularly where there is an explicit or implicit drive to accentuate or downplay the biologically determined voice in order to adapt to specific sexual roles and social contexts. Indeed, while the morphological dimorphism accounts for a substantial part of the voice differences between adult males and females, it cannot fully explain the intra-sexual differences in the femininity or masculinity of the voices of individuals (Rendall et al. 2008). Moreover, recent research on homosexual speech presents acoustic data that suggests the involvement of such adjustments in the expression of sexual orientation. While, contrary to popular stereotypes, 'gay speech' does not systematically reflect opposite sex patterns (Munson et al. 2006; Pierrehumbert et al. 2004), homosexual voices do display some characteristics associated with the opposite sex (e.g., sex-specific vowel formant values), even after controlling for body size (height and weight) (Munson and Babel 2007; Munson et al. 2006; Pierrehumbert et al. 2004; Rendall et al. 2008). Further, perceptual studies (Gaudio 1994; Munson and Babel 2007; Smyth et al. 2003) have shown that listeners' rating of the masculinity/femininity of a speaker's voice is correlated with their judgment of the speaker's sexual orientation: voices rated with higher femininity scores are more likely be judged as belonging to homosexual male speakers (and vice versa).

Here we investigate whether actors playing homosexual and heterosexual male characters in American TV modify their voice in line with gender stereotypes. More specifically, we hypothesize that actors playing homosexual characters feminize their voice by (1) increasing their mean fundamental frequency as well as its dynamic variation and (2) raising overall formant frequencies spacing. Increases in $F0$ and $F0$ variation can be achieved by raising the rate of vocal folds vibration, and its variability, and will respectively result in higher pitched and more melodious voices. Increases in formant frequency spacing reflect a shortening of the supralaryngeal vocal tract that can be achieved by spreading the lips, which effectively shortens the anterior end of the vocal tract, and also by raising the larynx, which shortens the posterior end of the vocal tract (Riordan 1977; Titze 1994). An increase in formant spacing results in a less resonant or baritone timbre (Fitch and Giedd 1999).

Method

Selection of Stimuli

We identified actors who played at least one homosexual role and one heterosexual role in American television comedies or dramas, and for which at least one interview was also available. The list of suitable programs (available from the authors upon request) was compiled from Wyatt (2008), the Internet Movie Database (http://www.imdb.com/), and television networks listings.

The characters' audio samples were extracted from randomly selected episodes from Home DVDs (PAL) and TV show recordings. Interviews were selected from talk shows and DVD-extras to match the genre of the actors' selected roles. All audio samples were extracted using iSkySoft DVD Audio Ripper 1.8.2.7 (Wondershare Software Co. Ltd 2010) and saved in WAV format (sample rate 44,100 Hz, bit rate 128 kbps).

For each actor, we used approximately five audio samples (SD = 1.4) from homosexual roles, with average duration 426 s (SD = 210 s), five audio samples from heterosexual roles (SD = .7) with average duration 410 s (SD = 180 s) and five (SD = 1.8) audio samples from interviews, with average duration 531 s (SD = 191 s). The criteria for sample selection were: no background noise (i.e., music, other people speaking), no crowded settings (i.e., office, bar) and no strong emotional content. As expressive speech could not be completely avoided, samples were categorized by two listeners into five categories: emotional neutrality, fear, anger, happiness, and sadness (following Costanzo et al. 1969; Frick 1986; Johnson et al. 1986). The few samples for which agreement was not reached (4%) were discarded. Samples were then selected in order to balance the emotional content across the two acted contexts. In the resulting dataset, homosexual and heterosexual character speech samples contained the same percentage of mild happiness (8%) and mild anger (17%) (Table 1). Samples were also selected to minimize content-dependent phonetic biases, by ensuring that vowels were similarly distributed across the three contexts. Pearson's Chi-squares showed no significant relationship between recording context and vowel type in either comedy $\chi^2(20) = 30.52$, $p > .05$ or drama, $\chi^2(20) = 27.94$, $p > .05$. The resulting data set consisted of a total of 200 samples from 15 actors, 8 actors playing in comedies and 7 playing in dramas. Samples were then randomly assigned to a numeric code and renamed accordingly, to ensure blind analysis.

Table 1 Distribution of emotional content between recording contexts

Emotions	Recording context		
	Homosexual characters (%)	Heterosexual characters (%)	Interviews (%)
Happiness	8	8	6
Neutral	75	75	88
Anger	17	17	6

Acoustic Analyses

All acoustic analyses (extraction of $F0$ contours and formant center frequencies) were conducted with Praat 5.1.19 (Boersma and Weenink 2009) using a custom written script (available from the authors on request). The script allows the experimenter to set all analysis parameters prior to processing, and to modify them manually if necessary (blind to sample and condition), to correct for tracking errors.

Fundamental Frequency

The script uses the built-in autocorrelation algorithm (to Pitch (ac) command) to extract the $F0$ contour, and then computes the mean ($F0mean$) and the standard deviation ($F0$SD). The analysis parameters were set as follows: pitch floor = 65 Hz, pitch ceiling = 300 Hz and time step = .01 s. The coefficient of variation ($F0CV$) was then calculated, as the ratio of SD to the mean. $F0CV$ describes the dispersion of $F0$ in a way that does not depend on its magnitude and thus corrects for correlative increases of $F0$SD with mean $F0$ and accounts for its logarithmic perception by human listeners (Gaudio 1994): voices with large $F0CV$ are perceived as more melodious than those with small $F0CV$ (Devillers and Vasilescu 2003; Hodges-Simeon et al. 2010).

Formants

The script uses Linear Predictive Coding (LPC: 'To Formants (Burg)' command) to estimate the center frequencies of the first four formants ($F1$–$F4$). The analysis parameters were set as follows: maximum number of formants to be extracted = 4, ceiling of the formant search range = 4,000 Hz, and effective duration of the analysis window = .03 s. In order to check the accuracy of formant tracking, the script displays a PRAAT Editor window (narrow band spectrogram with overlaid formant tracks) for each sample. In 12 samples, the tracks of the estimated formants were clearly not aligned with the formants visible in the spectrogram, indicating that the chosen number of poles with LPC analysis was inadequate. The ceiling of the formant search range ("maximum formant" parameter) was raised by 200 Hz-increments to match the formant tracks with the formants on the spectrogram, from the "Formant Settings…" dialogue in the Editor window.

Formant Spacing

Formant spacing (ΔF) is the average interval (in Hz) between each adjacent pair of formants. It is determined by, and inversely correlated to, the length of the vocal tract of the speaker (Titze 1994). We estimated ΔF by modeling the vocal tract as a straight uniform tube closed at the glottis and opened at the lips (full details are given in "Appendix 1").

This method of estimating ΔF is justified by the observation that, although formants vary from vowel to vowel, formant spacing (ΔF) approaches a constant determined by vocal tract length at supra-segmental level (Titze 1994). Furthermore, psychoacoustic experiments varying ΔF have shown that the linear scaling of formant spacing determines the perceived age, size, and gender of human voice by listeners (Feinberg et al. 2005; Pisanski and Rendall 2011; Smith and Patterson 2005).

Statistical Analyses

For each acoustic parameter, a two-way mixed ANOVA was carried out with genre as the group factor (comedy, drama) and context (heterosexual, interview, homosexual) as the repeated factor. Post-hoc analyses were conducted by applying contrasts to study differences between the three contexts and between the two genres. All calculations and graphics were completed using SPSS v.16 for Mac.

Results

Mean Fundamental Frequency

There was a significant main effect of context, $F(2,26) = 19.72$, $p < .001$ (see Fig. 1a). Contrasts revealed that $F0mean$ in homosexual roles ($M = 143.31$, SD = 13.42) was significantly higher than in heterosexual roles ($M = 122.93$, SD = 16.85), $F(1,13) = 38.27$, $p < .001$, and in interviews ($M = 112.75$, SD = 21.38), $F(1,13) = 47.64$, $p < .001$. $F0mean$ was not statistically different between heterosexual roles and interviews $F(1,13) = 2.42$, $p > .05$. A main effect of genre was also found, $F(1,13) = 5.25$, $p = .04$. However, while $F0mean$ was significantly higher in comedy ($M = 150.99$, SD = 12.02) than in drama ($M = 134.52$, SD = 9.12), parameter estimates show that this difference was only statistically significant for homosexual roles, $t(13) = 2.95$, $p = .01$. There was no significant interaction between context and genre, $F(2,26) = .06$, $p > .05$.

Fundamental Frequency Variation

There was a significant main effect of context on $F0$ standard deviation ($F0$SD), $F(2,26) = 12.56$, $p < .001$ (see Fig. 1b). Contrasts revealed that $F0$SD was significantly higher in homosexual roles ($M = 29.99$, SD = 8.20) than in heterosexual roles ($M = 22.33$, SD = 7.65), $F(1,13) = 10.65$, $p = .006$, and interviews ($M = 18.92$, SD = 6.64), $F(1,13) = 25.74$, $p < .001$. $F0$SD was not statistically different between heterosexual roles and interviews $F(1,13) = 2.42$, $p > .05$. There was also a main effect of genre $F(1,13) = 6.11$, $p = .028$, with higher $F0$SD in comedy roles than in drama roles.

A significant main effect of context was also found on normalized $F0$ variation ($F0CV$), $F(2,26) = 4.68$, $p = .018$ (see Fig. 1c). Contrasts revealed that $F0CV$ in actors playing homosexual roles ($M = .21$, SD = .05) was higher than in interviews ($M = .16$, SD = .04); $F(1,13) = 12.79$, $p = .003$ while $F0CV$ was not statistically different between heterosexual characters ($M = .18$, SD = .05) and interviews $F(1,13) = 1.16$, $p > .05$. However, in contrast with $F0$SD, $F0CV$ was not significantly different between homosexual and heterosexual roles, $F(1,13) = 2.87$, $p > .05$. Finally, there was no significant main effect of genre on $F0CV$, $F(1,13) = 2.78$, $p > .05$.

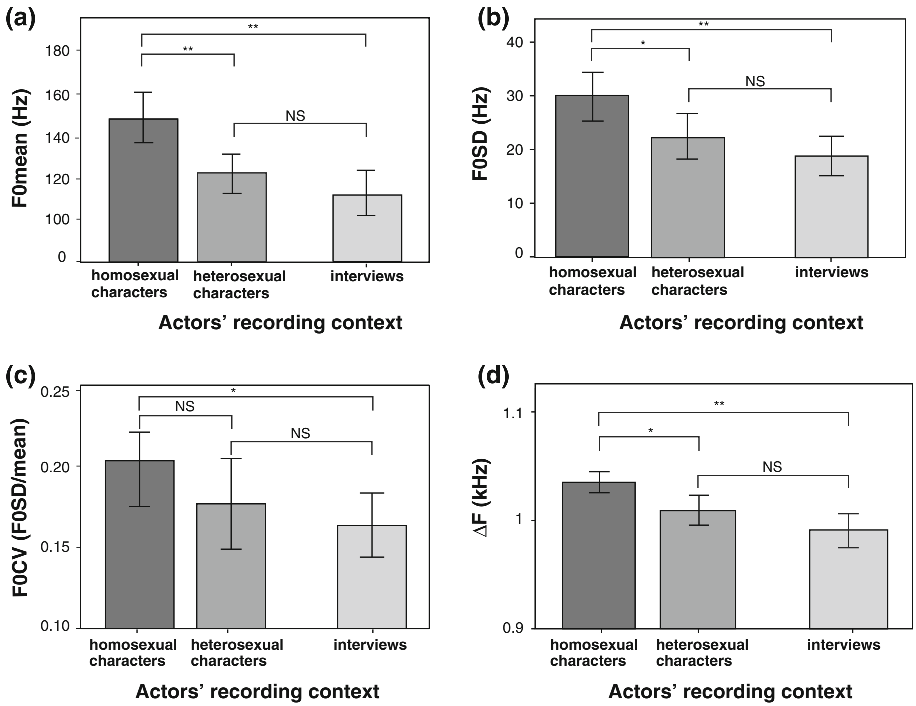

Fig. 1 Mean values of **a** fundamental frequency (*F0mean*)—*error bars*: 95% CI; **b** *F*0 standard deviation (*F*0SD); **c** coefficient of variation (*F0CV*)—*error bars*: 95% CI and **d** formant spacing (ΔF)—*error bars*: 95% CI for actors playing homosexual roles, heterosexual roles being interviewed across genres. *NS* not significant, $*p < .05$, $**p < .001$

Formant Frequencies

The mean formant values ($F1$–$F4$) measured across the contexts are presented in "Appendix 2". When playing homosexual roles, actors' voices were characterized by higher $F1$, $F2$, and $F4$ formants than when playing heterosexual roles ($F1$: $F(1,13) = 13.038$, $p = .003$; $F2$: $F(1,13) = 14.17$, $p = .002$; $F3$: $F(1,13) = 4.47$, $p = .054$; $F4$: $F(1,13) = 7.11$, $p = .019$) or during interviews ($F1$: $F(1,13) = 41.09$, $p < .001$; $F2$: $F(1,13) = 10.62$, $p = .006$; $F3$: $F(1,13) = 20.49$, $p < .001$; $F4$: $F(1,13) = 9.64$, $p = .008$). Contrasts revealed that $F3$ was also higher in homosexual acted speech than in interviews $F(1,13) = 20.49$, $p = .001$, and between the two acting contexts in the comedy genre, $F(1,13) = 8.49$, $p = .012$. Furthermore, in homosexual roles, $F3$ and $F4$ were significantly higher in the comedy genre than in the drama genre ($F1$: $F(1,13) = 2.22$, $p > .05$, $F2$: $F(1,13) = .44$, $p > .05$, $F3$: $F(1,13) = 5.89$, $p = .03$, $F4$: $F(1,13) = 7.01$, $p = .02$).

Formant Spacing

There was a significant main effect of context on ΔF, $F(2.26) = 16.00$, $p = .002$ (see Fig. 1d). Contrasts revealed that actors playing homosexual roles ($M = 1{,}035.16$ Hz, SD = 17.5 Hz) spoke with a higher ΔF, than when playing heterosexual roles ($M = 1{,}009.47$ Hz, SD = 24.94 Hz), $F(1,13) = 14.98$, $p = .002$, or than when interviewed ($M = 991.19$ Hz, SD = 28.33 Hz), $F(1,13) = 30.22$, $p < .001$. While actors playing heterosexual roles spoke with a slightly higher ΔF than when being interviewed, this difference approached significance $F(1,13) = 4.61$, $p = .051$.

Furthermore, parameter estimates showed that while in homosexual roles, ΔF was significantly higher in the comedy genre than in drama, $t(13) = 2.63$, $p = .021$, in heterosexual roles genre had no significant effect on ΔF $t(13) = .28$, $p > .05$. Finally the context by genre interaction had no significant effect on ΔF, $F(2,26) = .56$, $p > .05$.

Discussion

We found that actors playing homosexual characters produced higher pitched, more melodious, and less baritone voices, by respectively increasing the mean $F0$, $F0$ variation, and formant frequency spacing of their voice. To the extent that adult female voices are characterized by higher $F0$ and formants than adult male voices (Titze 1994), these manipulations created voice profiles that were less masculine and more feminine. Moreover the increased $F0$ variation observed in the voice of actors playing homosexual characters suggests that they attempt to increase the melodic quality of their voice, another stereotypical correlate of perceived femininity (Avery and Liss 1994; Henton 1989, 1995; Terengo 1966). These results confirm that the stereotypical portrayal of male homosexuals by the media, which attributes feminine values to their appearance and behavior (Chung 2007), also involves the feminization of their voices. Our observation that actors playing homosexual roles in comedy further accentuated the feminine qualities of their voice is in line with previous research showing that homosexual male characters in comedy draw from stereotypes of femininity (Battles and Hilton-Morrow 2002), as well as being the subject of jokes based on these stereotypes (Battles and Hilton-Morrow 2002; Dow 2001).

The frequency values achieved by actors in the different contexts can be examined in the context of published differences between men and women and between homosexual and heterosexual male speakers. While the $F0$ reported for heterosexual roles (123 Hz) and interviews (113 Hz) was comparable to that reported in male speakers (100–120 Hz: Simpson 2009), for homosexual roles (143 Hz) it remained within the range of male values, but was shifted approximately 40% towards female values (200–220 Hz: Simpson 2009). This is despite the fact that no differences in $F0$ are reported between homosexual and heterosexual male voices (Rendall et al. 2008). Fundamental frequency variability (speech melody) expressed as standard deviation from mean $F0$ ($F0SD$), was approximately 8 Hz higher when actors played homosexual roles (30 Hz) than when they played heterosexual roles (23 Hz) or were being interviewed (19 Hz). When $F0$ variability was expressed as the coefficient of variation ($F0CV$), homosexual acted speech remained characterized by the highest $F0$ variability. However, the difference was only significant when compared to interview recordings. The fact that overall, $F0$ variability was higher in homosexual characters than in heterosexual characters or than in interviews is consistent with stereotypical notions that women's voices are more melodious than men's (Avery and Liss 1994; Henton 1989; Terengo 1966) and that less monotonous voices are perceived as more feminine (Ko et al. 2006). However this stereotype is only partly supported by acoustic studies, which do not consistently identify significant differences in $F0$ variation between sexes (Simpson 2009), nor between homosexual and heterosexual male voices (Gaudio 1994; Munson and Babel 2007; Smyth et al. 2003).

Formant spacing (ΔF) values in actors playing heterosexual roles (1,009 Hz) and interviews (991 Hz) are comparable to heterosexual men's ΔF reported in the literature (1,005 Hz in Feinberg et al. 2005; and 991 Hz, as calculated from $F1$–$F4$ values in

Pisanski and Rendall 2011). The ΔF observed in homosexual roles (1,035 Hz) is approximately 26 Hz higher than in heterosexual roles. While this ΔF represents a 14.8% shift towards normal adult female values (1,167 Hz, as estimated from formant values in Pisanski and Rendall 2011), it remains well within the normal range for adult male speakers. The observed ΔF values for homosexual acted speech are also higher than those reported in the voices of self-identified homosexual speakers (1,005 Hz, as estimated from formant values in Pisanski and Rendall 2011).

The voice stereotype identified here is likely to result from interactions between existing acoustic cues to gender and sexual orientation in non-acted speech, and perceptual and cultural biases affecting audience expectations (Hajek and Howard 2005). Production studies on homosexual male speech have identified a partial shift of frequency-related components towards female values (Gaudio 1994; Munson et al. 2006; Rendall et al. 2008) and voice perception studies have found that self-identified sexual orientation was a strong predictor of how listeners rate speakers' sexual orientation (Gaudio 1994; Munson et al. 2006) and femininity (Munson and Babel 2007; Riordan 1977) from their voice. The likely acoustic bases for such observations can be described in terms of source-filter theory of voice production (Fant 1960). At the level of the source, there are no significant differences in mean $F0$ (Rendall et al. 2008) and $F0$ variability (Gaudio 1994) between heterosexual and homosexual men. However, at the perceptual level, listeners rate male speech with higher $F0$ as more feminine- and gay- sounding and listeners' ratings of $F0$ variability correlate positively with perceived homosexuality in men (Smyth et al. 2003). While a study of vowels /IY/, /UY/, /AA/, /EY/ and /AE/ embedded in four sentences spoken by Chicago-area speakers failed to find significant differences in average formant values between self-identified homosexual and heterosexual male speakers, it showed that the vocalic space was more dispersed in homosexual than heterosexual male speakers (Pierrehumbert et al. 2004). More recently, a study of homosexual males speakers from the St. Paul/Minneapolis metropolitan area found that the /AE/ and /EH/ vowels (embedded in CVC words) were characterized by higher $F1$ and $F2$ (Munson et al. 2006). Similarly, /IY/ and /UY/ were characterized by higher $F1$ and /AX/ by lower $F1$ in homosexual male speakers from southern Alberta, Canada (Rendall et al. 2008). Furthermore, perception experiments (Munson and Babel 2007; Munson et al. 2006) confirm that higher $F1$ and $F2$ values correlate with listeners' ratings of male voices as gay sounding. Thus, whilst our acoustic study suggests that acted gay speech is characterized by a shift of spectral components towards female values specific to media stereotyping, acoustic and perceptual observations of non-acted homosexual speech indicates that this shift may also partly reflect the representation and exaggeration by the media of female voice patterns adopted by some gay male speakers (Rendall et al. 2008).

Interestingly, whilst actors' voices displayed the highest frequency values for homosexual roles, the lowest values were registered for the interviews (although the difference between heterosexual characters and interviews was non-significant). Lower levels of emotional intensity and intended voice projection may account for the observed low interview values: in natural speech, vocal effort is often accompanied with an increase of fundamental frequency (Plant and Younger 2000) and rising formant frequencies (especially $F1$), due to amplification of articulatory movements (Audibert et al. 2010; Tom et al. 2001). Moreover, $F0$ has been found to be constantly higher in acted speech than in non-acted speech, presumably due to greater levels of emotional intensity in the former (Kienast and Sendlmeier 2000). Besides, the homosexual and heterosexual acted contexts contained higher percentages of mild emotional context ("anger" and "happiness"), which are known to raise $F1$ and $F2$ (Kienast and Sendlmeier 2000; Murray and Arnott 1993), suggesting that the lower values for interviews may reflect a bias due to the fewer emotional speech instances in such a context.

More generally, *F*0 and formant frequency adjustments similar to that identified here have been hypothesised and observed to play a role in mammal vocal communication. The "frequency code" theory (Ohala 1984) posits that, across species, signallers can vary the expression of their dominance by raising *F*0 and formants to sound smaller, and thus less threatening, while *F*0 and formant lowering are associated with greater body size and aggressiveness.

Whilst studies of animal (Davis 1987; Fitch and Reby 2001; Lopez et al. 1988; Reby et al. 2005) and human (Puts et al. 2006) vocal communication support this hypothesis, there is also growing evidence that in human speech, *F*0 and formant manipulations are involved in the vocal expression of gender-related attributes (Feinberg et al. 2005; Pisanski and Rendall 2011). As mentioned in the introduction, despite negligible differences in anatomy between the two sexes in the pre-pubertal stage (Lee et al. 1999; Sachs et al. 1973; Vorperian and Kent 2007), boys have lower formants with consequently narrower formant spacing than girls, suggesting that children acquire the ability to behaviorally achieve gender-specific formant patterns during development (Mattingly 1966; Sachs et al. 1973). Furthermore, increases in *F*0 (by shortening the length and/or increasing the tension of the vocal folds) and formant frequencies (by raising the larynx and/or spreading the lips), convey 'friendliness', 'politeness', 'vulnerability', and 'femininity' (Sachs et al. 1973), which are typically considered female characteristics, while decreases in these frequency components convey 'aggressiveness', 'assertiveness', and 'masculinity' (Chuenwattanapranithi et al. 2006, 2008; Puts et al. 2007).

The specific gestures at the basis of the acoustic variation reported in this study remain to be investigated. While increases in mean *F*0 (pitch) can be achieved by adjusting vocal fold length (Titze 1994), upward shifts of formant frequencies (which will result in a less baritone timbre) can involve either vowel-specific or more global adjustments. For example, research on vowel fronting shows that North-American male speakers from Northern states produce raised and fronted /AE/, thus lowering *F*1 and raising *F*2 (Clopper et al. 2005), while speakers from Southern states tend to front back-vowels /u/ and /o/ which would lead to higher mean *F*1 and *F*2, due to the combination of tongue, lip, and laryngeal movements (Thomas 2003). Here the upward shift is identified at supra-segmental level and involves most formant frequencies (with the exception of *F*3, which is not significantly raised in the drama genre), suggesting a global adjustment of vocal tract length, which could be obtained via lip spreading and/or larynx lowering. In an idealized uniform linear vocal tract with a constant cross-sectional area, vocal tract length variation by lip rounding or by larynx lowering should uniformly affect the frequency of all formants (Titze 1994). However, vocal tract modeling (Fagel 2010; Lasarcyk and Trouvain 2003; Sundberg and Nordstrom 1976) and production (Fagel 2010; Lasarcyk and Trouvain 2003; Tivoli and Gordon 2008) studies show that lip and larynx movements affect formants differently and that these differences are vowel-specific. The retraction of the mouth corners ("smiling") is characteristic of female speakers across cultures (Drahota et al. 2008; Tartter 1980) and the effect of the associated shortening of the vocal tract on the quality of the voice has been hypothesized to contribute to the expression voice gender (Sachs et al. 1973) due to associated raising of formant frequencies (Tartter 1980). Future studies could investigate these interactions between facial and vocal behaviors and their contribution to gender expression in general, and to the "effeminacy" stereotype attributed to male homosexual characters in particular.

Conclusions

This study shows that the vocal behavior of actors playing homosexual characters conforms with the effeminacy stereotype, as they alter the frequency components of their voice along the existing sexual dimorphism in adult human voices: vocal tract resonances are raised towards female values, and $F0$ mean and $F0$ variation are increased towards female values. In perceptual terms, these manipulations result in actors having higher-pitched, lighter, and more expressive voices when playing homosexual roles than when playing heterosexual ones. These results on stereotypical acted speech show that speakers can use behavioral strategies to adjust gender-related acoustic properties at the source ($F0$) and filter (formants) level, in order to vary their expression of gender and gender-related attributes. The ontogeny of these vocal gestures, and the extent to which they are used for the expression of gender and sexual orientation, in both acting and everyday life, is an exciting area for future research.

Acknowledgments Many thanks to Karen McComb and Robin Banerjee for their helpful comments on earlier versions of the paper.

Appendix 1

Calculation of Formant Spacing

Formant spacing (ΔF) was calculated by fitting a model that assumes that the vocal tract is an open (lips)—closed (glottis) tube with a uniform cross-section (quarter-wave resonator) to the observed formant values (Reby and McComb 2003). In this model individual formant frequencies are inversely related to the length of the vocal tract by the following formula:

$$F_i = \frac{(2i-1)c}{4VTL}, \quad (1)$$

where c is the speed of sound in air (approximated as 350 m/s in the vocal tract), i is the number of the formant ($i = 1, 2, \ldots$) and VTL is the length of the vocal tract (Titze 1994).

Since the formant frequency spacing can be expressed as the difference between any two adjacent formants, ΔF is inversely related to VTL:

$$\Delta F = F_{i+1} - F_i = \frac{c}{2VTL} \quad (2)$$

By replacing VTL in Eq. (1) with ΔF estimated in Eq. (2), individual formants are directly related to ΔF:

$$F_i = \frac{(2i-1)}{2}\Delta F \quad (3)$$

Thus, ΔF can be derived from Eq. (3) as the slope of the linear regression of observed formant frequency values Fi (y-axis) over the expected formant positions $(2i-1)/2$ (x-axis), and with the intercept set to 0 (Reby and McComb 2003).

Appendix 2

See Tables 2, 3 and 4.

Table 2 Acoustic characteristics (in Hz) of actors' voices in homosexual roles

ID	*F*0*mean*	*F*0SD	*F*0*CV*	*F*1	*F*2	*F*3	*F*4	Δ*F*
1	136.45	22.38	.16	658.17	1,703.44	2,567.29	3,496.64	1,025.75
2	138.66	29.46	.21	710.72	2,069.90	2,639.85	3,479.50	1,058.96
3	169.38	30.08	.18	732.49	1,868.02	2,798.73	3,506.23	1,068.42
4	141.79	35.92	.25	590.46	1,677.16	2,620.69	3,423.58	1,016.44
5	118.84	21.35	.18	661.30	1,683.68	2,556.70	3,526.29	1,028.09
6	142.55	38.23	.27	637.36	1,791.07	2,667.11	3,496.28	1,043.33
7	138.51	15.57	.11	584.81	1,851.97	2,635.30	3,457.02	1,036.10
8	131.84	21.48	.16	668.06	2,021.73	2,571.41	3,412.64	1,035.21
9	147.08	48.37	.33	663.12	1,747.89	2,685.96	3,459.51	1,036.98
10	127.77	27.17	.21	647.93	1,885.65	2,586.81	3,360.44	1,018.14
11	178.38	34.17	.19	702.62	1,900.28	2,685.17	3,524.04	1,059.47
12	151.5	28.75	.19	627.93	1,692.68	2,615.17	3,416.65	1,016.63
13	145.58	28.12	.19	626.64	1,861.49	2,633.46	3,399.26	1,027.93
14	137.32	31.46	.23	596.65	1,798.63	2,501.30	3,414.93	1,009.61
15	164.00	37.37	.23	646.24	1,772.70	2,633.49	3,545.31	1,046.40
Mean	143.21	29.99	.21	650.30	1,821.75	2,626.57	3,461.23	1,035.16
SD	13.42	8.20	.05	42.93	118.44	69.42	54.78	17.15

*Note F*0*mean* (Hz), *F*0SD (Hz), *F*0*CV* (SD/mean), mean *F*1–*F*4 formant frequency values (Hz) and spacing Δ*F* (Hz) for each actor (ID) and across actors playing homosexual roles

Table 3 Acoustic characteristics (in Hz) of actors' voices in heterosexual roles

ID	*F*0*mean*	*F*0SD	*F*0*CV*	*F*1	*F*2	*F*3	*F*4	Δ*F*
1	120.73	18.31	.15	573.13	1,546.60	2,477.16	3,467.66	996.96
2	117.88	23.69	.20	601.72	1,683.77	2,614.79	3,400.04	1,012.55
3	127.45	17.16	.13	574.90	1,757.37	2,471.06	3,421.19	1,003.59
4	124.51	26.48	.21	618.23	1,765.44	2,598.68	3,301.90	1,000.51
5	92.37	6.11	.07	623.88	1,644.86	2,731.25	3,430.20	1,029.19
6	118.19	21.17	.18	637.22	1,763.14	2,582.43	3,547.54	1,039.80
7	114.70	13.91	.12	604.32	1,867.71	2,669.23	3,469.71	1,043.85
8	117.93	30.29	.26	615.13	1,493.35	2,606.53	3,367.10	992.80
9	117.39	27.50	.23	588.28	1,642.46	2,407.20	3,275.31	963.78
10	103.79	21.24	.20	638.19	1,713.14	2,533.68	3,245.82	980.16
11	160.04	31.80	.20	638.29	1,925.41	2,582.14	3,523.35	1,047.35
12	146.07	33.08	.23	641.70	1,726.94	2,633.55	3,429.19	1,023.68
13	144.35	28.40	.20	600.48	1,714.24	2,659.09	3,426.67	1,024.41
14	112.99	12.91	.11	555.28	1,524.64	2,455.22	3,400.85	981.22
15	125.62	23.01	.18	585.36	1,541.22	2,458.04	3,513.46	1,002.23
Mean	122.93	22.33	.18	606.41	1,687.35	2,565.38	3,414.67	1,009.47
SD	16.85	7.65	.05	27.11	124.79	96.08	88.24	24.94

*Note F*0*mean* (Hz), *F*0SD (Hz), *F*0*CV* (SD/mean), mean *F*1–*F*4 formant frequency values (Hz) and spacing Δ*F* (Hz) for each actor (ID) and across actors playing heterosexual roles

Table 4 Acoustic characteristics (in Hz) of actors' voices in interviews

ID	*F*0*mean*	*F*0SD	*F*0*CV*	*F*1	*F*2	*F*3	*F*4	Δ*F*
1	107.73	11.83	.11	543.24	1,501.34	2,430.73	3,240.24	949.58
2	108.52	21.75	.20	589.44	1,576.22	2,264.63	3,297.41	945.79
3	156.62	27.63	.18	579.77	1,715.74	2,602.84	3,372.69	1,008.33
4	104.76	15.73	.15	563.30	1,608.11	2,507.31	3,305.99	977.77
5	87.79	13.11	.15	571.58	1,731.28	2,523.78	3,251.82	979.69
6	113.97	24.63	.22	609.04	1,560.51	2,491.88	3,555.07	1,015.13
7	100.82	11.64	.12	538.81	1,659.24	2,609.09	3,472.88	1,020.77
8	139.77	27.16	.19	619.83	1,471.86	2,527.42	3,290.53	969.20
9	100.93	21.48	.21	578.89	1,921.19	2,573.93	3,475.56	1,036.69
10	96.50	11.33	.12	536.56	1,701.42	2,460.72	3,412.12	995.94
11	98.60	15.67	.16	605.08	1,646.72	2,424.97	3,604.06	1,021.39
12	109.36	20.76	.19	605.79	1,748.47	2,708.65	3,337.66	1,018.05
13	113.64	14.20	.12	589.46	1,753.40	2,566.68	3,234.25	983.88
14	95.30	14.98	.16	545.65	1,621.10	2,411.03	3,251.98	957.81
15	156.94	31.98	.20	543.52	1,519.56	2,462.29	3,438.97	987.77
Mean	112.75	18.92	.16	574.66	1,649.08	2,504.40	3,369.42	991.19
SD	21.38	6.64	.04	28.38	117.86	104.81	118.68	28.33

*Note F*0*mean* (Hz), *F*0SD (Hz), *F*0*CV* (SD/mean), mean *F*1–*F*4 formant frequency spacing Δ*F* (Hz) for each actor (ID) and across actors in interviews

References

Audibert, N., Aubergé, V., & Rilliard, A. (2010). Prosodic correlates of acted vs. spontaneous discrimination of expressive speech: a pilot study. In *Proceedings from 5th International Conference on Speech Prosody*. Chicago, USA.

Avery, J. D., & Liss, J. M. (1994). Acoustic characteristics of less-masculine-sounding male speech. *Journal of the Acoustical Society of America, 99*, 3738–3748.

Battles, G. K., & Hilton-Morrow, W. (2002). Gay characters in conventional spaces: Will and Grace and the situation comedy. *Critical Studies in Media Communication, 19*, 87–105.

Blashill, A. J., & Powlishta, K. K. (2009). Gay stereotypes: The use of sexual orientation as a cue for gender-related attributes. *Sex Roles, 61*, 783–793.

Boersma, P., & Weenink, D. (2009). PRAAT: Doing phonetics by computer (Version 5.1.19) [Computer program]. Retrieved from: http://www.praat.org.

Busby, P., & Plant, G. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America, 97*, 2603–2606.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., & Maneewongvatana, S. (2006). Expressing anger and joy with the size code.In *Proceedings of Speech Prosody*. Dresden, Germany.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., & Maneewongvatana, S. (2008). Encoding emotions in speech with the size code: A perceptual investigation. *Phonetica, 65*, 210–230.

Chung, S. K. (2007). Media literacy art education: Deconstructing lesbian and gay stereotypes in the media. *International Journal of Art and Design Education, 26*, 98–107.

Clopper, C., Pisoni, D. B., & Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America, 118*, 1661–1676.

Costanzo, F. S., Markel, N. N., & Costanzo, P. R. (1969). Voice quality profile and perceived emotion. *Journal of Counseling Psychology, 16*, 267–270.

Davis, M. S. (1987). Acoustically mediated neighbor recognition in the North-American bullfrog, Rana-Catesbeiana. *Behavioral Ecology and Sociobiology, 21*, 185–190.

Devillers, L., & Vasilescu, I. (2003). Prosodic cues for emotion characterization in real-life spoken dialogs. In *Proceedings of Eurospeech* (pp. 189–192).

Dow, B. (2001). Television, and the politics of gay and lesbian visibility. *Critical Studies in Media Communication, 18*, 123–140.

Drahota, A., Costall, A., & Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Communication, 50*, 278–287.

Fagel, S. (2010). Effect of smiled speech on lips, larynx and acoustics. *Lecture Notes in Computer Science, 5967*, 294–303.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton and Co.

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.

Fitch, W., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America, 106*, 1511–1522.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. In *Proceedings of the Royal Society*. London, UK: Royal Society Publishing.

Frick, R. W. (1986). The prosodic expression of anger: Differentiating threat and frustration. *Aggressive Behavior, 12*, 121–128.

Gaudio, R. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech, 96*, 30–57.

Hajek, C., & Howard, G. (2005). Intergroup communication schemas: Cognitive representations of talk with gay men. *Language & Communication, 25*, 161–181.

Henton, C. G. (1989). Fact and fiction in the description of female and male pitch. *Language & Communication, 9*, 299–311.

Henton, C. G. (1995). Pitch dynamism in female and male speech. *Language & Communication, 15*, 43–61.

Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97*, 3099–3111.

Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nature, 21*, 406–427.

Hollien, H., Green, R., & Massey, K. (1994). Longitudinal research on adolescent voice change in males. *Journal of the Acoustical Society of America, 96*, 2646–2654.

Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry, 43*, 280–283.

Kienast, M., & Sendlmeier, W. F. (2000). Acoustical analysis of spectral and temporal changes in emotional speech. In *Proceedings from the ISCA ITRW on Speech and Emotion*. Newcastle, UK: Textflow.

Kite, M. E., & Deaux, K. (1987). Gender belief systems: Homosexuality and the implicit inversion theory. *Psychology of Women Quarterly, 11*, 83–96.

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the voice reveals: Within- and between-category stereotyping on the basis of voice. *Personality and Social Psychology Bulletin, 32*, 806–819.

Lasarcyk, E., & Trouvain, J. (2003). Spread lips + raised larynx + higher F0 = smiled speech?: An articulatory synthesis approach. In *Proceedings of the 8th International Seminar on Speech Production*. Strasbourg, France: INRIA.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455–1468.

Linneman, T. J. (2008). How do you solve a problem like Will Truman? The feminization of gay masculinities on Will and Grace. *Men and Masculinities, 10*, 583–603.

Lopez, P. T., Narins, P. M., Lewis, E. R., & Moore, S. W. (1988). Acoustically induced call modification in the white-lipped frog, Leptodactylus albilabris. *Animal Behavior, 36*, 1295–1308.

Madon, S. (1997). What do people believe about gay males? A study of stereotype content and strength. *Sex Roles, 37*, 663–685.

Mattingly, I. G. (1966). Speaker variation and vocal-tract Size (A). *The Journal of the Acoustical Society of America, 39*, 1219(A).

Munson, B., & Babel, M. (2007). Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass, 1*, 416–449.

Munson, B., McDonald, E., & DeBoe, N. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics, 24*, 202–240.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of Acoustic Society of America, 93*, 1097–1106.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica, 41*, 1–16.

Perry, T., Ohde, R., & Ashmead, D. (2001). The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America, 109*, 2988–2998.

Pierrehumbert, J., Bent, T., & Munson, B. (2004). The influence of sexual orientation on vowel production (L). *The Journal of the Acoustical Society of America, 116*, 1905–1908.

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *Journal of the Acoustical Society of America, 129*, 2201–2212.

Plant, R. L., & Younger, R. M. (2000). The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech. *Journal of Voice, 14*, 170–177.

Puts, D., Gaulin, S., & 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.

Raley, A. B., & Lucas, L. J. (2006). Stereotype or success? Prime-time television's portrayals of gay male, lesbian, and bisexual characters. *Journal of Homosexuality, 51*, 19–38.

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behavior, 65*, 519–530.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., & Fitch, W. T. (2005). *Red deer stags use formants as assessment cues during intrasexual agonistic interactions*. London: Proceedings of the Royal Society, Royal Society Publishing.

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (*F*0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *Journal of the Acoustical Society of America, 117*, 944–955.

Rendall, D., Vasey, P., & McKenzie, J. (2008). The Queen's english: An alternative, biosocial hypothesis for the distinctive features of "gay speech". *Archives of Sexual Behavior, 37*, 188–204.

Riordan, C. J. (1977). Control of vocal-tract length in speech. *The Journal of the Acoustical Society of America, 62*, 998–1002.

Sachs, J., Lieberman, P., & Erikson, D. (1973). Anatomical and cultural determinants of male and female speech. In R. W. Shuy & R. W. Fasold (Eds.), *Language attitudes* (pp. 74–84). Washington: Georgetown University Press.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass, 3*, 621–640.

Smith, D. R. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgments of speaker size, sex, and age. *The Journal of the Acoustical Society of America, 118*, 3177–3186.

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society, 32*, 329–350.

Staats, G. R. (1978). Stereotype content and social distance: Changing views of homosexuality. *Journal of Homosexuality, 4*, 15–27.

Sundberg, J., & Nordstrom, P. E. (1976). Raised and lowered larynx: The effect on vowel formant frequencies. *Speech Transmission Laboratory Quarterly Progress, 17*, 35–39.

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics, 27*, 24–27.

Terengo, L. (1966). Pitch and duration characteristics of the oral reading of males on a masculinity-femininity dimension. *Journal of Speech and Hearing Research, 9*, 590–595.

Thomas, E. R. (2003). Secrets revealed by southern vowel shifting. *American Speech, 78*, 150–170.

Titze, I. R. (1994). *Principles of voice production*. Iowa City: IA7 National Center for Voice and Speech.

Tivoli, M., & Gordon, M. J. (2008). The [+ spread] of the Northern Cities Shift. In *University of Pennsylvania Working Papers in Linguistics*. (Vol. 14, pp. 111–120).

Tom, K., Titze, I. R., & Hoffman, E. A. (2001). Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness. *The Journal of Acoustical Society of America, 109*, 742–747.

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*, 1510–1545.

Vorperian, H. K., Kent, R. D., & Lindstrom, M. J. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America, 117*, 338–350.

Wolfe, V. I., Ratusnik, D. L., Smith, F. H., & Northrop, G. (1990). Intonation and fundamental frequency in male-to-female transsexual. *Journal of Speech and Hearing Research, 55*, 43–50.

Wondershare Software Co. Ltd. (2010). iSkysoft Audio Ripper (Version 1.8.2.7) [Computer program]. Retrieved from: http://www.iskysoft.com.

Wyatt, D. A. (2008). Gay/lesbian/bisexual television characters. [Website] Retrieved from: http://home.cc.umanitoba.ca/~wyatt/tv-characters.html.