Male voices and perceived sexual orientation: An experimental and theoretical approach

RON SMYTH, GREG JACOBS and HENRY ROGERS

Language in Society / Volume 32 / Issue 03 / June 2003, pp 329 - 350
DOI: 10.1017/S0047404503323024, Published online: 21 October 2003

Language in Society **32**, 329–350. Printed in the United States of America
DOI: 10.1017/S0047404503323024

Male voices and perceived sexual orientation: An experimental and theoretical approach

RON SMYTH

Departments of Linguistics and Psychology
University of Toronto
130 St. George St.
Toronto, Ontario M5S 3H1, Canada
smyth@utsc.utoronto.ca

GREG JACOBS*

Department of Languages, Literatures and Linguistics
York University, Toronto, Ontario

HENRY ROGERS

Department of Linguistics
University of Toronto
rogers@chass.utoronto.ca

ABSTRACT

This article describes the development of a data bank of 25 male voices spanning the range from very gay-sounding to very straight-sounding, according to listener ratings. These ratings allowed the researchers to examine the effects of different discourse types (scientific, dramatic, and spontaneous) and listener groups (gay males vs. a mix of males and females of unknown sexual orientation) on how listeners perceived the voices. The effects of lexical and pragmatic content were explored by a comparison of spoken and written presentations of the same spontaneous speech samples. The effect of asking participants to rate the voices using different constructs (e.g., masculine/feminine vs. gay-sounding/straight-sounding) is discussed. The ultimate goal of this research program is to examine correlations between these ratings and a range of phonetic variables in order to shed light on the specific features to which listeners attend when judging whether a man's voice sounds gay or straight. (Gay men, homosexuality, phonetics, sexual orientation, voice)**

INTRODUCTION

The central goal of the research reported here is to understand why some men's voices are perceived as sounding gay (homosexual) and others as straight (heterosexual). In this article we describe the development of a data bank of

25 male voices spanning the range from very gay-sounding to very straight-sounding, according to listener ratings. The listening tasks allowed us to examine the effects of different discourse types (scientific, dramatic, and spontaneous), different listener groups (gay males vs. a mix of males and females of unknown sexual orientation), and types of ratings (gay/straight vs. masculine/feminine) on how listeners perceived the voices. We also explore the effects of lexical and pragmatic content by comparing a spoken and a written presentation of the same spontaneous speech samples. In other research, we have examined correlations between these ratings and a range of phonetic variables: voiced and voiceless fricative frequency and duration, vowel duration, vowel formants, /l/ fronting, aspiration, and various pitch measures (Jacobs, Rogers & Smyth 1999, 2000, 2001; Smyth & Rogers 2000; Rogers & Smyth 2000; Rogers, Smyth, & Jacobs 2000a, 2000b). In this article we report one such finding: the relationship between mean pitch and both gay/straight and masculine/feminine ratings. We also discuss various social theoretical frameworks for interpreting our results related to the "gay voice" phenomenon.

To begin our discussion, we will review the literature on the phonetic correlates of gender, which has focused primarily on the acoustic differences between the speech of males and females (Schwartz 1968, Ingemann 1968, Schwartz & Rine 1968, Coleman 1971, Coleman 1976, Lass, Hughes et al. 1976, Lass, Mertz, & Kimmel 1978, Lass, Tecca, et al. 1979, Edelsky 1979, Lass, Almerino et al. 1980, Bennett & Montero-Diaz 1982, Günzburger 1984). Generalizing the results, it is clear that listeners can judge with almost perfect accuracy whether a voice belongs to an adult male or an adult female, at least with regard to the North American English voices used in most of these studies. Although listeners are less able to distinguish between the voices of male and female children, the success rate is nonetheless better than chance (Weinberg & Bennett 1971, Sachs, Lieberman, & Erickson 1973, Sachs 1975, Bennett & Weinberg 1979a, Bennett & Montero-Diaz 1982). The studies from outside North America provide similar results (Swedish children, Fichtelius, Johansson, & Nordin 1980; Dutch children, Bresser & Günzburger 1985, Günzburger, Bresser, & ter Keurs 1987; Scottish children, Lee, Hewlett & Nairn 1995).

The layperson typically assumes that identifying a person's sex by the voice is related to male-female differences in anatomy and physiology. Bergvall 1999 briefly discusses how popular interpretations of sex-based behavioral differences are explained in terms of physical differences. Certainly, the difference between men's and women's vocal fold length would account for the almost uniform finding that the average pitch in adult men is lower than in women (Weaver 1924, Snidecor 1951, Linke 1973, Coleman 1976, Lass, Hughes et al. 1976, Loveday 1981, Günzburger 1984, Gilmore et al. 1992). But if differences in pitch are purely physiological, this could not account for the better-than-chance ability to distinguish between the voices of preadolescent boys and girls, whose physical articulatory properties and average pitch are similar (Weinberg & Zlatin 1970,

Weinberg & Bennett 1971, Sachs, Lieberman & Erickson 1973, Bennett & Weinberg 1979a, 1979b, Bresser & Günzburger 1985, Günzburger, Bresser & ter Keurs 1987, Lee, Hewlett & Nairn 1995). In their review of the literature, Lee, Hewlett & Nairn conclude that "the current balance of evidence from anatomical studies would suggest that the vocal organs of prepubertal boys and girls are not significantly different" (1995:199–200). Interestingly, even when pitch differences between adult male and female speakers are eliminated through electronic manipulations, listeners can still distinguish between the two (Schwartz & Rine 1968, Coleman 1971, Lass, Hughes et al. 1976, Lass, Almerino et al. 1980, Günzburger 1984). Therefore, listeners must be relying on other cues beyond the average fundamental frequency to make their judgments. It is clear, then, that differences are not entirely biological and must be the result of other, more social factors. Bergvall 1999 argues that a comprehensive theory of language and gender must also include the social (speaker agency) and the ideological (more macro-level societal prescriptions).

Pitch variability and range are other pitch-related properties (in addition to AVERAGE pitch, discussed above) that have been explored with regard to male/female speech differences. The general consensus of the published literature is that males tend to use less of the pitch range available to them, and to shift their pitch less frequently, than do females (Brend 1975, Lass, Hughes et al. 1976, Fichtelius, Johansson, & Nordin 1980, Gilmore et al. 1992). However, there are various practices of measuring pitch variability, as well as disputes about how this should be done (cf. Henton 1989), and not all research supports the prevailing evidence (Snidecor 1951, Linke 1973, Bennett & Weinberg 1979b, Henton 1989). Gaudio 1994 questions the reliability of ANY study of pitch variability that relies on gross averages, and urges a more nuanced approach that considers the interaction of other factors (e.g., prosodic, segmental, and lexical).

Other possible acoustic cues that have been explored as a potential source of differentiation between male and female voices include vowel length and formant frequencies (Schwartz & Rine 1968, Coleman 1971, Sachs, Lieberman, & Erickson 1973, Sachs 1975, Coleman 1976, Lass, Hughes et al. 1976, Bennett & Weinberg 1979a, 1979b, Bennett & Montero-Diaz 1982, Bresser & Günzburger 1985, Lee, Hewlett, and Nairn 1995, Avery & Liss 1996, Högberg 1996), sibilant and other consonant production (Schwartz 1968, Ingemann 1968, Avery & Liss 1996, Linville 1998), loudness (Markel, Prebor, & Brandt 1972, von Raffler-Engel & Buckner 1983), the use of high rising terminals (Edelsky 1979, Britain 1992), breathiness (Henton & Bladon 1985, Klatt & Klatt 1990, Hillenbrand, Cleveland, & Erickson 1994), and creakiness (Henton & Bladon 1988).

We now turn to the literature on phonetics and sexual orientation. A prevailing belief that has frequently surfaced from our discussions with students in the classroom is that gay men's speech mirrors the patterns of stereotypical high-pitched women's voices, including highly variable intonation.[1] One experiment that refutes this stereotype is by Lerman & Damsté 1969, who found no significant

differences in average pitch between their gay and straight speakers. In later studies, researchers recognized the drawbacks of using sexual orientation as an independent variable, since there are no clear criteria, other than self-identification, for classifying subjects as gay or straight. A reliance on self-report means that it is impossible to construct a truly random sample based on sexual orientation. In Kulick's (2000) literature review, he notes that one "basic conceptual difficulty that is not resolved in studies like these is that even if listeners had correctly identified the gay and lesbian speakers with 100% accuracy, we would still not know exactly what it was that was being identified. Is it sexual orientation as such, and therefore applicable to all (most? some?) gays and lesbians, even those who are not 'out'?" (273). Although Kulick's critique is aimed at Gaudio 1994 and Moonwomon-Baird 1997, the former did attempt to correct for this shortcoming by ensuring that his four gay and straight speakers did at least sound "gay" and "straight" respectively, based on listener judgments, and by correlating these judgments with pitch behavior (range and variability). In other words, his independent variable was not the actual sexual orientation of the speakers ("gay" versus "straight"), but rather listener perceptions of the sexual orientations of the speakers. Simplifying his conclusions considerably, we note that Gaudio claimed that his gay-sounding speakers employed greater pitch variation than their straight-sounding counterparts: they used more of the pitch range and changed pitch more frequently. However, this difference was found in only one of the two speaking tasks, and of the 13 measures, only one showed a statistically significant difference, and eight only approached significance.

Some researchers have abandoned the notion of sexual orientation altogether by having their speech samples rated on the dimensions of Masculine-Feminine (Terango 1966), or More-Masculine-Sounding (MMS) versus Less-Masculine-Sounding (LMS) (Avery & Liss 1996), and correlating these judgments with specific cues. Terango found that his more feminine-sounding male subjects had higher average fundamental frequencies, but only one of his eight measures for range and variability was statistically significant. In contrast, Avery and Liss's analyses found no differences in fundamental frequency between their two groups. In addition, some of their measures of pitch variability proved significant, but most did not. If pitch does not provide sufficient cues to allow listeners to infer sexual orientation, or degree of masculinity/femininity, then the question remains as to what cues the listeners are attending to in order to make these judgments.

It is important to stress again that we are not exploring whether gay and straight men have different voices, but rather what properties of a man's voice make listeners judge it as gay- or straight-sounding, regardless of his sexual orientation. When we do refer to subjects as "gay" or "straight," this is based on their self-identification; we recognize that these are social labels and may or may not correlate with actual sexual orientation.

METHODOLOGY

The reading task

As a first step, we collected a sample of voices from both straight- and gay-sounding men. The snowball technique (inviting our acquaintances, and inviting them to invite their acquaintances, etc.) was used to recruit 25 speakers. Our goal was to have a fairly even distribution of voices along a continuum from gay- to straight-sounding; in particular, we did not want a truly random sample, since that could yield a larger proportion of straight-sounding voices. The speakers ranged in age from 25 to 50 years, and all were native speakers of a variety of Canadian English. Eight self-identified as straight, and 17 as gay.

To begin, the speakers were asked to complete an informed consent statement that did not tell them the purpose for which we were recording their voices. One ethical issue that arose was that some participants might react adversely to knowing that their voices were being collected with the aim of potentially identifying them as gay-sounding – a socially stigmatized speech variety both inside and outside the gay community. However, to inform the participants of the study's aim prior to the recordings could have influenced their speaking style. Therefore, participants were informed that there would be some deception, and that the purpose of the research would be explained to them after the recording session, at which point they would have the right to request that their speech sample be erased. No participant asked to have his recording erased, and in fact, most were intrigued with the project and asked further questions.

The speakers were asked to complete three tasks:

(i) Read a SCIENTIFIC PARAGRAPH, devised by Fairbanks 1966, whose intention was to create a phonetically balanced passage. We used this "Rainbow Passage" because its subject matter (the history and science of rainbows[2]) evokes little emotional involvement in the speaker.

(ii) Read a DRAMATIC PARAGRAPH, created by Crist 1997 to investigate the use of the phoneme /s/ as a stereotype that men draw upon to make their voices sound more gay. We are using it because its dramatic content (the use of the first person to tell a story about a fire) promised to create a more "excited sounding" voice.

(iii) Respond to an OPEN-ENDED QUESTION, intended to divert the speaker's attention from the tape-recorder and to elicit a more spontaneous speech sample. Participants were asked to tell a true incident that had happened to them, based on one of six scenarios (e.g., a recent argument with someone, a bad driving experience).

The speaking tasks took approximately 30 minutes in total, and participants were given a $10 music store gift certificate, which they were informed they could keep even if they requested that their recordings be erased after the debriefing.

TABLE 1. *Sex and mean age of listener-subjects.*

	n	Mean age
All Gay Listener Group	14 Males	37.7
Mixed Group	32	26.3
	13 Males	*25.5*
	19 Females	*26.9*
Total	46	29.8

The listening task

Three separate listening tasks were constructed, using the tapes from the three speaking tasks. For the scientific and dramatic readings, three of the middle sentences were extracted for each of the 25 speech samples. These were identical for each of the 25 speakers, and the approximate duration was 30 seconds per sample. For the spontaneous speech samples, we selected approximately 30 seconds for analysis. Obviously, identical portions could not be obtained, since each participant's response was unique. However, caution was exercised to ensure that the content of the passage did not explicitly reveal the sexual identity of the speaker (e.g., by reference to the sex of romantic or sexual partners), and that it did not refer to stereotypical gay or straight behavior (e.g. lack of familiarity with sports).

For each of the speaking tasks, a master tape was constructed with the 75 thirty-second speech samples (25 speakers $\times$ 3 speaking tasks), with a five-second response interval between samples. The voices were presented in a different random order for each task.

Forty-six listener-subjects were recruited, 14 of whom were explicitly identified as gay males. The remainder formed a mixed group, by which we mean that we did not ask about their sexual orientation and we presume that most identified as heterosexual. The mixed group (13 males and 19 females) was recruited from the University of Toronto community, and the gay males from the experimenters' friends and acquaintances. The only information we collected on the listeners was their sex and age. Table 1 provides some demographic information.

The listeners were presented first with the scientific (Rainbow) passage. They were instructed to listen to each of the voice samples, and during the five-second pause that followed each token, to indicate whether the voice sounded gay or straight; this was a forced-choice response. In addition, they had to mark their confidence in their judgment on a scale from zero ("total guess") to six ("100% positive"). The same listeners were given the same instructions for speaking task (ii), the dramatic (fire) paragraph, and this was followed by speaking task (iii), the spontaneous responses to the open-ended question. The listening task took approximately 45 minutes, and participants were given a $10 music store gift certificate.

TABLE 2. *Mean "sounds gay" judgments and confidence ratings for 25 male voices, collapsed across the three discourse types (mixed listener group only).*

Speaker Identification	Proportion "sounds gay"	Mean confidence score	Speaker's Sexual Orientation
4	0.10	4.2	Gay
23	0.12	4.3	Straight
13	0.14	4.2	Gay
21	0.14	4.5	Straight
24	0.16	4.2	Straight
8	0.24	3.8	Straight
15	0.25	4.1	Gay
14	0.25	4.1	Straight
22	0.28	3.6	Straight
1	0.28	3.7	Gay
10	0.28	4	Gay
17	0.35	3.6	Straight
19	0.38	3.7	Gay
18	0.39	3.2	Gay
11	0.43	4.1	Gay
6	0.52	3.9	Gay
5	0.53	3.6	Gay
12	0.60	3.8	Straight
7	0.79	3.9	Gay
16	0.80	4.2	Gay
2	0.81	4.1	Gay
20	0.83	4.1	Gay
3	0.84	4.1	Gay
9	0.92	4.5	Gay
25	0.98	5.0	Gay

RESULTS

Distribution of gay- and straight-sounding voices

The results were first examined against our original goal of creating a data bank of voices for further analysis and judgment studies. We wanted the voices to span the range from very gay-sounding to very straight-sounding in order to meet the statistical assumptions underlying our subsequent correlational analyses. Based on the ratings of the group of listeners of unknown sexual orientation (the "mixed" group), Table 2 shows that we do have a good distribution of mean "sounds gay" ratings. In particular, we have five voices that 80% or more of the listeners from the mixed group rated as straight-sounding (Speakers 4, 23, 13, 21, and 24), and six voices that 80% or more of these listeners rated as gay-sounding (25, 9, 3, 20, 2 and 16). This clearly supports one of our main research objectives, which was

to find evidence for the existence of the "gay/straight voice" phenomenon. It is also interesting to note that these 11 extreme voices were rated with high confidence (4.1 or higher). In addition, we have 14 voices in the middle range, where the judgments were made with somewhat less confidence: only four had confidence ratings of greater than 4.0.

Thus, when mean judgments from the mixed listeners are plotted against their mean confidence ratings, the result is a U-shaped curve. Statistically speaking, a significantly U-shaped curve should yield a significant polynomial regression with a quadratic trend. This was true for all three discourse types (scientific passage: $R^2 = .527$; dramatic passage: $R^2 = .653$; spontaneous speech sample: $R^2 = .568$; in all three cases both the linear and quadratic components were significant at the .0001 level). Clearly, then, listeners were in agreement as to which voices sounded clearly gay or straight, and their confidence judgments reflected their uncertainty about the other voices.

Main analysis

Our main statistical method was a mixed analysis of variance design. To understand this analysis, it is important to note that the participants in this study were the 46 listeners. The 25 speakers were not participants in the statistical sense; rather, they were the source of the two kinds of speech samples: straight men's voices (n = 8) and gay men's voices (n = 17).

The dependent variable was the mean proportion of "sounds gay" judgments given by the 46 listeners. These listeners were divided into three Listener Groups (the between-groups variable). There were six means for each listener, one for each combination of the two within-groups factors: Speaker Sexual Orientation (gay vs. straight) × Discourse Type (scientific, dramatic, and spontaneous).

For the "sounds-gay" judgments, the main effect of Listener Group was not significant: $F(2,43) = 2.69$, $P = .08$. The mean ratings, collapsed across all speakers and all discourse types, were .46 for gay male listeners, .43 for mixed female listeners, and .37 for mixed male listeners. In other words, all three listener groups made approximately the same number of "gay" judgments overall. On the other hand, results between $P = .05$ and $P = .10$ are often viewed as marginally significant and given tentative interpretations. We therefore compared the differences among these three means with the Newman-Keuls test. Only the difference between the males from the "gay" group and the males from the "mixed" group approached significance ($P = .055$). If this is a replicable result, it means that gay males are more likely to label voices as sounding gay. The mean for "mixed" female listeners fell between those of the other two groups, but was not significantly different from either of them.

The main effect of Discourse Type was significant: $F(2,86) = 6.55$, $P = .002$. The mean "sounds-gay" rating, collapsed across all speakers and listeners, was higher for the scientific passage (.45) than for either the dramatic passage or the spontaneous speech sample (both .40) ($P < .005$ in both cases), although the

TABLE 3. *Mean proportion of "sounds gay" judgments by discourse type for gay and straight speakers (collapsed across listener groups).*

	Gay Speakers	Straight Speakers
Scientific Passage	.56	.35
Dramatic Passage	.53	.26
Spontaneous Speech	0.58	0.22

latter two did not differ significantly from each other. This suggests that the more formal speaking style associated with reading a scientific passage may be interpreted as more gay-sounding (but note the interaction with Speaker's Sexual Orientation to be discussed below).

The main effect of Speaker's Sexual Orientation (.55 for the gay speakers and .28 for the straight speakers) was also significant: $F(1,43) = 165.42$, $P < .0001$. This is hardly surprising given that the speakers were not randomly selected. Some of the men were invited to be recorded precisely because we hoped that they would be judged to have gay-sounding voices. This recruitment strategy was successful, as we now have a varied sample for our database, including a sufficient number of gay-sounding voices for future acoustic analysis. It is likely that a random sample would have yielded a much smaller proportion of gay men, which in turn would have resulted in an even smaller proportion of gay-sounding voices.

In addition to the main effects of Discourse Type and Speaker's Sexual Orientation, there was a significant interaction between these two factors: $F(2,86) = 10.16$, $P = .0001$. The means for this interaction are shown in Table 3. The Newman-Keuls test showed that the ratings of the gay speakers were higher than those of the straight speakers ($P < .0001$ for all three discourse types). However, the straight speakers were rated as significantly more gay-sounding on the scientific passage than on either the dramatic passage ($P < .0003$) or the spontaneous speech sample ($P < .0001$); the ratings on the latter two types did not differ significantly ($P > .10$). On the other hand, the ratings for the gay speakers were similar for all three discourse types ($P > .16$ for each comparison). In other words, only the straight speakers were judged to be more gay-sounding when reading the scientific passage. This supports the idea that formal speech is considered more gay-sounding, as suggested above, and it further suggests that the gay men tended to use this style in all three discourse types, while the straight men used it only for the scientific passage.

Individual differences

We also conducted an individual differences analysis, comparing the mean judgments and confidence ratings given to each voice by the gay and mixed groups of

TABLE 4. *Significant and marginal differences for between-groups t-tests between gay male and mixed listener groups in the "sounds gay" judgments (proportions) and confidence ratings.*

Speaker Identification	Scores		*P*	Speaking Task
	Gay Listeners	Mixed Listeners		
1 (gay)	0.5	0.22	.06	Spontaneous Speech
2 (gay)	1	0.78	.06	Spontaneous Speech
4 (gay)	0.43	0.06	.00001	Spontaneous Speech
5 (gay)	0.86	0.41	.00001	Scientific Passage
7 (gay)	0.93	0.66	.06	Spontaneous Speech
8 (straight)	4.71	3.50 (conf. rating)	.02	Spontaneous Speech
13 (gay)	5.29	4.28 (conf. rating)	.05	Spontaneous Speech
18 (gay)	4	2.75 (conf. rating)	.03	Spontaneous Speech
18 (gay)	0.57	0.28	.06	Dramatic Passage
19 (gay)	0.5	0.22	.06	Spontaneous Speech
20 (gay)	1	0.75	.04	Scientific Passage
25 (gay)	0.79	0.97	.04	Spontaneous Speech

listeners. In this analysis, the males and females from the "mixed" group of listeners were treated as one group. The significant and marginal differences are presented in Table 4. Of the 150 possible comparisons (25 speakers $\times$ 3 discourse types $\times$ 2 measures, i.e. judgment and confidence) only twelve significant differences emerged between the gay and mixed listeners, nine involving "sounds gay" judgments and three involving confidence ratings.[3] With *P* set at .05 for 150 comparisons, one might expect about seven spuriously significant results. Nonetheless, the pattern of the twelve differences is striking. Eleven of them involved gay men's voices. In addition, nine were for the spontaneous speech sample, while two were for the scientific passage, and one was for the dramatic passage. Finally, and perhaps most interestingly, in eleven of the twelve cases the difference was in the direction of the gay listeners rating a voice as more gay-sounding, or rating it with more confidence. Only one voice (Speaker 25) was more often rated as gay-sounding by the mixed group.

"Gaydar" analyses

Leap 1996 devotes an entire chapter to the phenomenon of "gaydar," which he defines as simply the ability to identify men who are likely to be gay. (In a personal communication, Rudolf Gaudio reminded us that this is a PRESUMED ability.) Leap's analyses focus on the more pragmatic discursive strategies in specific daily communities of practice that allow speakers to negotiate their gay identities jointly. By contrast, we are primarily interested in listener judgments based solely

on the acoustic cues available to listeners in an experimental context. We are interested in exploring whether one group of participants would have better auditory gaydar than another. For example, would the women in the "mixed" group have better gaydar than the men in that group? Would the openly gay listeners have the best gaydar?

Note that in this part of the analysis, we deviate from our usual practice and consider a "correct" judgment to be one that corresponds to the speaker's self-identification as gay or straight. Like the previous analysis of "sounds gay" judgments, the gaydar analysis is a mixed analysis of variance with one between-groups factor (Listener Group), and two within-groups factors (Discourse Type and Speaker's Sexual Orientation). In the analyses reported in the sections above, the proportions cited were listeners' perceptions of the talkers' voices, whereas in this section, the proportions given are CORRECT judgments of the talkers' actual sexual orientation (based on self-identification).

There is, of course, considerable overlap between these two data sets (i.e., perceptions vs. correct judgments). For the gay speakers, the proportion of sounds-gay judgments corresponds exactly to the proportion correct, while for the straight speakers the proportion correct is the complement of the proportion of "sounds gay" judgments. Simply put, if the talker is gay, then the percentage of "sounds gay" judgments matches the percentage of "correct" judgments; if the talker is straight, the "sounds gay" judgments are now considered to be errors. If a straight speaker has a mean "sounds gay" rating of .60, then the gaydar score is .40, a "sounds gay" rating of .70 would correspond to a gaydar score of .30, .80 corresponds to .20, and so forth.

There was no main effect of Listener Group: $F(2,43) = 1.08$, $P = .35$. This means that listeners from all three groups (gay, "mixed" female, and "mixed" male) were equally accurate in identifying the correct sexual orientations of all the speakers, with mean proportions correct of .63, .66, and .62, respectively.

The main effect of Speaker's Sexual Orientation was highly significant ($F(1,43) = 28.75$, $P < .00001$): the straight talkers were more accurately identified than the gay talkers (.72 vs. .55). This cannot be taken as an indication of how difficult it is to identify gay men by their voices, randomly chosen from the general population, since (for the reasons outlined above) our sample includes a disproportionate number of gay men who were purposely chosen because the researchers thought they sounded gay.

The main effect of Discourse Type was also significant ($F(2,86) = 10.16$, $P < .0001$). Overall, correct judgments were lowest for the scientific passage (.60), significantly higher for the dramatic passage (.64), and significantly higher still for the spontaneous speech samples (.68). Table 5 shows that this factor interacted significantly with Speaker's Sexual Orientation ($F(2,86) = 6.55$, $P = .002$). Similar to the "sounds gay" judgments, there were no significant differences among the discourse types for the gay speakers (all $P > .20$), but the straight speakers were misidentified as gay most often in the scientific passage

TABLE 5. *Gaydar analysis: Proportion of correct identification of gay and straight speakers by discourse type.*

	Gay Speakers	Straight Speakers	Mean
Scientific Passage	0.56	0.65	0.60
Dramatic Passage	0.53	0.74	0.64
Spontaneous Speech	0.58	0.78	0.68
Mean	0.55	0.72	0.68

($P < .0005$), and the proportion correct for the dramatic and spontaneous conditions did not differ ($P = .13$).

These analyses provide some evidence that gaydar, in general, is not reliable: 309 of the listener judgments (or 28%) of the 1,104 straight talkers' tokens (8 straight talkers $\times$ 3 speaking tasks $\times$ 46 listeners) were incorrectly identified as gay. Furthermore, the accuracy in identifying the gay speakers was significantly lower than for the straight speakers in all three discourse types ($P < .001$ in all cases). Recall that for the gay men, the overall mean proportion of correct judgments was .55, and for the straight men .72. This is particularly interesting given that our sample is heavily biased; we included a disproportionate number of gay-sounding gay men. This suggests that auditory gaydar is not very reliable: we presume that in a random sample of gay men, the proportion judged to be gay in a study of this type would be far lower than the overall mean proportion of .55 observed here. Thus, we can conclude that most gay men do not sound gay, and that a substantial number of straight men do sound gay. Simply put, being right about a straight man is easier than being right about a gay man. This is not surprising given that many, if not most, gay men do not have a gay-sounding voice.

Although our results do not provide clear evidence that gay men in general have a better ability to detect other gay men by voice alone, it is still interesting to note that for two particular speakers, there were large differences between the ratings of the gay and "mixed" listener groups. From Table 4, we note that in the open-ended question, Speaker 4 was given a "sounds gay" rating of .06 by the mixed listener group, but .43 by the gay listeners. We have already hinted that this could be due to pragmatic content, but it is still interesting to note that the mixed group of listeners were oblivious to any cues. On the scientific passage, Speaker 5 received a rating of .41 from the mixed listener group, and .86 from the gay listeners. In this case, we cannot attribute the difference to pragmatics because all the speakers read the same passage. Although further testing is warranted, these results nonetheless open up the possibility that for certain speakers, gay men may indeed have better auditory gaydar.

Lexical and pragmatic factors

Recall that, because each speaker's spontaneous response to the open-ended question was unique, we selected for the listening task only those sentences the researchers subjectively deemed not to reveal the speaker's sexual orientation through lexical or pragmatic content. This, of course, was not an issue for the scientific and dramatic passages, where all speakers read the same texts and the only possible differences were in their spoken renditions.

On the other hand, in the individual differences analyses (Table 4), most of the cases in which "sounds gay" ratings were higher for the gay listeners involved the spontaneous speech samples (9 of 12 cases). This led us to suspect that our attempts to remove gay content had not been successful, and that the gay listeners may have been more attuned to cues of this type.

In order to test this hypothesis, we conducted a follow-up study in which 16 University of Toronto students (eight male and eight female) were presented with written transcripts of the 25 spontaneous speech samples, using the same rating method as in the main listening task. If the lexical and pragmatic content was in fact neutral between gay and straight, then the responses in the reading task should be random.

However, there was a significant correlation between the judgments on the listening task and those on the new reading task ($R = .51$, $P = .009$), and the mean ratings for the two tasks did not differ ($F(1,24) = 1.12$, $P = .30$). This provides strong support for our suspicion that despite our best efforts, lexical and pragmatic content did influence the responses of the listeners in the spontaneous speech task. Even after removing the acoustic cues, and all the most obvious lexical and pragmatic cues, the remaining, more subtle information still allowed the readers to rate the speakers in about the same way as those who actually heard the identical tokens on the tape.

"GAY/STRAIGHT" VS. "MASCULINE /FEMININE"

We were also interested in whether the construct "gay- or straight-sounding" would give different results from other, similar constructs. Gaudio 1994 found a high correlation between listener's judgments of "gay-sounding" and "effeminate," and between "straight-sounding" and "masculine." Terango 1966 examined the relationship between the phonetic properties of male voices and ratings defined as "Masculine-Feminine," and Avery & Liss 1996 conducted similar tests using "More-Masculine-Sounding" versus "Less-Masculine-Sounding." However, it is not clear whether, or to what extent, the results of such studies are relevant to our research on straight/gay judgments. In other words, would different constructs yield different results? Although there does seem to be a relationship between popular notions of gayness and femininity, there may be some differences between these concepts. For example, listeners might resist labeling a low-pitched voice as "feminine," even if they consider it to be very gay-sounding.

To address this question directly, we conducted another follow-up study in which 16 University of Toronto students, 8 males and 8 females, listened to the same 75 speech samples as in our main experiment. The only difference was that the forced-choice judgment was between "masculine" and "feminine," rather than between "gay" and "straight."

There was a significant correlation between the two ratings for all three discourse types (scientific: $R^2 = .900$; dramatic: $R^2 = .799$; spontaneous: $R^2 = .874$; in all cases, $P < .000001$). However, the mean ratings were significantly lower for the "masculine/feminine" ratings than for the straight/gay ratings (scientific .495 vs. .350; dramatic .493 vs. .365; spontaneous .516 vs. .399; in all cases $P <$ .00001 by correlated samples t-tests). These findings suggest that sounding "gay" and sounding "feminine" are related but not identical concepts. In particular, the differences between the two ratings ranged from -0.19 to $+0.37$ for the scientific passage, from -0.19 to $+0.40$ for the dramatic passage, and from -0.14 to $+0.31$ for the spontaneous speech samples. This substantial variation could be related to other factors not examined in this study. For example, larger differences might be associated with lower-pitched voices, since pitch is a much more reliable feature for distinguishing male (masculine) from female (feminine) voices. This hypothesis is tested in the next section.

PITCH ANALYSIS

Methodology

In order to test the hypothesis that differences between straight/gay ratings and masculine/feminine ratings are related to mean pitch, we measured the pitch characteristics of samples extracted from the 75 utterances heard by the listeners (i.e. 25 speakers $\times$ 3 speaking tasks). An intonation curve for each portion of the speech sample was extracted and then analyzed acoustically. The taped samples were digitized at 22050 Hz and analyzed using *Signalyze 3.12*, an acoustic analysis program, on a Power Macintosh 7100/66AV. For the Rainbow Passage, the following sentence taken from the middle of the recording was analyzed:

> *Since then physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbows.*

The average duration for this sentence was 6.99 seconds. One utterance was discarded because the speaker stumbled and corrected himself in the middle of the sentence. In the second task, the dramatic fire passage, the following sentence taken from the middle of the recording was analyzed:

> *There were all these people in the apartments upstairs screaming out of the windows; they must have been trapped.*

The average duration for this sentence was 4.99 seconds. For the responses to the open-ended question, the portion for analysis was selected by starting at a break

in production which coincided with the beginning of a sentence. The following six seconds of speech was then analyzed; typically, the analyzed portion stopped in the middle of a sentence.

Using the sentence from the Rainbow Passage, the pitch was analyzed using the "Temporal Structure Analysis" function in *Signalyze*. The accuracy of the pitch analysis was verified intermittently (at least twice for each speaker) by measuring the length of the period in the wave form at the corresponding point in time. The nature of Temporal Structure Analysis requires that the settings be adjusted for each speaker to define fairly carefully the lowest and highest expected pitches. The settings for each speaker were recorded. The utterances of the other two tasks were analyzed similarly using the same settings as used in the first task.

Results

The main statistical question is whether there is a significant correlation between mean pitch and either the "straight/gay" ratings or the "masculine/feminine" ratings. As discussed above, previous studies using small samples sizes have found little evidence of such an association.

There was no significant relationship between mean pitch and the "sounds gay" ratings for any of the discourse types (scientific: $R = -0.01$, $P = .95$; dramatic: $R = -0.17$, $P = .42$; spontaneous: $R = -0.02$, $P = .93$). The same was true of the relationship between mean pitch and the "masculine/feminine" ratings (scientific: $R = +0.15$, $P = .48$; dramatic: $R = -0.004$, $P = .99$; spontaneous: $R = -0.10$, $P = .64$).

We then addressed the question of whether an individual talker's mean pitch might be correlated with the difference between his "straight/gay" rating and "masculine/feminine" rating. These differences were marginally correlated with mean pitch for the scientific passage ($R = -0.37$, $P = .07$) and the dramatic passage ($R = -0.39$, $P = .06$), but nonsignificant for the spontaneous speech sample ($R = -0.31$, $P = .14$). Collapsing across all three discourse types, the correlation is highly significant: $R = -0.54$, $P = .007$. This significant negative relationship means that low-pitched voices showed a larger discrepancy between the "straight/gay" ratings and the "masculine/feminine" ratings, while higher-pitched voices showed a smaller discrepancy. In other words, listeners felt that low-pitched voices could be rated as "gay-sounding," but they were much more hesitant to rate them as "feminine-sounding." For higher-pitched voices, these ratings tended to be more similar. In retrospect, this might explain Terango's (1966) report that his more feminine-sounding voices had higher pitches than those rated as more masculine-sounding. Our results are compatible with findings that suggest a relationship between pitch and masculine/feminine ratings, even though there is no such relationship for straight/gay ratings.

Our failure to find a relationship between mean pitch and "sounds gay" ratings corresponds to the findings of Gaudio 1994, who had a much smaller sample.

They also agree with Avery & Liss 1996 in that their "less-masculine-sounding" and "more-masculine-sounding" voices did not differ in fundamental frequency. The finding of a strong negative correlation between mean pitch and discrepancies between the two kinds of ratings confirms that these scales are not equivalent. Thus, the parameter "masculine/feminine" elicits responses that are pitch-dependent: a low-pitched, gay-sounding voice is more likely to be judged "gay" than "feminine."

DISCUSSION

We have successfully created a data bank of gay- and straight-sounding voices. The listener judgments are highly consistent at the extreme ends, thus providing us samples of reliable tokens that do indeed sound gay or straight. In addition, we have several intermediate voices, thus providing a continuum from gay- to straight-sounding, necessary for any type of correlational analysis. Our findings are summarized as follows:

- The straight talker-subjects were rated as more gay-sounding in the scientific passage than in the other two conditions.
- The gay male listeners were more likely to rate a voice as "gay-sounding" than the other male listeners in our study. (The female listeners were in between.)
- The accuracy in correctly identifying the sexual orientation of the speakers ("gaydar") was significantly lower for the gay speakers than for the straight men. In particular, many gay men were judged as sounding straight. Indeed, the accuracy for correctly identifying gay voices was quite poor, which is surprising because we stacked the sample with gay-sounding voices. If indeed the phenomenon of auditory gaydar exists (i.e., if listeners can accurately identify gay people by their voices) then we should have seen even higher accuracy for gay men. We are not arguing against the existence of auditory gaydar. Rather, for those gay men who have a gay-sounding voice, listeners are obviously detecting particular features of such a voice, and responding to them; for these gay men, gaydar is quite accurate. However, there are many gay men who simply do not sound gay. Therefore, identifying the sexual orientation of men randomly chosen from the general population, by their voices alone, will not be very accurate.
- There are no significant correlations between average fundamental frequency and either "straight/gay" judgments or "masculine/feminine" judgments.
- The two constructs "straight/gay" and "masculine/feminine" are highly correlated in terms of listener judgments. However, the "masculine/feminine" ratings were on average lower than the "straight/gay" ratings. This is because, as our pitch analysis demonstrates, listeners were reluctant to rate a gay-sounding voice as "feminine" if it had a low fundamental frequency.

This is an important point because it demonstrates that different constructs will yield different results.

- Overall, we did not find differences in the ratings between the two listener groups (gay male vs. mixed). However, when we examined the ratings of the 25 individual voices separately, we did find twelve significant differences (when only seven would have been expected by chance alone). This means that there may be some individual voices that are perceived differently by different listener groups, but in our study we did not have enough of them to detect an overall effect.
- With regard to the open-ended responses, our subjective attempt to eliminate any lexical and pragmatic cues that would reveal the talker's sexual orientation was not successful. This has implications for future researchers who want to collect naturalistic speech samples for similar listening studies.

The consistency of these results leads us to conclude that our experimental approach is effective. Researchers who wish to replicate or extend this study need to consider that reading tasks, listener groups, constructs, sample sizes, and talker selection criteria will all influence the results in unique and interacting ways.

CONCLUSION

Ultimately, our research program is to use this collection of voices in order to explore the specific phonetic characteristics that allow listeners to rate a voice as "gay-" or "straight-sounding". Zwicky 1997 suggests there is a "menu" of cues from which speakers make different selections. All these features may sound gay, but in different ways depending on the feature, and the combinations thereof. It may be the case that in a male-as-norm culture that values masculinity, any feature that does not sound "straight" may be labeled as gay-sounding, even if just one speaker does it. Thus, one of the goals of our future research is to discover whether gay-sounding features tend to co-occur in the speech of gay-sounding men, or whether each gay-sounding voice tends to have its own combination of these features.

To what extent do female voices share these phonetic characteristics? As previously noted, the prevalent stereotype is that gay-sounding men speak like women. If it were possible to eliminate the most obvious male/female difference – fundamental frequency – would listeners be able to distinguish correctly between the gay-sounding male speakers and the female speakers? If so, then we must assume that gay men share only some of the characteristics of female speech, and that there may be some factors that are unique to gay-sounding speech. This raises the deeper question of whether gay-sounding speech is in fact modeled on women's speech, and if so, how and why some men and boys base their articulation on opposite-sex models.

To what extent do phonetic markers interact with visual cues? When we were brainstorming about possible phonetic cues to explore, we invited a small group

of acquaintances to watch several videos of *Q-TV*, a now defunct lesbian and gay TV talk show that aired briefly on a Canadian cable network in the late 1990s. Initially, the television screen was covered, forcing the participants to base their comments solely on the interviewees' voices. It is interesting to note that in one case, the participants could not readily identify the speaker's sexuality until the screen was exposed, at which point they quickly identified him as gay. Judgments about gay- and straight-sounding voices clearly will differ depending on whether the raters hear the voice alone, or in conjunction with body language.

How do listeners learn to detect these markers? Listeners might be able to distinguish gay- and straight-sounding voices, but are they all attending to the same cues and combinations of cues? Interestingly, when we watched the listener-subjects making their decisions, we noted that they made them at similar points in the tape. Rarely did any of the raters wait until the end of the 30-second voice sample to mark judgments on the response sheets. Purnell, Idsardi & Baugh 1999 found that their listener-subjects accurately judged a speaker's race with very short tokens (just by the word *hello*). We need to investigate further at which point each individual listener makes a decision. The phonetic features that occur prior to the decision may shed light on which cues were most salient to the listener. Another strategy, based on Crist 1997, is to have speakers mimic stereotypical gay speech, which may reveal which cues are drawn upon when listeners make judgments about the gayness or straightness of voices.

Does social contact with gay males make listeners more sensitive to the phonetic cues carried by gay-sounding voices? Although we did not ask our listeners about their contacts with gay men, we believe that this is not a major factor, since there were few differences between our gay and mixed listener groups even though the former are likely to have more frequent contact with gay men. It seems that most speakers, regardless of sexual orientation, are familiar with phonetic variation along a gay-straight continuum.

How much variation is there in the presence of the phonetic characteristics in different social settings? What is the effect of an experimental setting? Do these phonetic characteristics vary according to race, culture, class, or even language? Henton cautions about the use of specific measurements in making sweeping generalizations: "This conglomerate value should not be the mainstay for the argument here, since it is not entirely legitimate to compare values across studies, across languages, and across differing linguistic environments" (1989:303–4). Gilmore et al. 1992 found interaction effects between speaking tasks and speaker sex in their measures of fundamental frequency and pitch range. Other studies have also investigated the interaction of race and sex (Lass, Tecca et al. 1979, Lass, Almerino et al. 1980). Loveday 1981 investigated the difference between intonation and politeness formulas between English and Japanese male and female speakers. Henton & Bladon 1988 factored into their design the interaction effects of class and gender. And, of course, we found that discourse type affected our results in that our straight talkers sounded gayer in the scientific passage.

This article has been a foray into the methodological issues surrounding how researchers may go about investigating the specific phonetic characteristics of the gay voice. However, theorists in language and identity have reminded us that such experimental studies cannot be entirely separated from social theory. Indeed, many of the experimental studies we have cited undertheorize the issues involved in the acquisition or adoption of particular voices. Bergvall 1999 argues that any comprehensive theory of language and gender must consider the extent to which language is innate (biology), achieved (performance), and ascribed (ideology). In the past decade, language and gender researchers have adopted Butler's notion of PERFORMATIVITY: "Gender is the repeated stylization of the body, a set of repeated acts within a rigid regulatory frame which congeal over time to produce the appearance of substance, of a 'natural' kind of being" (1990:33). Cameron offers a summary of how Butler's framework can be applied to language: " 'Feminine' and 'masculine' are not what we are, nor traits we HAVE, but effects we produce by way of particular things we DO" (1997:49). Delph-Janiurek's (1999) study of vocal performances in English university teaching spaces reminds us that speaking is not the only relevant performative act; equally important are the acts of interpretation. Through participant observation and interviews, he concludes that "the everyday reading of voices clearly seems to involve attributing them to recognisable performances of roles and gendered and sexualised identities" (1999:150).

Equally interesting is the fact that the "gay voice" occurs within a homophobic culture. Cameron 1995 explores "verbal hygiene" whereby speakers overtly pass judgment on particular language forms in reference to established norms. She notes:

> There are penalties, ranging from being judged 'eccentric' to being ostracized or persecuted to being locked up and stripped of your rights, for constructing an identity in defiance of cultural prescriptions, or for failing to construct a proper identity at all. Butler's account, in other words, makes use of the idea of normativity – the 'highly rigid regulatory frame' she mentions in relation to gender. There are codes which define what is intelligible, acceptable and normal: individuals transgress those codes at their peril. (1995:16)

In North America, having a gay voice can be stigmatized both within and outside the gay community(ies). One only has to enter a gay chat room, or peruse the "men seeking men" advertisements, to confirm this observation. "Straight-acting" and "masculine" are desired prototypes for many gay men, behaviors that presumably encompass the way one talks. Outside the gay community, high school students perceived as being gay are open to cruel teasing and gay-bashing. Gay-acting men in homophobic workplaces are faced with hostility from their coworkers and denied promotions. The question remains as to why a male would construct an identity in defiance of rigidly controlled gender prescriptions, and

what consequences these individuals face because they resist or subvert these codes.

REFERENCES

* We regret to inform the sociolinguistics community that Greg Jacobs passed away suddenly on October 4, 2002. His commitment to research on gay men's speech and language has helped shape the direction the field has taken, and he will be sorely missed.

** We would like to acknowledge research support from the York University Contract Faculty Fund awarded to Greg Jacobs. We thank Michelle-Renée Carroll, William Kehoe, and John Duncan for assistance with the listening and reading tasks. We also thank Ruth King and the participants in the NWAV 1999 special session on language and sexuality, and the audience at our 2000 talk at the Canadian Lesbian and Gay Studies Association for their helpful comments. We are also grateful for the insightful comments of the three referees: Robin Queen, William Leap, and Rudolf Gaudio.

[1] In our linguistics classes over the years, one exercise we have performed is to elicit stereotypes of gay men's and lesbians' speech. Students almost immediately mention that gay men's voices sound like women's. When asked to elaborate, they refer to the high pitch and the "up and down" quality (i.e. highly variable intonation).

[2] The association between the subject matter of this passage (rainbows) and the icon that has recently become the symbol of gay pride is purely coincidental.

[3] We recognize the possibility that some differences that emerge between the two listener groups (gay and mixed) could be related to age differences between the groups, as noted in Table 1.

REFERENCES

Avery, Jack D., & Liss, Julie M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *Journal of the Acoustical Society of America* 99:3738–48.

Bennett, Suzanne, & Montero-Diaz, Luisa (1982). Children's perception of speaker sex. *Journal of Phonetics* 10:113–21.

_____, & Weinberg, Bernd (1979a). Sexual characteristics of preadolescent children's voices. *Journal of the Acoustical Society of America* 65:179–89.

_____, _____ (1979b). Acoustic correlates of perceived sexual identity in preadolescent children's voices. *Journal of the Acoustical Society of America* 66:989–1000.

Bergvall, Victoria L. (1999). Towards a comprehensive theory of language and gender. *Language in Society* 28:273–93.

Brend, Ruth M. (1975). Male-female intonation patterns in American English. In Barrie Thorne & Nancy Henley (eds.), *Language and sex: Difference and dominance*, 84–87. Rowley, MA: Newbury House.

Bresser A., & Günzburger, D. (1985). Voice recognizability of prepubertal boys and girls. *Progress Report Institute of Phonetics Utrecht* 10:25–32.

Britain, David (1992). Linguistic change in intonation: The use of high rising terminals in New Zealand English. *Language Variation and Change* 4:77–104.

Butler, Judith (1990). *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.

Cameron, Deborah (1995). *Verbal hygiene*. London and New York: Routledge.

_____ (1997). Performing gender identity: Young men's talk and the construction of heterosexual masculinity. In S. Johnson and U. H. Meinof (eds.), *Language and masculinity*, 47–64. Oxford: Blackwell.

Coleman, Ralph O. (1971). Male and female voice quality and its relationship to vowel formant frequencies. *Journal of Speech and Hearing Research* 14:565–77.

_____ (1976). A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech and Hearing Research* 19:168–80.

Crist, Sean (1997). Duration of onset consonants in gay male stereotyped speech. *University of Pennsylvania Working Papers in Linguistics* 4:53–70.

Delph-Janiurek, Tom (1999). Sounding gender(ed): Vocal performances in English university teaching spaces. *Gender, Place and Culture* 6:137–53.

Edelsky, Carole (1979). Question intonation and sex roles. *Language and Society* 8:15–32.
Fairbanks, Grant (1966). *Experimental phonetics: Selected articles*. Urbana: University of Illinois Press.
Fichtelius, Anna; Johansson, Iréne; & Nordin, Kerstin (1980). Three investigations of sex-associated speech variation in day school. *Women's Studies International Quarterly* 3:219–25.
Gaudio, Rudolf P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech* 69:303–18.
Gilmore, Stuart I.; Guidera, Angela M.; Hutchins, Susan L.; & van Steenbrugge, Willem (1992). Intra-subject variability and the effect of speech task on vocal fundamental frequency of young adult Australian males and females. *Australian Journal of Human Communication Disorders* 20:65–73.
Günzburger, D. (1984). Perception of some male-female voice characteristics. *Progress Report Institute of Phonetics Utrecht* 9:15–26.
_____; Bresser, A.; & ter Keurs, M. (1987). Voice identification of prepubertal boys and girls by normally sighted and visually handicapped subjects. *Language and Speech* 30:47–58.
Henton, Caroline G. (1989). Fact and fiction in the description of female and male pitch. *Language and Communication* 9:299–311.
Henton, Caroline, & Bladon, Anthony (1988). Creak as a sociophonetic marker. In Larry M. Hyman & Charles N. Li (eds.), *Language, speech and mind: Studies in honour of Victoria A. Fromkin*, 3–29. London & New York: Routledge.
_____, _____ (1985). Breathiness in normal female speech: Inefficiency versus desirability. *Language and Communication* 5:221–27.
Hillenbrand, James; Cleveland, Ronald A.; & Erickson, Robert L. (1994). Acoustic qualities of breathy vocal quality. *Journal of Speech and Hearing Research* 37:769–78.
Högberg, Jesper (1996). Some studies of the relationship between speech acoustics, articulation and phonetic structure. *Quarterly Progress and Status Report: Royal Institute of Technology, Department of Speech* 3:13–21.
Ingemann, Frances (1968). Identification of the speaker's sex from voiceless fricatives. *Journal of the Acoustical Society of America* 44:1142–44.
Jacobs, Greg; Rogers, Henry; & Smyth, Ron (1999). Searching for phonetic correlates of gay- and straight-sounding voices. Paper presented at New Ways of Analyzing Variation in English (and other Languages), Toronto.
_____, _____, _____ (2000). Phonetic correlates of gay- and straight-sounding voices. Paper presented at the Canadian Lesbian and Gay Studies Association, Edmonton.
_____, _____, _____ (2001). Searching for phonetic correlates of gay- and straight-sounding male voices. *Toronto Working Paper in Linguistics* 18:46–64.
Klatt, Dennis H., & Klatt, Laura C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87: 820–57.
Kulick, Don (2000). Gay and lesbian language. *Annual Review of Anthropology* 29:243–85.
Lass, Norman J.; Almerino, Celest A.; Jordan, Laurie F.; & Walsh, Jayne M. (1980). The effect of filtered speech on speaker race and sex identifications. *Journal of Phonetics* 8:101–12.
_____; Hughes, Karen R.; Bowyer, Melanie D.; Waters, Lucille T.; & Bourne, Victoria T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America* 59:675–78.
_____; Mertz, Pamela J.; & Kimmel, Karen L. (1978). The effect of temporal speech alterations on speaker race and sex identifications. *Language and Speech* 21:279–90.
_____; Tecca, John E.; Mancuso, Robert A.; & Black, Wanda I. (1979). The effect of phonetic complexity on speaker race and sex identifications. *Journal of Phonetics* 7:105–18.
Leap, William L. (1996) *Word's out: Gay men's English*. Minneapolis: University of Minnesota Press.
Lee, Alison; Hewlett, Nigel; & Nairn, Moray (1995). Voice and gender in children. In Sara Mills (ed.), *Language and gender: Interdisciplinary perspectives*, 194–204. London & New York: Longman.
Lerman, J. W., & P. H. Damsté (1969). Voice pitch of homosexuals. *Folia Phoniatrica* 21:340–46.
Linke, C. E. (1973). A study of pitch characteristics of female voices and their relationship to vocal effectiveness. *Folia Phoniatrica* 25:173–85.
Linville, Sue Ellen (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatrica et Logopaedica* 50:35–48.

Loveday, Leo. (1981). Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech* 24:71–89.
Markel, Norman N.; Prebor, Layne D.; & Brandt, John F. (1972). Biosocial factors in dyadic communication: Sex and speaking intensity. *Journal of Personality and Social Psychology* 23:11–13.
Moonwomon, Birch (1985). Toward a study of lesbian speech. In S. Brammer et al. (eds.), *Proceedings of the First Berkeley Conference on Women and Language*, 96–107. Berkeley: Women and Language Group. Reprinted in Anna Livia & Kira Hall (eds.), *Queerly phrased: Language, gender and sexuality*, 202–13. Oxford: Oxford University Press, 1997.
Purnell, Thomas; Idsardi, William; & Baugh, John (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18:10–30.
Rogers, Henry, & Ron Smyth (2000). Will and Jack: Gay but different. Paper presented at Lavender Languages, Washington, D.C.
_____; Smyth, Ron; & Jacobs, Greg (2000a). Vowel and sibilant duration in gay- and straight-sounding male speech. Paper presented at the International Gender and Language Association, Stanford.
_____, _____, _____ (2000b). The sibilant as a cue in distinguishing gay- and straight-sounding male speech. Paper presented at the Canadian Linguistic Association, Edmonton.
Sachs, Jacqueline (1975). Cues to the identification of sex in children's speech. In Barrie Thorne & Nancy Henley (eds.), *Language and sex: Difference and dominance*, 152–71. Rowley, MA: Newbury House.
_____; Lieberman, Philip; & Erickson, Donna (1973). Anatomical and cultural determinants of male and female speech. In Roger W. Shuy & Ralph W. Fasold (eds.), *Language attitudes: Current trends and prospects*, 74–84. Washington, DC: Georgetown University Press.
Schwartz, Martin F. (1968). Identification of speaker sex from isolated, voiceless fricatives. *Journal of the Acoustical Society of America* 43:1178–79.
_____, & Rine, Hellen E. (1968). Identification of speaker sex from isolated, whispered vowels. *Journal of the Acoustical Society of America* 44:1736–37.
Smyth, Ron, & Rogers, Henry (2000). Relationships among phonetic factors used in making gay-straight judgements. Poster presented at New Ways of Analyzing Variation in English (and Other Languages), Lansing.
Snidecor, John C. (1951). The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech and Hearing Disorders* 16:44–52.
Terango, Larry (1966). Pitch and duration characteristics of the oral reading of males on a masculinity-femininity dimension. *Journal of Speech and Hearing Research* 9:590–95.
von Raffler-Engel, Walburga, & Buckner, Janis (1983). A difference beyond inherent pitch. In Betty Lou Dubois & Isabel Crouch (eds.), *Proceedings of the conference Sociology of the Languages of American Women*, 115–18. San Antonio, TX: Trinity University.
Weaver, Andrew Thomas (1924). Experimental studies in vocal expression. *Journal of Applied Psychology* 8:23–56.
Weinberg, Bernd, & Bennett, Suzanne (1971). Speaker sex recognition of 5- and 6-year-old children's voices. *Journal of the Acoustical Society of America* 50:1210–13.
_____, & Zlatin, Marsha (1970). Speaking fundamental frequency characteristics of five- and six-year-old children with mongolism. *Journal of Speech and Hearing Research* 13:418–25.
Zwicky, Arnold. M. (1997). Two lavender issues for linguists. In Anna Livia & Kira Hall (eds.), *Queerly phrased: Language, gender, and sexuality*, 21–34. New York: Oxford University Press.

(Received 9 October, 2001; accepted 14 March, 2002)