

HEARING “GAY”: PROSODY, INTERPRETATION, AND THE AFFECTIVE JUDGMENTS OF MEN’S SPEECH

EREZ LEVON

New York University

ABSTRACT: This article describes a controlled experiment designed to determine what people listen to specifically when judging a speaker’s sexuality. Four experimental stimuli were produced by digitally shortening the syllable duration and narrowing the pitch of one male speaker reading a passage. Listeners rated various combinations of the four stimuli on 10 affective scales, including straight/gay and effeminate/masculine. Altering the two variables was insufficient to alter listeners’ perceptions of the speaker’s sexuality to a level of significance. However, significant correlations between the different attitudinal scales illustrated that perceptions of sexuality are ideologically linked to other perceptions of personality and personhood.

IN THIS ARTICLE, I examine whether two prosodic variables, pitch range and sibilant duration, affect the perceptual identification of the sexuality of a male speaker. Both of these variables have been widely discussed in the literature on language and sexuality and have been correlated, to varying degrees of success, with the perception of gayness in men. Because of the popular perception that gay men’s speech is characterized by high levels of pitch variability, numerous studies have attempted to identify wide pitch range as an index of gayness (e.g., Gaudio 1994; Rogers and Smyth 2003). None of these, however, have been able to demonstrate a direct and unqualified correlation between a speaker’s pitch range and his perceived sexuality. Research into sibilant duration has been somewhat more successful in this regard. Associated with the popular stereotype of a “gay man’s lisp,” several studies have been able to link a speaker’s sibilant durations and perceptions of that speaker’s sexuality (e.g., Linville 1998; Rogers, Smyth, and Jacobs 2000).

In addition to realizing, at best, mixed results, previous research has recently been subjected to a critical analysis of its underlying methodologies. This critique, developed primarily in Kulick (2000) and Cameron and Kulick (2003), asserts that much of this research relies upon static concep-

tualizations of reified identity categories, and thus essentializes both the social categories being discussed (e.g., "gay" or "straight") and the linguistic practices associated with them. In other words, much previous research has at least implicitly assumed homogeneity of linguistic practice among people who identify as either "gay" or "straight" and has subsequently analyzed this linguistic practice as a natural outcome of speakers' sexual identifications. Kulick (2000, 265) argues that this type of research has "started out by 'knowing' the identities whose very constitution ought to have been precisely the issue under scrutiny." Rather, Kulick (2000) and Cameron and Kulick (2003, 114) advocate research into the "social semiotic of desire," that is, research into the various iterable social signs (Derrida 1995) that serve to index desire and sexuality. Shifting the focus from "identity" to social practice allows research not to treat "sexuality" as a static category. By examining the various ways in which desire and sexuality are materialized through language, research will be able to better develop an understanding of how individuals personally negotiate the desire to present a distinctive psychological self (Irvine 2001) and the social and linguistic ideologies of sexuality and identity that limit that agency.

Eckert (2002, 105) reminds us that we cannot divorce semiotic resources from their social histories, histories which "relat[e] what an individual does to things that categories of people do." While Eckert acknowledges that these categories may come in all shapes and sizes, she argues that the popular identity categories (e.g., *gay*, *straight*, *lesbian*) possess a certain ideological status which renders them particularly salient social locations around which individuals can position themselves. In other words, though a particular individual may feel that no popular identity category matches his or her own conception of self, the fact that these categories and their associated semiotic practices are widely recognized and understood makes them effective social positions to adopt, even if only temporarily. Cameron and Kulick (2003), in fact, make a similar claim, though they state it somewhat differently. While certainly arguing for a methodological shift toward the study of social practice, they also claim that "both ideology and practice are real social phenomena; each of them is interesting, and the study of language and any form of identity is to a considerable extent the study of the relationship between them" (136). They argue that the crucial focus of research on language and sexuality is to examine the "complex and creative ways" in which "speakers 'take up' ideological resources to construct identities in practice" (136).

In this article, I propose an experimental design which attempts to incorporate the arguments put forth by both Cameron and Kulick (2003) and Eckert (2002), shifting the focus from identity categories to the semiotic

practices that are indexical of those categories. In this experiment, the speech of a single individual is digitally manipulated and tested in order to determine whether variation in pitch range and sibilant duration alters listeners' perceptions of his sexuality. This allows me to investigate the existence of specific ideological resources available to both speakers and listeners in the construction or perception of a gay identity while avoiding the assumption that these linguistic practices are an essential or universal characteristic of gay individuals.

PREVIOUS RESEARCH

PITCH RANGE. Gaudio (1994) examined the speech of gay and straight men in an effort to correlate listeners' perceptions of sexual identity with the actual sexual identity of the speakers. In his study, 13 raters listened to 16 segments of talk from 8 speakers (4 self-identified as gay men, 4 as straight men). After hearing these samples, the listeners rated each speaker on the following 4 affective scales: straight/gay, effeminate/masculine, reserved/emotional, and affected/ordinary. Gaudio found that his listeners (10 women and 3 men) were able to accurately guess the sexual orientation of the speakers 7 out of 8 times. However, when Gaudio analyzed the speakers' data for prosodic differentiation between the straight and gay speakers that may have served as a salient cue to the listener population, he found no significant difference with respect to pitch range.¹ Gaudio concluded that though pitch range may play a role in the identification of the sexual orientation of a speaker, it does not do so in isolation and can only be understood in relation to the entire speech signal and the context of the talk.²

Rogers and Smyth (2003) conducted a series of experiments to examine whether differences in mean pitch (F_0) and pitch variability (i.e., pitch range) affected a listener population's perceptions of the sexuality of 25 speaker-subjects.³ The 25 speaker-subjects were all men, 17 self-identified as gay, 8 as straight. For the first experiment, the listener population comprised 46 people, whom Rogers and Smyth describe as "fourteen gay males and thirty-two people from a mixed group of men and women of unknown sexual orientation" (1855). These listeners heard each of the 25 speakers read a portion of the *Rainbow Passage* (a standardized passage about rainbows often used in phonetic and dialectology research; see appendix A), with purportedly low affect; a passage about a burning building, with somewhat higher affect; and a description of an interesting event in each of the speaker-subjects' lives. In this initial experiment, Rogers and Smyth found that there was no significant correlation between the mean pitches or pitch ranges of the speakers

and the listeners' gayness ratings of the speakers.⁴ To follow up, Rogers and Smyth conducted a second experiment in which only the readings of the Rainbow Passage were used. From these samples, Rogers and Smyth removed all segmental and voice-quality information. A second listener population (a new group of 44 people of unidentified sexual orientation) therefore heard only what sounded like a schwa vowel pronounced with a varying pitch. In this test as well, however, Rogers and Smyth found no significant correlation between mean pitch or pitch range and the gayness ratings of the speakers. Finally, Rogers and Smyth conducted a third experiment in which they again presented a listener population (a third group composed of 9 linguists) with the stimuli from which all segmental and voice quality information had been removed. In this final test, the listeners were not asked to assess the speakers' sexualities but, instead, to provide a rating on the relative height of the pitch of the utterance and the intonational variability of the utterance. This third experiment therefore provided data on the perceived mean pitches and perceived pitch ranges of the speakers.

Multiple regression analyses were conducted which compared the results of perceived mean pitch and pitch range obtained in the third experiment to the results of perceived gayness ratings obtained from the first and second experiments. These analyses showed a significant effect between perception of pitch range and perception of sexuality but only after all segmental and voice quality information had been removed. This finding leads Rogers and Smyth to argue that while there may be no real acoustic correlate between actual pitch properties of a speaker and perceptions of that speaker's sexuality, listeners have an expectation that there will be a correlation and may therefore base their judgments on this expectation.

SIBILANT DURATION. Rogers, Smyth, and Jacobs (2000) examined sibilant duration as a potential cue for the perception of a gay-male subjectivity. In their study, 46 listeners heard a group of 25 men (17 self-identified as gay, 8 as straight) read the Rainbow Passage. The listeners were then asked to rate the speakers on a scale of "gayer-sounding" versus "straighter-sounding." Rogers, Smyth, and Jacobs found that those speakers who were rated as "gayer-sounding" had significantly longer mean normalized durations for both /s/ and /z/. Additionally, they found that those speakers rated as "gayer-sounding" also had significantly higher peak frequency values for these same fricatives.

Finally, Linville (1998) examined the contribution of /s/ duration and /s/ peak frequency values in the speech of 9 speakers (5 self-identified as gay, 4 as straight) to the perception of their sexual identifications. The listeners in her study, a group of 25 women, correctly identified the sexual orientation

of the speakers 79.6% of the time. Multiple regression analysis of the perceptual results revealed that correct identification of actual sexual orientation of the speakers could be statistically correlated with the characteristics of /s/ observed. The gay speakers in the study were all observed to have longer /s/ frication (106 msec was the shortest for a gay speaker, 103 msec was the longest for a straight speaker). The gay speakers' /s/ peak frequencies varied from 5882 Hz to 7333 Hz, while the straight speakers' /s/ peak frequencies varied from 4722 Hz to 6882 Hz. Though there is some obvious overlap between the straight and gay speakers in terms of /s/ peak frequency values, Beta weights in the multiple regression analysis indicated that both /s/ duration and /s/ peak frequency made a statistically substantial contribution to the identification of the sexual orientation of the speaker.

EXPERIMENTAL DESIGN

The experimental design reported here attempts to address what I believe to be the theoretical and methodological drawbacks of the previous research. All four of the experiments summarized above, as well as the majority of research on the linguistic perception of sexuality, adopts a two-step methodology. Two groups of speakers, one self-identified as "gay" and one self-identified as "straight," are used to obtain the experimental stimuli. These stimuli are then presented to listeners, who rate the speakers on various affective scales related to gender and sexuality. In all of the studies mentioned above, the listeners were shown to be remarkably accurate at correctly identifying the sexual identification of the speakers. The researchers then examined the stimuli recordings of the "gay" speakers and the "straight" speakers in an attempt to identify systematic linguistic differences between the two groups that would explain the perceptual accuracy of the listeners. This methodology has two main problems. First, it places the identity categories being studied at the center of the inquiry, and, second, it forces the linguistic analysis of the experimental stimuli to be necessarily post hoc. Below, I describe these problems in some more detail, while simultaneously discussing my methodological proposals for avoiding them.

Methodologies that employ a group of self-identified straight men to obtain a sample of "straight" speech and another group of self-identified gay men to obtain a sample of "gay" speech run dangerously close to assuming a necessary and fixed relationship between the constructs gay and straight and the language style associated with these social positionings. This methodology implicitly assumes that the practices of a group of people who self-identify as gay will be internally homogenous, as well as mutually exclusive and maximally

differentiated from a group of people who happen to self-identify as straight. By examining the speech of a small group of people and allowing that to stand as representative of “gay” or “straight” speech, these methodologies ignore the reality that linguistic practice is highly variable, both within and between speakers. Cameron and Kulick (2003) remind us that we cannot ignore the fact that individuals in their everyday lives change the way they talk across a range of situations, in which different social identifications are more salient than others. While topic- and audience-conditioned linguistic variation has been acknowledged in sociolinguistics for decades (Giles and Powesland 1975; Bell 1984, 2001), the majority of research on language and sexuality has ignored it (though cf., notably, Podesva, Roberts, and Campbell-Kibler 2002; Podesva 2003; and the discussion in Bucholtz and Hall 2004).

With respect to variation between speakers, while a certain individual’s linguistic practice may match the relevant stereotype exactly, another’s may not. Cameron and Kulick (2003, 136) warn that any claims about a gay speaking style “should not be taken as descriptive generalizations about the behavior of real individuals in particular gay communities, for those individuals may not in practice display the speech characteristics that make up the ideological construct.” By ignoring the reality of linguistic variation, these methodologies reify the identities which they attempt to examine.

I therefore propose a methodology that does not place these reified identity categories at the center and does not rely upon a comparison between representative “gay” speakers or “straight” speakers. Instead, the speech of one person, otherwise unrelated to the project, was used to obtain the experimental stimuli. This speaker, a white male in his mid-20s, was recorded reading a short passage, approximately 71 seconds in length. This passage was a neutral narrative about a typical New York City topic (i.e., a crowded subway platform in Manhattan) that was designed to control for the number and position of sibilants: the voiceless alveolar fricative /s/, the voiced alveolar fricative /z/, and the voiceless palato-alveolar fricative /ʃ/ (see appendix B for text of the passage).⁵ The sibilant composition of the passage is presented in table 1.⁶

The speaker’s recorded reading of this passage was presented to a pretest group of 10 men and women, who rated this original recording on the scales straight/gay and effeminate/masculine. This pretest group was composed of linguistics graduate students at New York University, who were familiar with the general topic of research (i.e., the sociolinguistic perception of sexuality) but unfamiliar with the speaker or his sexual identification. All 10 listeners rated this original recording as both “extremely gay” and “extremely effeminate.” Because I was not testing differences between “gay” and “straight” speech but rather focusing on the specific effect of particular

TABLE 1
Sibilants in Stimulus Text by Type and Position

	<i>Word Position</i>		<i>Syllable Position</i>		<i>Syllable Stress</i>		<i>Total</i>
/s/	Initial	13	Onset	20	+stress	25	39
	Medial	14	Coda	19	-stress	14	
	Final	12					
/z/	Initial	0	Onset	2	+stress	5	20
	Medial	7	Coda	18	-stress	15	
	Final	13					
/ʃ/	Initial	10	Onset	13	+stress	1	14
	Medial	3	Coda	1	-stress	13	
	Final	1					
TOTAL	Initial	23	Onset	35	+stress	31	73
	Medial	24	Coda	38	-stress	42	
	Final	26					

linguistic variables on the perception of sexuality, this recording, which was judged in its original form as “extremely gay” and “extremely effeminate,” was used to create additional experimental stimuli. This was the primary question: if pitch ranges and sibilant durations of this person’s speech were altered, would listeners’ judgments of his speech change? By changing nothing other than these specific linguistic features, would some listeners rate this speaker as straight or masculine? Formulating the research questions in this way avoids using identity categories as primitives, and instead allows us to examine whether controlled variation of discrete linguistic features could affect the perceived sexuality of the same person. The crucial difference between the current methodology and those used in previous research is the importance of the sexual identification of the speaker in the empirical design. In earlier research, linguistic analyses were conducted separately on those who self-identified as “gay” or “straight” in an effort to find differences between the two. Here, the PERCEPTION of a gay identity is used as the starting point, and linguistic manipulations are conducted in order to determine whether those PERCEPTIONS could be changed. The actual self-identification of the speaker is never invoked.

The methodology proposed here also has the benefit of addressing the other criticism of previous research, namely that the linguistic analyses were necessarily post hoc. Those methodologies which present the speech of differentially identified speakers encounter potentially insurmountable difficulties with respect to the identification of which specific linguistic features are being evaluated. Though raters have been able to consistently, and correctly, identify the sexuality of speakers, as in Gaudio (1994), Linville (1998), and Smyth, Jacobs, and Rogers (2003), the methodologies that these researchers

employ require that their analyses of which features listeners are attuning to are an empirical afterthought, largely motivated by the ideological suppositions of the researchers themselves.

Consider a fictional example of this type of methodology. A listener is presented with a short recording of a speaker, and then asked to rate whether that speaker is gay or straight. Following this, that same speaker is presented with another short recording of a different speaker, and asked to rate whether that second speaker is straight or gay. Let’s assume that the listener rates the first speaker as straight and the second speaker as gay. What was that listener attuning to? Was that listener paying attention to the pitch properties of the two speakers (their pitch ranges or levels of pitch dynamism)? Or was that listener paying attention to some segmental information (the length of their sibilants, or their stop-releases or their vowel placements)? Or was that listener paying attention to voice quality information (breathy voice or falsetto voice)? Or was it some combination of these things? These questions could continue *ad infinitum*. Exhaustively and conclusively identifying the salient linguistic cues from among the hundreds of different features and combinations of features that listeners are attuning to is nearly impossible in a methodology that does not control for variation between the stimuli.

Trying to address this, the original speaker’s recording of the passage was digitally altered with respect to the two variables, sibilant duration and pitch range. Because I was beginning with a passage that had been identified as “extremely gay” and “extremely effeminate,” both sibilant durations and pitch ranges were reduced in the digital manipulations, to assess whether there is a direct relationship between sibilant length and/or width of pitch range and ratings of gayness and effeminacy.⁷ Sibilants were shortened across the board by 17%,⁸ the threshold identified by Rogers, Smyth, and Jacobs (2000) and Linville (1998) at which listeners’ perceptions of sexuality changed. This allowed me to test whether previous findings of significance with respect to sibilant duration would hold when everything else was controlled for. This alteration used a relative measure (17%) rather than an absolute measure (e.g., 45 msec) in order to preserve the normal phonetic variation within the original speech sample. Variation in length of segments has been shown to be dependent upon word position (Oller 1973; Klatt 1974; Lehiste 1975; Umeda 1977), syllable position (Oller 1973), phrase position (Klatt 1975), syllable stress (Fry 1958; Lehiste 1975; Umeda 1977), and segment type (Nooteboom 1997). Using a relative measurement to alter the sibilant durations allows phonetic distinctions within the original speech sample to be preserved.

In pitch range, there does not exist as clear-cut a benchmark between perceptions of sexuality. Some researchers have reported differences as small

as 8%, while others have reported differences as large as 45%. In previous research that I conducted on representations of gay characters in film and television (Levon 2004), I found that the differences in pitch range between the natural speech of the actors I analyzed and the speech of their characters fell within the 22–28% range. For this reason, I reduced the pitch range of the original speech sample by 25%. This was done using the Pitch Manipulation editor in Praat 4.1.15. The original speech sample was first divided into intonational phrases using the Tone Break Indices (ToBI) coding system. Within each intonational phrase, the central pitch range (or 95.4% of the total pitch range) was calculated in semitones. Henton (1989, 1995) argues extensively for the calculation of pitch range in the logarithmic semitone scale, rather than the linear Hertz scale, insofar as the ear perceives pitch logarithmically. I adopt Henton's proposal here, and the reader is referred to her work for a complete justification. Additionally, Jassem (1971) argues that a restricted pitch range, composed of four standard deviations from the mean pitch of an utterance, should be used when investigating pitch variability. This ensures that the measured range will be more representative of the actual fluctuations in pitch, excluding extreme outliers and/or rogue data. I also adopt Jassem's proposal here. Once the pitch range of each of the intonational phrases was identified, the overall contour was flattened by 25%. The original shape of the contour is thus preserved, as are the relative rates of pitch dynamism between the peaks and valleys. This allowed me to isolate only pitch range as the variable being examined.⁹

Four experimental stimuli were created from the manipulations of the original passage, as shown in table 2, where the only differences between each passage are their respective settings for the two variables.¹⁰ In contrast to the majority of previous research on the subject (though reminiscent of Lambert et al. 1960), this methodological design allows us to attribute any differences in listeners' affective ratings of the passages to the specific variables under consideration.¹¹

The analysis is drawn from the affective judgments of 121 college undergraduates in Manhattan. These 121 listeners were all students at either New

TABLE 2
Experimental Stimuli

<i>Stimulus</i>	<i>Pitch Range</i>	<i>Sibilant Duration</i>
A (unaltered)	[–narrow]	[–short]
B	[+narrow]	[–short]
C	[–narrow]	[+short]
D	[+narrow]	[+short]

York University or the City University of New York, enrolled in linguistics, anthropology, or women's studies courses. None of these listeners knew me before the experiment or had any knowledge of my research. Prior to the experiment, the listeners were asked to participate in a listening task, in which they would hear passages and then be asked to rate the speaker on a variety of personality scales. The listener population ranged in age from 18 to 45 and was made up of approximately 65% women and 45% men. Following each passage, the listeners rated the speaker of that passage on each of the ten affective scales shown in table 3. These scales attempt to measure opinions of what Scherer (1972) labels the *COMPETENCE* of the speaker (scales 2 and 8) and the *BENEVOLENCE* of the speaker (scales 1, 5 and 9), as well as various characteristics related to the sexual and gender identification of the speaker (scales 3, 4, 6, 7, and 10; adapted from Gaudio 1994). The average ratings of the four stimulus passages are given in table 4.

The listener population was divided into six groups. Ideally, a population of 24 groups would have been used, so that each group could hear each of the four stimuli in all possible combinations. This, however, would have necessitated a much larger subject population than I was able to obtain. Instead, I chose to run the experiment as if it were three smaller experiments, in the manner shown in table 5. Each of the six listener groups heard passage D, with shortened sibilants and narrowed pitch range, as well as either passages A, B, or C. Order of presentation was also varied to test its significance on listener ratings. Quantitative analysis was then conducted on those groups who heard the same passages, namely groups 1a and 1b, groups 2a and 2b, and groups 3a and 3b.¹² These quantitative analyses assessed the effect, if any, of the feature values (i.e., + or -) of the variables considered. If an effect was borne out, then we could conclude that this effect was elicited by that

TABLE 3
Affective Scales

	1 = extremely		4 = neutral			7 = extremely		
1. generous	1	2	3	4	5	6	7	greedy
2. lazy	1	2	3	4	5	6	7	hardworking
3. prudish	1	2	3	4	5	6	7	promiscuous
4. effeminate	1	2	3	4	5	6	7	masculine
5. aloof	1	2	3	4	5	6	7	friendly
6. straight	1	2	3	4	5	6	7	gay
7. neat	1	2	3	4	5	6	7	messy
8. savvy	1	2	3	4	5	6	7	naive
9. kind	1	2	3	4	5	6	7	mean
10. genuine	1	2	3	4	5	6	7	fake

TABLE 4
Average Ratings for the Four Stimulus Passages

<i>Affective Scales (1/7)</i>	<i>Passage A</i>	<i>Passage B</i>	<i>Passage C</i>	<i>Passage D</i>
Generous/Greedy	4.30	4.45	4.53	4.28
Lazy/Hardworking	4.13	3.95	3.78	4.15
Prudish/Promiscuous	3.93	3.82	4.42	4.13
Effeminate/Masculine	2.61	3.29	2.73	2.96
Aloof/Friendly	3.85	3.51	3.73	3.56
Straight/Gay	5.35	4.66	5.50	4.92
Neat/Messy	3.26	3.74	3.49	3.48
Savvy/Naive	3.09	3.11	3.46	3.17
Kind/Mean	4.30	4.34	4.78	4.23
Genuine/Fake	3.96	4.50	4.68	4.22

TABLE 5
Presentation of Stimuli to Listeners

Group 1a	A, D	Group 1b	D, A
Group 2a	B, D	Group 2b	D, B
Group 3a	C, D	Group 3b	D, C

particular variable setting. For example, if speakers rated passage C, with [–narrow] pitch range and [+short] sibilant durations, as consistently gayer than passage D, with [+narrow] pitch range and [+short] sibilant durations, we could conclude that wide pitch range is a linguistic feature that these listeners ideologically associated with gay-identified speech. Finally, a correlation analysis was conducted for all three groups, which examined the extent to which ratings on a given scale (e.g., effeminate/masculine) are positively or negatively correlated with ratings on another scale (e.g., straight/gay).

RESULTS

LINGUISTIC VARIABLES. The quantitative analyses, which compared the three main groups, were conducted using the Linear Mixed Model Analysis function in SPSS 12. This analysis is a type of multivariate analysis of variation (MANOVA), which assesses the extent to which certain independent factors had an effect on the attestation of certain dependent factors. In the current study, the independent factors considered were stimulus (i.e., which recording the listeners were rating) and order (i.e., whether they heard that particular recording first or second in the testing session). The dependent

variables under consideration were the 10 affective scales listed in table 3. A linear mixed model analysis begins by assuming a null hypothesis—that the independent variables have no statistical effect on the dependent variables. The analysis then builds a statistical model to test this assumption, and outputs a probability figure. This probability represents the likelihood that the null hypothesis is correct, that is, the likelihood that any patterning seen in the data is equally due to chance coincidence as it is to an actual effect of the independent variables on the dependent variables. In the social sciences, the threshold of $p \leq 0.05$ is the standard for rejection of the null hypothesis. In other words, if the quantitative analysis reports that an independent variable had an effect on a dependent variable, and the model is at least 95% sure that this effect is not due to chance, we can conclude that the independent variable is statistically significant with respect to the attestation of the dependent variable.

Table 6 lists the results of a linear mixed model analysis of the three groups, where it was tested whether stimulus had a statistically significant effect on the listeners’ ratings on the 10 affective scales. Concerning ourselves first only with the probability figures for group 1, recall that group 1 heard passages A and D, where passage A had [–narrow] pitch range and [–short] sibilant durations and passage D had [+narrow] pitch range and [+short] sibilant durations. According to the quantitative analysis, there was no statistically significant difference between how listeners rated the speaker on the 10 affective scales, whether or not they were rating the recording with [–narrow] pitch range and [–short] sibilant durations or the recording with [+narrow] pitch range and [+short] sibilant durations. In terms of the straight/gay scale, the probability figure of 0.118 does not achieve the

TABLE 6
Linear Mixed Model Results

	<i>Group 1</i>		<i>Group 2</i>		<i>Group 3</i>	
	F (1, 87)	<i>p</i>	F (1, 67)	<i>p</i>	F (1, 69)	<i>p</i>
Generous/Greedy	0.560	0.814	0.337	0.564	1.158	0.286
Lazy/Hardworking	0.157	0.693	0.011	0.915	2.142	0.148
Prudish/Promiscuous	0.213	0.646	0.706	0.404	0.119	0.731
Effeminate/Masculine	3.312	0.080	1.080	0.302	0.169	0.682
Aloof/Friendly	3.112	0.081	0.156	0.694	0.000	0.988
Straight/Gay	2.498	0.118	0.504	0.480	0.680	0.412
Neat/Messy	0.063	0.802	0.651	0.423	0.194	0.661
Savvy/Naive	0.032	0.858	0.106	0.746	0.150	0.700
Kind/Mean	0.008	0.930	0.004	0.951	1.233	0.271
Genuine/Fake	0.002	0.965	0.159	0.691	0.201	0.656

required threshold. The only affective scales that do come close to showing significant effects are the effeminate/masculine and aloof/friendly scales, with probabilities of 0.080 and 0.081, respectively. While it may be possible to argue for a statistical trend approaching significance for some of the scales (e.g., effeminate/masculine), none of these figures meet the 0.05 threshold required for rejection of the null hypothesis.

The results are similar for group 2, which heard passage B ([+narrow] pitch range, [-short] sibilant duration) and passage D, and group 3, which heard passage C ([-narrow] pitch range, [+short] sibilant duration) and passage D. The probability values for all 10 affective scales in both of these groups show that the properties of the pitch range or sibilant durations of the recording had no significant effect on the listeners' ratings. In terms of the straight/gay scale in group 2, the analysis concludes that there is approximately a 50/50 chance that the ratings are due to the variable setting versus chance. Similarly, in terms of the effeminate/masculine scale in group 3, the analysis predicts that it is more likely that the ratings on this scale are due to chance rather than any effect of the independent variables.

Overall, the results of table 6 indicate that no statistically significant difference was found for the listeners' perceptions of the sexuality of the speaker for either variable in either setting. In other words, the listeners judged the speaker as having the same sexuality whether his pitch range was wide or narrow or whether his sibilants were long or short. For all three groups, changing either the pitch range or sibilant duration or both had no effect on listeners' judgments of the speaker across all 10 affective scales.

What this finding suggests is that in this experiment sibilant duration or pitch range alone do not contribute to the listeners' perceptions of the sexuality of the speaker. This stands in contrast to what some previous researchers have said about these variables. Yet, the empirical fact remains, attested to in much previous research, that listeners are remarkably accurate in judging the sexuality of speakers. And this fact remains puzzling. It is entirely possible that the association of a given speaker's linguistic practice with a gay subjectivity may be a gestalt phenomenon, where multiple auditory cues act in concert to produce a holistic index of gayness.¹³ Yet it is precisely the task of research on language and sexuality to determine whether this seemingly holistic phenomenon can be pried apart. Though the listeners in the current study were not affected by changing the pitch range or sibilant durations, perhaps they would have been affected by changing the vowel durations, the mean pitch, the formant frequencies, or some of the numerous other variables that have been identified as being part of a stereotypically gay speaking style. Perhaps, it is also insufficient to alter only one or two variables; rather these variables need to be considered in clusters. Maybe if the speaker's sample

had been altered by not only narrowing the pitch range and shortening the sibilants but also by lowering the mean pitch and slowing the rate of speech, listeners' judgments would have differed.

I am not claiming that sibilant duration and pitch range play no part in listeners' assessments of a speaker's sexuality. What I can claim, however, is that, for the time being, the null hypothesis of the current research cannot be rejected. These two variables, pitch range and sibilant duration, when presented alone with all else remaining constant were insufficient to alter listeners' perceptions of the sexuality of the speaker. Further research is required to identify which features or combinations of features would do so. And the only way to determine which features and combinations of features come together to index gayness for listeners is to continue conducting experiments that isolate specific linguistic variables and test these variables in the manner discussed here.

AFFECTIVE SCALES. While the linguistic variables had no significant effect on listeners' ratings across the 10 affective scales, there were significant correlations between these different personality characteristics themselves. The correlations highlight the fact that gender and sexuality as social positions do not exist in isolation; rather, they are ideologically linked with a host of personality characteristics that come together to form the total social perception of sexuality in a given community.

Correlation analysis was conducted on the combined data sets from groups 1, 2, and 3. This analysis assesses the extent to which two variables are linearly dependent on one another. In other words, a correlation analysis determines whether a given score on the straight/gay scale, for example, will likely co-occur with a given score on the effeminate/masculine scale. In assessing the possibility of a relationship between two variables, the quantitative model generates a Pearson's correlation coefficient.¹⁴ This figure is a number between -1 and 1. Values on the positive side of the scale indicate a direct relationship between the variables, that is, a high score for one variable co-occurs with a high score for the other variable. Values on the negative side of the scale indicate an inverse relationship, that is, a high score for one variable co-occurs with a low score for the other variable. A Pearson's figure of 0 indicates a lack of correlation between the two variables. The closer the figure is to either 1 or -1, the stronger the correlation between the two variables. Table 7 lists the results of the Pearson's correlation analysis obtained using the correlation function on SPSS 12 on all ten affective scales.

The results in table 7 show a wide variety of personality correlations that are not necessarily relevant to the present discussion. For example, greediness is correlated with promiscuity and meanness, while naïveté is correlated

TABLE 7
Correlation Analysis of All Ten Affective Scales

	<i>Generous/ Greedy</i>	<i>Lazy/ Hard</i>	<i>Prudish/ Promiscuous</i>	<i>Effem/ Masc</i>	<i>Aloof/ Friendly</i>	<i>Straight/ Gay</i>	<i>Neat/ Messy</i>	<i>Savvy/ Naive</i>	<i>Kind/ Mean</i>	<i>Genuine/ Fake</i>
<i>Generous/Greedy</i>	1	-0.260**	0.142*	-0.188**	-0.373**	-0.075	0.106	0.217**	0.592**	0.313**
<i>Lazy/Hardworking</i>	0.260**	1	0.044	0.081	0.171**	-0.100	-0.139*	-0.218**	0.301**	-0.347**
<i>Prudish/Promiscuous</i>	0.142*	0.044	1	0.084	0.074	0.078	0.091	0.040	0.034	-0.021
<i>Effeminate/Masculine</i>	0.188**	0.081	0.084	1	-0.153*	0.416**	0.354**	0.070	0.020	0.082
<i>Aloof/Friendly</i>	-0.373**	0.171**	0.074	-0.153*	1	0.172**	-0.069	-0.097	-0.489**	-0.340**
<i>Straight/Gay</i>	-0.075	-0.100	0.078	-0.416**	0.172**	1	-0.196**	-0.100	-0.081	-0.036
<i>Neat/Messy</i>	0.106	-0.139*	0.091	0.354**	-0.069	0.196**	1	0.169**	0.085	0.024
<i>Savvy/Naive</i>	0.217**	-0.218**	0.040	0.070	-0.097	-0.100	0.169**	1	0.221**	0.300**
<i>Kind/Mean</i>	0.592**	-0.301**	0.034	0.020	-0.489**	-0.081	0.085	0.221**	1	0.480**
<i>Genuine/Fake</i>	0.313**	-0.347**	-0.021	0.082	-0.340**	-0.036	0.024	0.300**	0.480**	1

* $p \leq 0.05$ (2-tailed); ** $p \leq 0.01$ (2-tailed)

with messiness and laziness. Recall that negative correlation values imply an inverse relationship, and that the polarity of the affective scales is not consistent throughout, as seen in table 3. So, the fact that the aloof/friendly scale is negatively correlated with the generous/greedy scale means that a listener who rated a speaker as "friendly" was also likely to rate them as "generous," whereas if the listener rated them as "aloof" they were also likely to rate them as "greedy." On the whole, we see that the ideological correlations between personality traits associate high levels of benevolence with high levels of competence. In other words, those characteristics which can be considered positive benevolence traits, i.e., generosity, friendliness, and kindness, are all significantly correlated with those characteristics which can be considered positive competence traits, i.e., hard-working and savvy.

Let us now turn to a more detailed consideration of the straight/gay and effeminate/masculine scales, reproduced in table 8. Masculinity is correlated with greediness, while effeminacy is correlated with generosity. Similarly, masculinity is correlated with aloofness and messiness, while effeminacy is correlated with friendliness and neatness. Those personality traits which presumably are considered positive characteristics, namely generosity, friendliness, and neatness, are all significantly correlated with effeminacy. With respect to the straight/gay scale and its correlated characteristics, gayness is strongly correlated with friendliness and neatness. Straightness, on the other hand, is strongly correlated with aloofness and messiness. Once again, the more positive personality characteristics, friendliness and neatness, cluster together and associate with a particular trait, gayness, while the presumably more negative characteristics, aloofness and messiness, are both significantly correlated with straightness. On the whole, an ideological perception of gay men as both friendly and neat is evidenced by the correlations in the data. Similarly, though none of the 124 listeners rated the speaker of the passage on the "straight" side of the straight/gay scale, the converse ideological perception of straight men as aloof and messy is also evident in the data.

In terms of the relationship between the effeminate/masculine scale and the straight/gay scale, effeminacy is strongly correlated with gayness,

TABLE 8
Correlations for Effeminate/Masculine and Straight/Gay

	<i>Generous/ Greedy</i>	<i>Effeminate/ Masculine</i>	<i>Aloof/ Friendly</i>	<i>Straight/ Gay</i>	<i>Neat/ Messy</i>
Effeminate/Masculine	0.188*	1	-0.153*	-0.416*	0.354*
Straight/Gay	-0.750	-0.416*	0.172*	1	-0.196*

* $p \leq 0.01$ (two-tailed)

while masculinity is strongly correlated with straightness. What this finding indicates is that, for the listeners in the current study, effeminacy in men is associated with gayness, while masculinity in men is associated with straightness. In other words, the listener population conflates the categories of gender and sexuality, and ideologically links the two, whereby normative gender behavior (i.e., masculinity in men) indexes heterosexuality, while nonnormative gender behavior (i.e., effeminacy in men) indexes homosexuality.¹⁵ Moreover, there is an almost total overlap between judgments of effeminacy and its correlated traits and judgments of gayness and its correlated traits. The listeners in the current study seem to view gender (see n. 15) and sexuality as part and parcel of the same identity construct. This finding parallels the results of Gaudio (1994, 48), who reports that “listener ratings on the straight/gay and effeminate/masculine scales were very strongly correlated . . . which means that speakers who were judged as ‘straight’ were also judged as ‘masculine’ and ‘gay’-sounding speakers were also judged as ‘effeminate.’”

These quantitative results provide empirical support for Cameron and Kulick’s (2003, 5) claim that “having a certain kind of body (sex), living as a certain kind of social being (gender), and having certain erotic desires (sexuality)—are not understood or experienced by most people in present day social reality as distinct and separate. Rather they are interconnected.” The interconnection of gender and sexuality has long been at the center of popular understandings of nonnormative sexual identities and practices. Cameron and Kulick, in fact, argue that the conflation of gender deviant behavior with sexuality (e.g., effeminacy in men implies homosexuality) is a product of the compulsory nature of heterosexuality (Rich 1980) in dominant ideologies of gender. Within this dominant ideology, a “real” man is the man who desires, and obtains, women. Therefore, a gay man cannot be a “real” man, and a mutually interdependent relationship between perceptions of gender and perceptions of sexuality arises.

This interdependence does not imply, however, that gender and sexuality are one and the same thing, nor that there are not compelling reasons to keep the two constructs distinct. From a sociolinguistic point of view, some research has found that linguistic features behave differently when acting as cues to gender versus sexuality. Smyth, Jacobs, and Rogers (2003), for example, found that variation in the mean pitch of a speaker had a different effect on listeners’ perceptions of gender than of sexuality. While the listeners in their sample were willing, at times, to rate certain voices as “gay” but not “feminine,” this was done only when the speaker’s mean pitch was low. In other words, listeners were willing to rate a “low-pitch, [otherwise] gay-sounding voice” (344) as “gay” but were unwilling to rate it as “feminine.”

This finding seems to suggest that listeners may judge gender but not sexuality by mean pitch, where feminine voices are ideologically assumed to be higher than masculine ones.

While this conceptualization has certainly been helpful and important within academic understandings of gender and sexuality (as well as within some political ones; cf. Valentine 2003), it remains important to realize that this division is not the lived experience of many people. For some, including the 121 listeners in the current research, a model of gender and sexuality which sees the two as tightly imbricated and interrelated is the way in which they conceptualize their own subjectivity, as well as that of others.

CONCLUSION

This article has presented a new methodological approach to the study of the ideological perception of gender and sexuality. Previous research has encountered both theoretical and analytical problems in using the natural speech of variably identified people as their experimental stimuli. Instead, this research uses controlled microlevel variation of specific linguistic features in an effort to determine what specifically people listen to when judging a speaker's sexuality. The results presented here have shown that, at least for the speaker considered, varying his pitch range and/or his sibilant durations was insufficient to alter listeners' perceptions of his sexuality to a level of statistical significance. Again, this is not to say that pitch range or sibilant duration play no part in judgments of sexuality. Further controlled research of this kind is needed to continue investigating potential symbolic indexes that may interact in the affective judgment of sexual identity. The significant correlations between the different attitudinal scales illustrated the ways in which perceptions of sexuality are ideologically linked to other perceptions of personality and personhood; similarly, by further investigating other linguistic features and their potential cross-correlations and interactions, we can begin to develop a more nuanced understanding of the relationship between ideology and practice and of what is ultimately behind the sociolinguistic perception of sexuality.

APPENDIX A

Rainbow Passage

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently

beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries people have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Others have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbows. Many complicated ideas about the rainbow have been formed. The difference in the rainbow depends considerably upon the size of the drops, and the width of the colored band increases as the size of the drops increase. The actual primary rainbow observed is said to be the effect of super-imposition of a number of bows. If the red of the second bow falls upon the green of the first, the result is to give a bow with an abnormally wide yellow band, since red and green light when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green or blue.

APPENDIX B

Stimulus Text

I was going down the steps to the Six train. It was right around five-thirty, it was rush hour, and the platform was really crowded—almost impossible to move. There was a guy sitting on the ground playing classical music on a keyboard, and another guy further down the platform playing drums. I wanted to get to the front end of the station, so that I could try and get a seat on the train, instead of having to stand. I'm walking along the edge of the platform, you know on those yellow bumpy things they have for blind people, and I glanced up and this woman is walking towards me from the other direction. She was playing with her phone, I don't know like choosing a new ring or something, and she wasn't looking where she was going. I assumed she knew I was there, but I tried to move out of her way. I guess I didn't move fast enough or far enough because I accidentally bumped into her. It wasn't very hard, but maybe since she wasn't paying attention it surprised her, you know? Anyway, she lost her balance and started to like teeter back and forth, almost like in a cartoon, and it looked like she was going to fall onto the subway tracks. I guess by instinct she kind of yelped and flung her arm out, missing my eye with her bracelet by like an inch. She caught hold of the collar of my jacket, and almost pulled me down with her. Luckily, I just fell on the platform, and managed to grab her waist and keep her from falling onto the track. She did drop her cell phone into the tracks—I guess that's the lesson—don't play with your phone when you're walking on the edge of the platform.

NOTES

I would like to thank the following people for their assistance and support: Victoria Bergvall, Renée Blake, Lisa Davidson, Penny Eckert, Rudy Gaudio, Keith Walters, and three anonymous reviewers for *American Speech*. Earlier versions of this paper were presented at the International Gender and Language Association (IGALA) 3 at Cornell University in June of 2004 and at New Ways of Analyzing Variation (NWAV) 33 at the University of Michigan in October 2004; I thank the audiences at these meetings for their questions and comments. All errors are, of course, my own.

1. Gaudio (1994) does report some significant findings with respect to pitch dynamism in his sample.
2. In terms of measuring pitch range, Gaudio reports results both for gross pitch range (i.e., absolute maximum F_0 values) and for restricted pitch range (i.e., the 95% mid-range of those values; Jassem 1971). Neither measurement of pitch range yielded a significant correlation.
3. Rogers and Smyth (2003) do not state explicitly how they measured pitch range, but rather claim that they measured mean F_0 , minimum F_0 , maximum F_0 , and F_0 range.
4. Rogers and Smyth (2003) use Pearson’s product moment correlation tests throughout their experiments. For details on this statistical procedure, see note 14.
5. One reviewer questioned the neutrality of the stimulus passage (see appendix). Since all of the listeners identified the speaker as male, there is the possibility that the content of the passage may have influenced their interpretation of his sexuality, irrespective of the phonetic content of the stimulus. My only response for this potential problem is that the semantic content of the passage was controlled throughout, hopefully controlling for any individual listener biases in the data set.
6. An investigation of the voiced palato-alveolar fricative /ʒ/, as well as both the voiced and voiceless palato-alveolar affricates /tʃ/ and /dʒ/, respectively, was excluded primarily because they have not featured prominently in previous research. The stimulus passage (see appendix) featured two tokens of /dʒ/ and one token of /tʃ/.
7. Obviously, this process can be undertaken in the opposite direction, beginning with a sample judged as “straight” and “masculine” and enlarging the pitch range and sibilant durations. The method employed here was chosen in order to better ensure the naturalness of the digital alterations. I am, however, currently conducting research that employs the second method in order to compare the two.
8. To shorten the sibilants, 17% of the sibilant was spliced out of the signal. The portion from which the splice was taken was always in the center of the segment, where frication amplitude was highest in order to preserve naturalness as much as possible.

9. The original sample and the digitally altered samples can be accessed through my Web site at <http://homepages.nyu.edu/~eml246>.
10. To my knowledge, no listener detected that the passages were digitally modified. However, in the debriefing period that followed each session, some listeners stated that they recognized the two stimuli as the same person, while others thought they were different people. This variability in judgments may have affected their ratings of the speakers.
11. A reviewer points out the potential danger of listeners' attention being drawn artificially to the two features that are experimentally controlled at the expense of other relevant phonetic features. I note this possibility, and see it as a drawback endemic to controlled perception research.
12. Note that quantitative analysis across all six groups is impossible because of the empirical design, which is actually a blend of a between-subjects and a within-subjects test. Were a full-factorial design used with 24 subject groups, then analysis across all groups would have been possible.
13. I credit Niko Besnier (pers. comm., June 6, 2004) for first describing this possibility to me in these terms.
14. The Pearson's product moment correlation coefficient (represented as r) is only one of many correlation statistics that can be derived. The Pearson's coefficient is used to determine how well a linear model describes the relationship between two variables. It is defined as the sum of the products of the standard scores of the two measures divided by the degrees of freedom (Spence et al. 1968).
15. I am describing ratings along the effeminate/masculine scale to be ratings of "gender." I use the term *gender* here to refer to the social (including linguistic) performance of a certain type of sex classification. All listeners identified the speaker as a man; their ratings on the effeminate/masculine scale therefore reflect their perception of his performance of this identity category, hence, what I am here calling "gender."

REFERENCES

- Bell, Allan. 1984. "Language Style as Audience Design." *Language in Society* 13: 145-204.
- . 2001. "Back in Style: Reworking Audience Design." In *Style and Sociolinguistic Variation*, ed. Penelope Eckert and John R. Rickford, 139-69. Cambridge: Cambridge Univ. Press.
- Bucholtz, Mary, and Kira Hall. 2004. "Theorizing Identity in Language and Sexuality Research." *Language in Society* 33: 469-515.
- Cameron, Deborah, and Don Kulick. 2003. *Language and Sexuality*. London: Cambridge Univ. Press.
- Derrida, Jacques. 1995. *Limited Inc.* Evanston, Ill.: Northwestern Univ. Press.

- Eckert, Penelope. 2002. "Demystifying Sexuality and Desire." In *Language and Sexuality: Contesting Meaning in Theory and Practice*, ed. Kathryn Campbell-Kibler, Robert Podesva, Sarah Roberts, and Andrew Wong, 99–110. Stanford, Calif.: Center for the Study of Language and Information, Stanford Univ.
- Fry, Dennis B. 1958. "Experiments in the Perception of Stress." *Language and Speech* 1: 126–52.
- Gaudio, Rudolf P. 1994. "Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men." *American Speech* 69: 30–57.
- Giles, Howard, and Peter F. Powesland. 1975. "A Social Psychological Model of Speech Diversity." In *Speech Style and Social Evaluation*, ed. Howard Giles and Peter Powesland, 154–70. New York: Harcourt Brace.
- Henton, Caroline G. 1989. "Fact and Fiction in the Description of Female and Male Pitch." *Language and Communication* 9: 299–311.
- . 1995. "Pitch Dynamism in Female and Male Speech." *Language and Communication* 15: 43–61.
- Irvine, Judith. 2001. "'Style' as Distinctiveness: The Culture and Ideology of Linguistic Differentiation." In *Style and Sociolinguistic Variation*, ed. Penelope Eckert and John R. Rickford, 21–43. Cambridge: Cambridge Univ. Press.
- Jassem, Wiktor. 1971. "On the Pitch and Compass of the Speaking Voice." *Journal of Phonetics* 1: 59–68.
- Klatt, Dennis H. 1974. "The Duration of [f] in English words." *Journal of Speech and Hearing Research* 17: 51–63.
- . 1975. "Vowel Lengthening Is Syntactically Determined in a Connected Discourse." *Journal of Phonetics* 3: 129–40.
- Kulick, Don. 2000. "Gay and Lesbian Language." *Annual Review of Anthropology* 29: 243–85.
- Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner, and Samuel Fillenbaum. 1960. "Evaluational Reactions to Spoken Languages." *Journal of Abnormal and Social Psychology* 60: 44–51.
- Levon, Erez. 2004. "Examining a Gay Prosody: Issues in Theory, Methodology, and Identity." Paper presented at Lavender Languages and Linguistics XI, Washington, D.C., Feb. 13–15.
- Lehiste, Ilse. 1975. "Some Factors Affecting the Duration of Syllabic Nuclei in English." In *Akten der 1: Salzburger Frühlingstagung für Linguistik, Salzburg vom 24. bis 25. Mai 1974*, ed. Gaberell Drachman, 81–104. Tübingen: Narr.
- Linville, Sue Ellen. 1998. "Acoustic Correlates of Perceived versus Actual Sexual Orientation in Men's Speech." *Folia Phoniatrica et Logopaedica* 50: 35–48.
- Nooteboom, Sieb. 1997. "The Prosody of Speech: Melody and Rhythm." In *The Handbook of Phonetic Sciences*, ed. William J. Hardcastle and John Laver, 640–73. Oxford: Blackwell.
- Oller, D. Kimbrough. 1973. "The Effect of Position in Utterance on Speech Segment Duration in English." *Journal of the Acoustical Society of America* 54: 1235–47.
- Podesva, Robert. 2003. "The Stylistic Use of Phonation Type: Falsetto, Fundamental Frequency, and the Linguistic Construction of Personae." Unpublished MS.

- Podesva, Robert J., Sarah J. Roberts, and Kathryn Campbell-Kibler. 2002. "Sharing Resources and Indexing Meaning in the Production of Gay Styles." In *Language and Sexuality: Contesting Meaning in Theory and Practice*, ed. Kathryn Campbell-Kibler, Robert Podesva, Sarah Roberts, and Andrew Wong, 175–90. Stanford, Calif.: Center for the Study of Language and Information, Stanford Univ.
- Rich, Adrienne. 1980. "Compulsory Heterosexuality and Lesbian Existence." *Signs: Journal of Women in Culture and Society* 5: 631–60.
- Rogers, Henry, and Ron Smyth. 2003. "Phonetic Differences between Gay- and Straight-Sounding Male Speakers of North American English." In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 3–9 August 2003*, 1855–58. Barcelona: Universitat Autònoma de Barcelona.
- Rogers, Henry, Ron Smyth, and Greg Jacobs. 2000. "Vowel and Sibilant Duration in Gay- and Straight-Sounding Male Speech." Paper presented at First International Gender and Language Association Conference (IGALA 1), Stanford, Calif., May 5–7.
- Scherer, Klaus R. 1972. "Judging Personality from Voice: A Cross-Cultural Approach to an Old Issue in Interpersonal Perception." *Journal of Personality* 40: 191–210.
- Smyth, Ron, Greg Jacobs, and Henry Rogers. 2003. "Male Voices and Perceived Sexual Orientation: An Experimental and Theoretical Approach." *Language in Society* 32: 329–50.
- Spence, Janet T., Benton J. Underwood, Carl P. Duncan, and John W. Cotton. 1968. *Elementary Statistics*. 2nd ed. New York: Appleton-Century-Crofts.
- Umeda, Noriko. 1977. "Consonant Duration in American English." *Journal of the Acoustical Society of America* 61: 846–58.
- Valentine, David. 2003. "'I Went to Bed with My Own Kind Once': The Erasure of Desire in the Name of Identity." *Language and Communication* 23: 123–38.