

Sexuality in context: Variation and the sociolinguistic perception of identity

EREZ LEVON

*Department of Linguistics
New York University
26 Broadway, 7th Floor
New York, NY 10003
erez.levon@nyu.edu*

ABSTRACT

This article illustrates the use of an empirical method for examining the perceptual identification of gayness in male speakers. It demonstrates how, by digitally manipulating the speech of isolated individuals, it is possible to obtain reliable evidence that pitch range and sibilant duration may act as indexical of a gay male identity. Further scrutiny of this result, however, illustrates that linguistic indexicality is not as straightforward as it originally appears. Subsequent analyses of the data highlight the ways in which the perceptual evaluation of sexuality is a highly contingent process, dependent upon a variety of sociolinguistic factors. An envelope of variation in listeners' affective judgments of a speaker is shown to exist, and it is argued that research on the perception of identity must go beyond identification of salient features, and also consider when and why these features are not salient. (Prosody, perception, sexuality and gender, contextualization)¹

INTRODUCTION

There is a common intuition among native speakers of American English that it is possible to identify the sexuality of individuals simply by listening to their speech. Consequently, numerous sociolinguists, linguistic anthropologists, and speech pathologists over the past 30 years have attempted to quantify this intuition, and, in so doing, to identify precisely what aspects of a speech signal listeners attune to when making judgments of a speaker's sexuality (see Jacobs 1996 and Kulick 2000 for full reviews). Overall, this body of research has been relatively ineffective at conclusively identifying the particular sociolinguistic variables that both speakers and listeners may stereotypically associate with gay sexuality. In this article, I attempt to isolate particular linguistic features and then examine them within a certain type of empirical context (Levon 2006) in order to see whether these features can be said reliably to index the sexuality of different speakers in different speech environments.

I focus on two prosodic variables, PITCH RANGE and SIBILANT DURATION, both of which have been widely discussed in the literature. Because of the pop-

ular perception that gay men's speech is characterized by high levels of pitch variability, numerous studies have attempted to correlate pitch range with the identification of a gay sexuality (e.g., Gaudio 1994, Rogers & Smyth 2003). None of these, however, has been able to establish a statistically significant link between variation in terms of pitch range and the perception of sexuality. Research on sibilant duration has been somewhat more successful in this regard. Proceeding from the popular stereotype of a "gay man's lisp," several studies have been able to demonstrate concrete evidence of a correlation between the acoustic properties of sibilants in an individual's speech and the perception of that speaker's sexuality (e.g., Linville 1998, Rogers, Smyth & Jacobs 2000).

This article reexamines these two variables using an empirical methodology that is designed to impose more stringent experimental conditions and thus allow for a more reliable linguistic analysis, while simultaneously responding to the recent critical discussion regarding research on language and sexuality (Bucholtz & Hall 2004, Cameron & Kulick 2003, Eckert 2002). I begin, below, by summarizing some recent studies that have examined pitch range and sibilant duration, and the roles that these variables may play in the perception of sexuality. I then outline the empirical methodology, which uses the digitally manipulated speech of single individuals to obtain the experimental stimuli (as developed in Levon 2006). Finally, I present the results of the experiments run using this methodology, and discuss the implications of these results for future research in this area.

PREVIOUS RESEARCH

Pitch range

Both Gaudio 1994 and Rogers & Smyth 2003 examined the effects of pitch range on the perception of the sexuality of a speaker. In Gaudio's study, 13 raters listened to 16 segments of talk from eight speakers (four who self-identified as gay [homosexual] men and four who self-identified as straight [heterosexual] men). After hearing these samples, the 13 listeners rated each speaker on affective scales related to the speaker's sexuality, gender, and personality. Gaudio found that his 13 listeners were able to guess accurately the sexuality of the speakers seven out of eight times. Gaudio then proceeded to analyze the speech signals of the eight speakers, in an attempt to identify any prosodic differentiation between those who self-identified as gay and those who self-identified as straight that may have served as a salient cue to the listener population. No significant difference with respect to pitch range was found, and Gaudio concludes that while pitch range may play a role in the identification of the sexuality of a speaker, it does not do so in isolation and can be understood only in relation to the entire speech signal and context of talk.²

Rogers & Smyth 2003 conducted a series of experiments to examine whether differences in terms of mean pitch (F0) and pitch range affected a listener

population's perceptions of the sexuality of 25 speaker-subjects. These speaker-subjects were all men, 17 of whom self-identified as gay and eight as straight. In the first experiment, the listener population (46 people) heard each of these 25 read a portion of the Rainbow Passage,³ 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 speakers' lives. In this first experiment, Rogers & Smyth found no significant correlation between the pitch ranges of the speakers and the listeners' ratings of the speakers' sexuality. In a second experiment, Rogers & Smyth used only the readings of the Rainbow Passage. From these recordings, all segmental and voice quality information was removed. These manipulated recordings were then presented to a second population (44 listeners), who heard what sounded like a schwa vowel pronounced with varying pitch. Once again, Rogers & Smyth found no significant correlation between the pitch ranges of these manipulated recordings and the listeners' perceptions of sexuality. Finally, in a third experiment, Rogers & Smyth presented the manipulated recordings to a group of nine linguists, who were asked to judge the relative pitch variability of the utterance. When these results, relating to the perception of variability, were compared with the results of the earlier groups, relating to the perception of sexuality, a significant correlation was obtained. Rogers & Smyth conclude that while there may be no real acoustic correlation between the pitch range of a speaker and listeners' judgments of his sexuality, listeners have an expectation that there will be one, and so may base their judgments on this expectation.

Sibilant duration

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

Linville 1998 also examined sibilant properties as a potential cue to the identification of the sexuality of a speaker. In her study, 25 listeners rated the speech of nine speakers (five self-identified as gay, four self-identified as straight). The listeners accurately guessed the self-identified sexuality of the speakers 79.6% of the time. Multiple regression analyses of the perceptual results showed that correct identification of actual sexual orientation could be significantly correlated with the properties of the voiceless alveolar fricative /s/. The gay speakers in the study were shown to have longer /s/ frication and well as higher /s/ peak frequency values. While there is some overlap in the data with respect to peak

frequency and length between the gay and straight speakers, Beta weights in the statistical analysis indicated that /s/ duration and /s/ peak frequency made a statistically significant contribution to the identification of the sexuality of the speaker.

Discussion

The four studies summarized above all employed essentially the same methodology. In each, a speaker population was recorded reading a short passage. This passage was then presented to a listener population who were asked, without being given any other information, to attempt to identify the sexuality of the speakers. In all four of the studies, listeners were shown to be remarkably adept at accurately identifying the sexuality of the speaker-subjects. Following this listening task, the researchers then analyzed the speech of those speakers who self-identified as “gay” and those who self-identified as “straight,” in an attempt to discover any systematic differences between the two groups that could help explain the ability of listeners to deduce the speakers’ sexualities correctly.

I argue that this type of methodology encounters two difficulties, one theoretical and one empirical. I have already extensively discussed the theoretical critique elsewhere (Levon 2006), and so I will not attend to it here. Suffice it to say that, from a theoretical perspective, this body of previous research has in some ways ignored the fact that linguistic indexicality is a semiotic tool that gets variably implemented by speakers across contexts. Instead, the methodology employed has implicitly assumed gay-identified speech to be an essential component of gay people, and then set out to quantify it. While I agree with the assertion that certain identity categories, however artificial, can act as useful and salient locations around which people position their own construction of a social persona (Eckert 2002), I contend that an empirical methodology should focus more on the semiotic practices people use to enact this social positioning (Cameron & Kulick 2003, Bucholtz & Hall 2004), rather than focus on those positions themselves. To do so, what is needed is a methodology that examines the perception of a gay identity – that is, what semiotic resources people associate with the linguistic performance of gayness, as opposed to a methodology that attempts to classify certain linguistic features as components of a “gay speech” (cf. Boellstorff & Leap 2004). Here, I attempt to address this issue by using an alternative methodology for the study of gay-indexed language that builds on the insights of previous research and examines them in a more empirically conclusive way.

In doing so, the current research also avoids an analytical shortcoming of the previous work described above. Recall the schematic of the methodology used in these studies, in which listeners first rate the speech of differentially self-identified individuals, and then the researchers attempt to tease out differences between the speech samples. In this scenario, because listeners are presented with the speech of entirely different individuals, there are potentially hundreds of linguistic differences among the samples. Exhaustively and conclusively iden-

tifying which precise difference (or differences) listeners are attuning to when making judgments of sexuality is practically impossible. Instead, I make use of newer developments in acoustic technology to create an empirical environment in which variation among the samples is controlled and limited to the particular variable in question. In this way, any change in perceptual reaction to the samples can be reliably identified with the specific linguistic variable(s) under consideration.

EXPERIMENTAL DESIGN

The experimental design of the current research is an adaptation and expansion of the methodology originally developed in Levon 2006. Let me begin by describing the methods there, before going on to discuss the ways in which the current project makes use of and eventually extends those methods to suit our present needs. In Levon 2006, the speech of a single individual was used as the basis from which to derive empirical stimuli for a perceptual study of male sexuality. Doing this allowed me to avoid comparing representative “gay” and “straight” speakers, while also allowing me to control the linguistic variation among the stimuli.

Stimuli

A speaker, a white male in his mid-twenties otherwise unrelated to the project, was recording reading a short passage, approximately 71 seconds in length (see Appendix). This passage was a neutral narrative about a typical New York City occurrence (a crowded subway platform during rush hour in Manhattan).⁴ The passage 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/.⁵

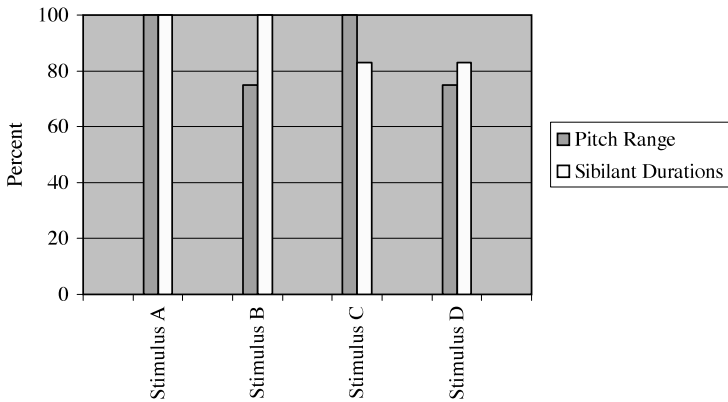
The speaker’s recorded reading of this passage was presented to a pre-test group of 10 listeners (all graduate students in Linguistics at New York University). This pre-test group rated the speaker of the original recording on the scales Straight/Gay and Effeminate/Masculine.⁶ All 10 of these pre-test listeners rated the speaker of the original recording as both “extremely gay” and “extremely effeminate.” The primary goal of that research was to evaluate the ways in which controlled variation of linguistic content may affect listeners’ perceptions of the sexuality of the speaker. Therefore, the original recording, judged as both “extremely gay” and “extremely effeminate,” was digitally altered with respect to the two variables under consideration, sibilant duration and pitch range. By changing only this variable content to create the experimental stimuli, I could determine whether these two linguistic features affect the subjective judgments of the speaker’s sexuality. In other words, if changing the pitch range, and only the pitch range, of a speech sample, led listeners to rate a speaker previously judged as “gay” or “effeminate” as “straight” or “masculine,” I could have safely concluded that pitch range was responsible for that change in listener evaluation.

Building on the ideological assumption that gay men have wider pitch ranges and longer sibilant durations than straight men, the digital manipulations of the original recording shortened the sibilants and reduced the pitch ranges across the board. Note that while I was not focusing on identity categories as primitives, I used the ideological assumptions regarding these categories to motivate the experimental testing. In other words, I was not testing what “gay” and “straight” speakers actually do with language, but rather what a listener population may THINK they do with language. For this reason, I used the stereotypical assumptions characterizing gay men’s speech as having wide pitch ranges and long sibilant durations to investigate the PERCEPTION of sexuality (cf. Bucholtz & Hall 2004).

Sibilants were shortened by 17%.⁷ This is the figure that Rogers, Smyth & Jacobs 2000 and Linville 1998 report as the threshold at which listeners’ perceptions of the sexuality of the speaker changed. By shortening the sibilants in this way, I was able to verify whether previous findings of significance with respect to sibilant duration would hold when everything else was controlled for. I 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. Much previous research has shown that segment length is highly dependent upon such things as word position (Klatt 1974, Lehiste 1975, Oller 1973, Umeda 1977), syllable position (Oller 1973), phrase position (Klatt 1975) and syllable stress (Fry 1958, Lehiste 1975, Umeda 1977). By using a relative measure to alter the sibilant durations, I was able to preserve the internal phonetic distinctions within the speech sample.

With regard to pitch range, there does not exist as clear-cut a benchmark in the literature differentiating speech that is judged as “gay” from speech judged as “straight.” Some researchers have reported differences as large as 45%, and others differences as small as 8%. In previous research I conducted on the representations of gay characters in television and film (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 ranges of the original sample by 25%. Pitch manipulations were done using the Pitch Manipulation Editor in Praat 4.1.15. The original sample was first coded for intonational phrasing using the Tone Break Indices (ToBI) method. Within each intonational phrase, the restricted pitch range (95.4% of the gross pitch range) was calculated in semitones.⁸ Once the pitch range of each intonational phrase was identified, I flattened the overall contour of each phase by 25%. This preserved the original shape of the pitch contour, as well as the relative rates of pitch dynamism between the peaks and valleys. In this way, pitch range was isolated as the variable under examination.

From these digital manipulations, four experimental stimuli were created, as shown in Figure 1. Stimulus A refers to the speaker’s original recording, with 100% of pitch ranges and sibilant durations intact. Stimulus B refers to that recording in which sibilant durations were kept constant, while pitch ranges were reduced by



- A: [+wide] pitch range, [+long] sibilant duration (*original*)
 B: [−wide] pitch range, [+long] sibilant duration
 C: [+wide] pitch range, [−long] sibilant duration
 D: [−wide] pitch range, [−long] sibilant duration

FIGURE 1: Derived stimuli.

25%. In stimulus C, pitch ranges were kept constant and sibilant durations were reduced by 17%. Finally, in stimulus D, both pitch ranges and sibilant durations were reduced by 25% and 17%, respectively. These four stimuli thus represent an exhaustive typology of the two variables, when considered binarily.⁹

These four stimuli were then presented to a listener population who rated the four stimuli on the ten affective scales shown in Table 1. These scales (adapted from Scherer 1972 and Gaudio 1994) are designed to gauge listeners' ideological perceptions of various personality characteristics of the speakers, as well as traits related to the speakers' gender and sexual identifications.¹⁰

In the interest of space, I will not discuss at length the results of this prior experiment (see Levon 2006). In brief, the results of that study were inconclusive as to the effect of either pitch range or sibilant duration on the affective judgments of the speaker's sexuality, insofar as no statistically significant results were obtained. Controlled variation of neither pitch range nor sibilant duration was shown to have a significant effect on listeners' perceptions across all 10 affective categories. The current research is a second run of this previous experiment that offers two additional innovations: (i) the variables were manipulated and examined in both directions (i.e., by both shortening sibilants and narrowing pitch ranges of a "gay"-identified voice, as well as lengthening sibilants and widening pitch ranges of a "straight"-identified voice); and (ii) the ways in which the stimuli are presented to subjects were changed to allow for a

TABLE 1. *Affective scales.*

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

(S) – adapted from Scherer (1972); (G) adapted from Gaudio (1994)

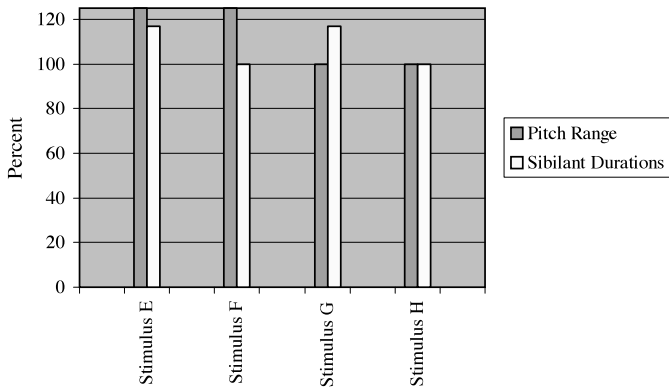
fully between-subjects design and an examination of order of presentation as a potential factor.

The four experimental stimuli derived from the “gay”-identified speaker in Levon 2006 are used again here. In addition to this, and in order to examine more fully the intricacies of the potential indexical properties of pitch range and sibilant duration, four new experimental stimuli were created from the speech of a “straight”-identified speaker. A second speaker, a white male in his mid-thirties otherwise unrelated to the project, was recorded reading the same short passage as in the first experiment (see Appendix). This speaker’s original recording was then played for the same pre-test group of linguistics graduate students as before, who this time all labeled the speaker as “straight” and “masculine.” Digital manipulations of this new recording were undertaken in a process identical to the one employed previously, except that this time pitch ranges were widened by 25% and sibilant durations were lengthened by 17%.¹¹ This process resulted in a new exhaustive typology of the two variables, as presented in Figure 2. In this new set of experimental stimuli, stimulus H represents the speaker’s original recording with 100% of pitch ranges and sibilant durations intact. Stimulus G is the one in which pitch range has been kept constant and sibilant durations have been lengthened by 17% across the board, while stimulus F is the one in which sibilant durations have remained constant and pitch ranges have been widened by 25% across the board. Finally, stimulus E is the sample in which both sibilant durations and pitch ranges were lengthened by 17% and widened by 25%, respectively.

Participants

These two sets of experimental stimuli, one “gay”-derived and one “straight”-derived, were then presented to a listener population of 123 undergraduate stu-

SEXUALITY IN CONTEXT



- E: [+wide] pitch range, [+long] sibilant duration
 F: [+wide] pitch range, [-long] sibilant duration
 G: [-wide] pitch range, [+long] sibilant duration
 H: [-wide] pitch range, [-long] sibilant duration (*original*)

FIGURE 2: Second set of derived stimuli.

dents in New York City. These listeners were all students at either New York University or the City University of New York, enrolled in women's studies, linguistics, or anthropology courses.¹² None of these listeners knew me before the experiment nor had any knowledge of the research project. The listener population varied in age from 18 to 45, and was made up of approximately 65% women and 35% men. After having heard each of the experimental stimuli, listeners were asked to rate the speakers of each passage on the 10 affective scales shown in Table 1 above.

Methodology

This listener population was divided into eight groups. Each of these groups heard four recordings of the stimulus text, as diagrammed in Figure 2, including one of the "gay"-derived stimuli (A, B, C or D), one of the "straight"-derived stimuli (E, F, G or H), and two decoy recordings (X and Y). The decoy recordings, which were kept constant throughout, were samples of the same passage, this time read by two additional people, also otherwise unrelated to the project. These decoy recordings were judged by the pre-test group to sound both "straight" and "masculine," and were used to keep the target recordings non-adjacent.

Although subjects were divided into eight groups to actually run the experiment, the analysis of the data considers them as only four; the pairs of groups who heard the same stimuli but in different orders are considered together, and order of presentation is examined. Take, for example, groups 1 and 8. Both of

TABLE 2. *Presentation of stimuli by order.*

Group 1	X, A, Y, H	
Group 2	X, B, Y, G	
Group 3	X, C, Y, F	
Group 4	X, D, Y, E	
Group 5	X, E, Y, D	
Group 6	X, F, Y, C	
Group 7	X, G, Y, B	
Group 8	X, H, Y, A	

these groups heard stimuli A, X, Y and H. In group 1, A was heard in the second position and H was heard in the fourth, while in group 8, H was heard in the second position and A was heard in the fourth. Since groups 1 and 8 heard exactly the same stimuli, they are analyzed together, and the fact that they heard them in different orders allows us also to test order of presentation as a potentially significant factor in the affective judgment of the speakers.

This is a between-subjects empirical design, whereby the ratings of the stimuli between the groups are compared. This type of statistical analysis is not necessarily as rigorous as a within-subjects test, where each rater hears all of the passages, but is an unavoidable drawback of this methodology. Because the differences between each of the target stimuli are restricted to a single variable – pitch range or sibilant duration – all of the target passages that are derived from the same person are obviously spoken by the same speaker. Asking listeners, then, to rate the speakers of the different passages when they are obviously read by the same person is somewhat counterintuitive. For this reason, each listener heard only one passage from each speaker, and ratings were compared across the four groups. Although this may reduce the statistical power of the analysis, it is an unavoidable artifact of the current methodology.

RESULTS

Since the primary goal of the current research is to assess the extent to which controlled variation of particular features can change listeners' perceptions of the same speaker, rather than comparing gay- and straight-identified speakers with one another, we will consider the listeners' judgments of the "gay"-derived stimuli (A–D) and the "straight"-derived stimuli (E–H) separately.¹³ Moreover, analyzing the two sets of stimuli separately avoids comparing stimuli whose differences from one another may include much more than the target variables in question (cf. discussion of previous research, above). My central goal is to examine the effect, if any, of specific linguistic features on the perception of sexuality. The analysis must therefore consider only those effects obtained across stimuli for the same speaker, as opposed to across speakers. In the interest of

TABLE 3. *Affective judgments of the “gay”-derived speaker by stimulus and order.*

	Stimulus	Order	Listener Rating
Effeminate/Masculine ¹	A	First	2.31
		Second	1.90
		Total	2.15 (<i>extremely effeminate</i>)
	B	First	3.00
		Second	1.73
		Total	2.28 (<i>extremely effeminate</i>)
	C	First	2.46
		Second	2.00
		Total	2.21 (<i>extremely effeminate</i>)
	D	First	4.08
		Second	2.36
		Total	3.19 (<i>effeminate</i>)
	Total	First	2.93 (<i>effeminate</i>)
Second		1.97 (<i>extremely effeminate</i>)	
Gay/Straight ²	A	First	2.59
		Second	1.20
		Total	2.07 (<i>extremely gay</i>)
	B	First	3.71
		Second	2.45
		Total	3.00 (<i>gay</i>)
	C	First	2.23
		Second	2.27
		Total	2.25 (<i>extremely gay</i>)
	D	First	4.00
		Second	2.57
		Total	3.26 (<i>gay</i>)
	Total	First	3.13 (<i>gay</i>)
Second		2.23 (<i>extremely gay</i>)	

¹ Stimulus: $F(3,112) = 4.463$, $p = 0.005$; Order: $F(1,112) = 17.622$, $p = 0.000$; Stimulus * Order: $F(3,112) = 1.833$, $p = 0.145$

² Stimulus: $F(3,113) = 4.180$, $p = 0.008$; Order: $F(1,113) = 9.989$, $p = 0.002$; Stimulus * Order: $F(3,113) = 1.166$, $p = 0.326$

³ For this table, a score between 1–2.5 and 5.5–7 is labeled “extremely” and a score between 3.5–4.5 is labeled “neutral”

space, I discuss only those results related to two of the ten affective scales: Effeminate/Masculine and Straight/Gay.

“Gay”-derived stimuli

Table 3 presents the results of a quantitative analysis of listeners’ affective ratings of the “gay”-derived speakers as a function of stimulus passage and order of presentation. In Table 3, scores represent listeners’ mean ratings on the 7-point

scale, where scores of 1–2.5 and 5.5–7 are taken to represent “extreme” ratings of the adjectives considered, and scores between 3.5 and 4.5 are taken to represent neutral ratings. Also recall that the expected alignment of gayness with effeminacy and straightness with masculinity is reversed (see Table 1 and n. 10). The scores in Table 3 for the Straight/Gay scale represent the figures obtained when the original polarity is reversed and the ideologically anticipated alignment is restored (for ease of comparison).

In Table 3, we see that both stimulus and order of presentation are selected as significant main effects on the Effeminate/Masculine (stimulus: $F = 4.463$, $p = 0.005$; order: $F = 17.622$, $p = 0.000$) and Gay/Straight (stimulus: $F = 4.180$, $p = 0.008$; order: $F = 9.989$, $p = 0.002$) scales. Their interaction, however, is not (stimulus * order: $F = 1.833$, $p = 0.145$ on the Effeminate/Masculine scale; $F = 1.166$, $p = 0.326$ on the Gay/Straight scale). This finding indicates that while the changes in the phonetic content of the stimulus passage and the order in which these passages are heard by listeners both have significant effects on perceptual ratings of the gender/sexuality of the speaker, these two effects can, and should, be considered separately. In other words, statistically, the significant effect of stimulus can be considered independently from the significant effect of order, and vice versa. For our current purposes, let us begin with the stimulus effect.

First, note the similarity between the scores obtained for the Effeminate/Masculine scale and the (now reversed) Gay/Straight scale with respect to stimulus passage (in the rows marked Total for each stimulus A–D). For both, scores range from the low twos (2.15 for the Effeminate/Masculine scale and 2.07 for the Gay/Straight scale) to the low threes (3.19 for the Effeminate/Masculine scale and 3.26 for the Gay/Straight scale). In addition, the ratings for both of these scales are essentially centered on the same space in the scale.¹⁴ This finding seems to indicate that listeners are treating the two scales, Effeminate/Masculine and Straight/Gay, similarly (cf. Levon 2006, Valentine 2003).¹⁵ In fact, correlation analyses of listeners’ ratings on these two scales show them to be significantly correlated ($r = -0.487$; $p \leq 0.001$). Notwithstanding this correlation, however, listeners were shown to treat these scales differently when reacting to variation between the stimuli. I therefore keep the two scales separate, while remaining aware of the fact that doing so is not necessarily statistically motivated.

Although Table 3 shows us that stimulus, in general, is a significant factor, it is unable to pinpoint the precise effect that either pitch range or sibilant duration has on listeners’ perceptions. To get at this information, we must consider the results of a pairwise analysis of all the stimuli, where each stimulus is compared to every other, and the significance of changes in listener ratings is assessed. Table 4 presents these results, where all of the possible combinations of the four stimuli are examined (i.e., A/B, A/C, A/D, B/C, B/D, C/D).

On the Effeminate/Masculine scale, three of the six comparisons proved significant: Stimulus A (wide pitch ranges, long sibilant durations) and Stimulus D

TABLE 4. *Pairwise comparisons of stimuli for “gay”-derived speaker.*

	Stimulus	Significance ¹
Effeminate/Masculine	A × B	0.698
	A × C	0.876
	A × D	0.008
	B × C	0.843
	B × D	0.008
	C × D	0.014
Gay/Straight	A × B	0.038
	A × C	0.704
	A × D	0.018
	B × C	0.093
	B × D	0.578
	C × D	0.043

¹ Two-tailed *t*-test; equal variances not assumed

(narrow pitch ranges, short sibilant durations); Stimulus B (narrow pitch ranges, long sibilant durations) and Stimulus D (narrow pitch ranges, short sibilant durations); and Stimulus C (wide pitch ranges, short sibilant durations) and Stimulus D (narrow pitch ranges, short sibilant durations). The significant differences between mean ratings on the last two of these (e.g., B vs. D and C vs. D), indicates that both shortening sibilant durations (i.e., B, mean score = 2.28, vs. D, mean score = 3.19) and narrowing the pitch ranges (i.e., C, mean score = 2.21, vs. D, mean score = 3.19) have the effect of making listeners rate the speaker as significantly more masculine. When these two are combined (i.e., A vs. D), listeners' ratings of the speaker change the most – from a mean score of 2.15 to a mean score of 3.19. What is crucial here, however, is the fact that we cannot claim, based on the results shown in Table 4, that pitch range or sibilant duration straightforwardly affects listeners' judgments of masculinity in isolation. The comparisons of Stimulus A (wide pitch ranges, long sibilant durations) to Stimulus B (narrow pitch ranges, long sibilant durations) and Stimulus A to Stimulus C (wide pitch ranges, short sibilant durations) do not yield significant results. In other words, narrowing the pitch ranges without shortening the sibilant durations, as in the change from A to B, or shortening the sibilant durations without narrowing the pitch ranges, as in the change from A to C, has no significant effect on listeners' judgments. Only when the two coincide – short sibilant durations with narrow pitch ranges – do listeners rate the speaker as significantly more masculine.

In terms of the Gay/Straight scale, three of the six pairwise comparisons yield a significant result: Stimulus A to Stimulus B, Stimulus A to Stimulus D and Stimulus C to Stimulus D. The fact that these three, and only these three, are

selected as significant points out the operative role that pitch range appears to be playing in the listeners' perceptual judgments of the sexuality of the speaker. In all three of the comparisons, it is the change in pitch range from wide to narrow that affects listener ratings, seemingly irrespective of the sibilant durations. In the change from Stimulus A to Stimulus B, where pitch ranges are narrowed while sibilant durations remain long, listener ratings go from a mean score of 2.07 ("extremely gay") to 3.00 ("gay"). Similarly, in the change from Stimulus C to Stimulus D, where pitch ranges are narrowed while sibilant durations remain short, listener ratings go from a mean score of 2.25 ("extremely gay") to 3.26 ("gay"). And finally, in the change from Stimulus A to Stimulus D, where pitch ranges are narrowed and sibilant durations are shortened, listener mean ratings go from 2.07 ("extremely gay") to 3.26 ("gay"). These results indicate, therefore, that pitch range is acting as indexical of sexuality for the listener population. Sibilant duration, on the other hand, does not appear to affect listener judgments of sexuality at all. Neither the change from Stimulus A to Stimulus C nor that from B to D yields any significant results. Moreover, listeners' ratings are affected by changes in pitch ranges regardless of the status of sibilant durations (i.e., as either long or short). It therefore appears that pitch range alone is the operative linguistic indexical with respect to the Gay/Straight scale.

To summarize what we can claim at this point: Both pitch range and sibilant duration were shown to significantly affect listeners' judgments of gender and/or sexuality. Pitch range was shown to be a salient index of both gender and sexuality, while sibilant duration was found to affect only the former. Moreover, both of these variables were shown to be at times dependent on each other for their indexical felicity: On the Effeminate/Masculine scale, pitch range was effective only when coupled with short sibilant duration, and sibilant duration was effective only when coupled with narrow pitch range. I take this interdependence of the two variables as evidence in support of a gestalt-like understanding of indexicality (Eckert 2000, Levon 2006), whereby linguistic features are not only salient on their own but can also work in clusters to achieve social-indexical significance. On the Gay/Straight scale, however, pitch range does work alone to alter listeners' judgments of the sexuality of the speaker, and thus serves to highlight the fact that linguistic indexicality may operate in various and multiple ways. Yet, regardless of the particular instantiation of linguistic indexicality in a given situation, these results demonstrate how the experimental methodology described here is able to tease out which specific linguistic features, or combinations of features, are affecting listeners' perceptions of a speaker.

Results from the current research also indicate that indexical speech is highly sensitive to context. Recall from Table 3 that order of presentation of the stimulus passage is also selected as a significant main effect. In the experimental design, each target stimulus was heard in either the second of four positions (following one decoy and preceding one decoy and one target stimulus) or in the fourth of four positions (following both decoys and the other target stimu-

lus). When looking at the combined ratings across stimuli in the Total rows at the bottom of Table 3, we see that when heard in the second position, the “gay”-derived speaker receives a mean score of 2.93, or “effeminate,” on the Effeminate/Masculine scale, and a mean score of 3.13, or “gay,” on the Gay/Straight scale. When, however, he is heard in the fourth position, after all the other recordings, he is rated as significantly more “effeminate” (with a mean score of 1.97), as well as significantly more “gay” (with a mean score of 2.23). In rating the stimuli, listeners were therefore also attuned to the context in which these recordings were heard. This fact can be taken to underscore the notion that linguistic indexicality itself is not an isolated phenomenon, whereby there is a direct link between linguistic form and social function. Rather, the data suggest an understanding of indexicality as a mediated phenomenon, wherein linguistic indexicals are inseparable from the context in which they are heard (Ochs 1990, 1992).

In the current example, the relational contrast between target stimuli and the other recordings (i.e., decoy passages and a “straight”-derived passage) was shown to have a significant effect on listeners’ ratings, independent of the manipulation of the variable content of the stimuli. In other words, on the Effeminate/Masculine scale, listeners rated those passages heard in the second position (after all other passages) as significantly more effeminate than those heard in the first position (after one decoy and before the second decoy and the second target stimulus), regardless of the specific phonetic content of that passage. In fact, estimates of effect sizes for stimulus and order on the Effeminate/Masculine scale indicate that order of presentation has a larger effect on listener ratings ($\eta_p^2 = 0.136$) than stimulus does ($\eta_p^2 = 0.107$; confidence interval for both = 95%). Listeners, therefore, appear to be slightly more attuned to relational contrasts across the stimuli when making judgments about gender, as evidenced in the order of presentation effect, than they are to the phonetic content of the target stimuli themselves. On the Gay/Straight scale, the attention paid to order of presentation is somewhat diminished; estimates of effect size indicate that order is actually less important than the phonetic content of the stimuli (for order, $\eta_p^2 = 0.081$; for stimulus, $\eta_p^2 = 0.100$; confidence interval for both = 95%). Yet, for both of the scales, the differences in effect sizes between order and stimulus are relatively small, leading to the conclusion that stimulus and order of presentation affect listener judgments in similar and important ways. The significant effect of order, then, serves as a check on the findings with respect to stimulus, and recalls Cameron & Kulick’s (2003) caution that results from perceptual experimentation, however significant, may not carry the same weight in practice as they do in the controlled empirical setting.¹⁶

“Straight”-derived stimuli

Let us now consider the results for the “straight”-derived stimuli. As seen in Table 5, this set of stimuli yielded no significant results with respect to listeners’

TABLE 5. *Affective judgments of the second speaker by stimulus and order.*

	Stimulus	Order	Listener Rating
Effeminate/Masculine ¹	E	First	3.57
		Second	4.08
		Total	3.81 (<i>neutral</i>)
	F	First	3.67
		Second	4.08
		Total	3.86 (<i>neutral</i>)
	G	First	3.91
		Second	3.88
		Total	3.90 (<i>neutral</i>)
	H	First	3.70
		Second	4.06
		Total	3.92 (<i>neutral</i>)
Total	First	3.74 (<i>neutral</i>)	
	Second	4.02 (<i>neutral</i>)	
Gay/Straight ²	E	First	3.64
		Second	4.00
		Total	3.81 (<i>neutral</i>)
	F	First	3.60
		Second	4.15
		Total	3.86 (<i>neutral</i>)
	G	First	4.45
		Second	3.76
		Total	4.15 (<i>neutral</i>)
	H	First	4.10
		Second	4.35
		Total	4.26 (<i>neutral</i>)
Total	First	4.00 (<i>neutral</i>)	
	Second	4.07 (<i>neutral</i>)	

¹ Stimulus: $F(3,112) = 0.019$, $p = 0.996$; Order: $F(1,112) = 1.924$, $p = 0.168$; Stimulus * Order: $F(3,112) = 0.321$, $p = 0.810$

² Stimulus: $F(3,113) = 0.494$, $p = 0.687$; Order: $F(1,113) = 0.199$, $p = 0.656$; Stimulus * Order: $F(3,113) = 1.297$, $p = 0.279$

³ For this table, a score between 1–2.5 and 5.5–7 is labeled “extremely” and a score between 3.5–4.5 is labeled “neutral”

judgments on either the Effeminate/Masculine scale or Gay/Straight scale for either variable. Whether pitch ranges were wider or narrower, sibilant durations longer or shorter, listeners’ judgments of the gender and sexuality of the “straight”-derived speaker did not change. Nor was order of presentation shown to be significant, as it was for the “gay”-derived speaker, insofar as listeners’ ratings also did not change depending on whether this second speaker was heard in the second or fourth position. This result certainly complicates the assertion made above

TABLE 6. *Comparison of certain voice quality features between speakers.*

Feature	Speaker Mean		F statistic	Significance
	“gay”-derived	“straight”-derived		
F0 floor	106.64 Hz	112.22 Hz	(1, 56) 2.119	0.151
F0 range	13.12 st	8.64 st	(1, 56) 20.64	0.000
Jitter	1.84%	3.63%	(1, 56) 107.59	0.000
Shimmer	10.08%	16.45%	(1, 56) 126.66	0.000
HNR	9.52 dB	8.66 dB	(1, 56) 8.317	0.006

that pitch range and sibilant duration can act, in certain environments, as salient indexicals of gender and sexuality.

Yet this result, or more appropriately the lack thereof, serves to reinforce the notion that indexicality is not a straightforward process by which particular linguistic variables (or clusters of variables) are linked with social positions. Similar to Ochs’s (1990, 1992) caution that indexicality must be viewed as a contextualized phenomenon that is always situated within a particular speech event, recall Eckert’s (2000) reminder, noted above, that indexical features of language rarely operate alone, but rather should be considered in clusters. This very idea is demonstrated with respect to the results of listener ratings of the “gay”-derived stimuli on the Effeminate/Masculine scale, where pitch range and sibilant durations were shown to be effective only when working together. It seems likely that in the case at hand, some linguistic factor or factors differentiating the speech of the two speakers is responsible for the lack of saliency with respect to pitch range and sibilant duration among the “straight”-derived stimuli. In other words, perhaps there is some characteristic feature in the speech of the “straight”-identified speaker, absent in the speech of the “gay”-identified one, that overrides any effect that may be caused by varying pitch ranges or sibilant durations.

Comparison of “gay”-identified and “straight”-identified speakers

Conclusively identifying what these other linguistic factors may be goes beyond the scope of the current research, but let me enumerate certain differences between the speech of the two speakers used here to highlight possible avenues for future work. Table 6 presents a sampling of voice quality features and their mean values for each speaker. Of the five voice quality characteristics considered – F0 floor, F0 range, jitter, shimmer, and Harmonics-to-Noise Ratio (HNR) – only mean F0 floor values are not significantly different between the two speakers. F0 range, though not shown to have a significant effect on listeners’ judgments of the stimuli derived from the “straight” speaker, is significantly narrower for him than for the other speaker (for whom variation of pitch range did have a significant effect). Perhaps this fact indicates that altering the “straight” speaker’s

pitch ranges by 25% was not enough, and had they been widened further, significant results might have been obtained.

In terms of the other voice quality measurements, mean levels of jitter, shimmer, and HNR are all significantly different between the two speakers. Jitter, which here measures the average absolute difference between periods in a waveform expressed as a percentage of the average period, was shown to be lower for the “gay”-derived speaker than the “straight”-derived speaker. Johnstone & Scherer 1999 have shown lower levels of jitter to be associated with the affective perception of “tension” or “stress” on the part of the speaker. It seems possible that this feature, then, could have an indirect effect on listeners’ perceptions of gender or sexuality (Ochs 1992), dependent upon a potential correlation between personality traits such as “tense” and gender or sexual identity categories. Shimmer levels, which here measure the average absolute differences between amplitudes of consecutive periods in a waveform expressed as a percentage of the average amplitude, were also shown to be significantly lower for the “gay” speaker than for the “straight”. While I know of no empirical study that has examined the affective reactions to different shimmer levels, it seems equally possible that shimmer like jitter, may play an indirect role in the indexation of gender and/or sexuality. Finally, measurements of the Harmonics-to-Noise Ratios (HNRs), which describe the distribution of energy within a sound signal, show HNR to be significantly higher for the “gay” speaker than for the “straight”. In general, lower HNR levels correspond with the perception of hoarseness in the voice, and it seems plausible that this type of psychoacoustic characteristic may be linked to relative perceptions of masculinity vs. femininity and/or sexuality (cf. Podesva n.d.).

In addition to the features mentioned above, there come to mind a host of other features that previous research has indicated may affect listeners’ perceptual evaluations of speakers, including sibilant quality (Linville 1998, Rogers, Smyth & Jacobs 2000), spectral slope (Fox & Nissen 2002), voice type (i.e., modal, falsetto, or creaky; Podesva n.d.), and vowel formants (Avery & Liss 1996). Yet whatever variable is to be tested, be it one of those mentioned in Table 6 above or any other, the central focus of the current study has been to illustrate the necessity of examining these variables in such a way as to minimize both the theoretical and empirical shortcomings of earlier research. I argue that only by isolating specific variables and presenting them in exhaustive typologies to listeners, as in the experiment described here, can we hope to determine reliably any function related to the indexation of gender and/or sexuality that a linguistic feature may possess.

CONCLUSION

This article assesses the effect of digitally manipulating the sibilant durations and pitch ranges of two speakers, one identified as sounding “straight” and one

identified as sounding “gay,” on listeners’ affective judgments of their sexuality. For the “gay”-derived stimuli, reliable empirical evidence was obtained that pitch range and sibilant duration can act as indices of sexuality and/or gender either alone (in the case of sexuality) or together (in the case of gender). The ability of features or clusters of features to act indexically was then complicated by the significant effect of order of presentation on listener judgments. The fact that listeners judged the speaker, regardless of the phonetic content of the sample, as more “gay” and more “effeminate” when hearing him after multiple other speakers whom they judged as “masculine” and “straight” emphasizes the need to consider linguistic indexicality broadly as a highly contextualized phenomenon (Ochs 1990, 1992).

The contingent nature of indexicality was further underscored by the lack of significant effect of either pitch range or sibilant duration on listeners’ evaluations of the “straight”-derived stimuli. It was proposed that this lack of effect may be due to linguistic elements in the speech of the original speaker that render ineffective any variation in terms of pitch range or sibilant duration. While certain linguistic features are listed as potential factors in this differentiation, it is stressed that subsequent controlled research, of the kind described here, is required to determine which, if any, of these differences is perceptually salient for listeners. Future research is encouraged that would examine these other features, as well as address certain shortcomings of the current study, such as the non-random and relatively small nature of the sample size, the potential semantic bias of the stimulus text, and the possibly problematic use of the Straight/Gay and Effeminate/Masculine affective scales. In doing so, this work would be able to continue the investigation of linguistic indexicality as a complex process involving a constellation of sociolinguistic factors, thus enabling a more comprehensive understanding of the linguistic perception of identity.

APPENDIX

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

¹ I greatly benefited from the insight and assistance of the following people, whom I would like to thank: Renée Blake, Lisa Davidson, Penny Eckert, Rudi Gaudio, Ron Butters, Keith Walters, Barbara Johnstone, and three anonymous reviewers. Earlier versions of this article were presented at the Linguistic Society of America Annual Meeting in Oakland, California, in January 2005, and the 72nd Southeastern Conference on Linguistics (SECOL) in Raleigh, North Carolina in April 2005. I thank the participants at those events for their questions and comments. All errors are, of course, my own.

² Gaudio measured two types of pitch range (gross pitch range and restricted pitch range), as well as pitch dynamism. While Gaudio does report some significant findings for pitch dynamism (the reader is referred to his work for the full analysis), neither measurement of pitch range yielded significant results.

³ The Rainbow Passage is a public domain text commonly used in acoustic and perceptual research. It briefly describes the scientific and historical characteristics of rainbows.

⁴ The semantic neutrality of the passage is not necessarily guaranteed. 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. Ron Butters (p.c.) has stated that some may view the narrator's actions as overly masculine, thus encouraging an assignment of "straight" sexuality, while others may view his actions as unconcerned about women, thus perhaps encouraging the assignment of a "gay" sexuality. In light of this potential variability in interpreting the narrator's behavior, as well as the fact that the semantic content was controlled throughout the experiment, I believe that any individual listener bias in the data set is sufficiently controlled.

⁵ An investigation of the voiced palato-alveolar fricative /ʒ/, as well as both the voiceless and voiced palato-alveolar affricates, /tʃ/ and /dʒ/, respectively, was excluded primarily because these variables do not feature prominently in the previous literature. The stimulus passage (see Appendix) featured two tokens of /dʒ/ and one token of /tʃ/.

⁶ All scales, both for the pre-test group and the experimental groups, were 7-point semantic differential scales (Gaudio 1994). These scales presented two opposing adjectives with a 7-point range in between. Raters are asked to circle one of the numbers that they feel most closely represents the speaker they are rating (see Table 1).

⁷ Shortening sibilants was done by deleting 17% of the total length of frication from the center of the segment, where frication amplitude was the highest.

⁸ Henton 1989, 1995 argues extensively for the calculation of pitch ranges in the logarithmic semitone scale, as opposed to the linear Hertz scale, insofar as the ear perceives pitch in a logarithmic fashion. Moreover, Jassem 1971 argues that a restricted pitch range (comprised of four standard deviations from the mean pitch) should be used when analyzing an utterance. This is done to ensure that the measured range will exclude any outliers or rogue data and more accurately reflect actual fluctuations in pitch.

⁹ There is the possibility that if we alter two, and only two, variables between the passages, listeners' attention is artificially drawn to these variables, which then become more salient than they would be otherwise. I note this possibility and see it as an unavoidable side effect of this type of controlled perception research.

¹⁰ Note that the polarity of the adjectives is not always what would be expected. For example, compare Straight/Gay and Effeminate/Masculine. While popular conceptualizations of gayness would align it with effeminacy, the scales used here reverse that polarity.

¹¹ Widening pitch ranges was done in exactly the same way as narrowing them had been done. The new recording of the passage was ToBI coded. Within each intonational phrase, the overall pitch range was stretched by 25%, while preserving the shape of the pitch contours and the levels of pitch dynamism. In terms of sibilant length, a portion of sibilant frication corresponding to 17% of the original total segment was copied and spliced into the center of the sibilant duration, where frication amplitude was the highest. This allowed for the lengthening of segments with material from those same segments, which were pasted in the portion of the segment where alteration was least likely to be noticed.

¹² A reviewer points out that this subject population can certainly not be said to be entirely random, insofar as students in these particular disciplines are perhaps more attuned to language, culture,

and gender differences than others are. I concede this as a valid critique and a potential shortcoming of the current research.

¹³ My use of quotation marks around “gay-” and “straight-” is intended to highlight that I am not considering the two speakers’ own sexual identifications, but rather what the pre-test group perceived to be the speakers’ sexualities.

¹⁴ The Effeminate/Masculine scale has a mean score across the four stimuli of 2.48, with a standard deviation of 0.48; the Straight/Gay scale (when reversed) has a mean score across the four stimuli of 2.65, with a standard deviation of 0.57.

¹⁵ I use the Effeminate/Masculine scale here as an assessment of the gendered identity of the speaker. Since all listeners identified the speaker as a man, I take their affective judgments of his relative effeminacy or masculinity to represent their opinions of his performances of “man-ness,” i.e., his gender. Likewise, I use the Straight/Gay scale to represent listeners’ perceptual judgments of his sexual identification.

¹⁶ Note that this order effect can also be taken to speak to the caution regarding artificially drawing listeners attention to certain variables mentioned in n. 9. In other words, the artificiality of the attention paid to pitch range and sibilant duration caused by the empirical design may be attenuated when the target stimuli are placed in the fourth position, wherein a contrast between all the speakers is instead highlighted.

REFERENCES

- Avery, Jack, & Liss, Julie (1996). Acoustic characteristics of less-masculine-sounding male speech. *Journal of the Acoustical Society of America* 99:3738–48.
- Bucholtz, Mary, & Hall, Kira (2004). Theorizing identity in language and sexuality research. *Language in Society* 33:501–47.
- Boellstorff, Tom, & Leap, William (2004). Introduction: Globalization and “new” articulations of same-sex desire. In William Leap & Tom Boellstorff (eds.), *Speaking in queer tongues: Globalization and gay language*, 1–21. Chicago: University of Illinois Press.
- Cameron, Deborah, & Kulick, Don (2003). *Language and sexuality*. London: Cambridge University Press.
- Eckert, Penelope (2000). *Linguistic variation as social practice: The linguistic construction of social meaning in Belten High*. Oxford: Blackwell.
- (2002). Demystifying sexuality and desire. In Kathryn Campbell-Kibler et al. (eds.), *Language and sexuality: Contesting meaning in theory and practice*, 99–110. Stanford: Center for the Study of Language and Information.
- Fox, Robert Allen, & Nissen, Shawn (2002). Age related changes in the acoustic characteristics of voiceless English fricatives. *Proceedings of the 2002 Conference of Linguistics and Phonetics*. Chiba City, Japan: Meikai University.
- Fry, Dennis (1958). Experiments in the perception of stress. *Language and Speech* 1:126–52.
- Gaudio, Rudy (1994). Sounding gay: Pitch properties of gay and straight men. *American Speech* 69:30–57.
- Henton, Caroline (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.
- Jacobs, Greg (1996). Lesbian and gay male language use: A critical review. *American Speech* 71:49–71.
- Jassem, Wiktor (1971). On the pitch and compass of the speaking voice. *Journal of Phonetics* 1: 59–68.
- Johnstone, Tom, & Scherer, Klaus (1999). The effects of emotion on voice quality. *Proceedings of the 14th International Conference of Phonetic Sciences*. Berkeley: University of California.
- Klatt, Dennis (1974). The duration of [s] 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.

- Lehiste, Ilse (1975). Some factors affecting the duration of syllabic nuclei in English. In G. Drachman (ed.), *Proceedings of the First Salzburg Conference on Linguistics*, 81–104. Tübingen: Gunter Narr.
- Levon, Erez (2004). Examining a gay prosody: Issues in theory, methodology and identity. Paper presented at Lavender Languages and Linguistics XI, American University, Washington, D.C.
- (2006). Hearing “gay”: Prosody, interpretation and the affective judgments of men’s speech. *American Speech* 81:56–78.
- Linville, Sue Ellen (1998). Acoustic correlates of perceived versus actual sexual orientation in men’s speech. *Folia Phoniatrica et Logopaedica* 50:35–48.
- Ochs, Elinor (1990). Indexicality and socialization. In James Stigler, Richard Schweder & Gilbert Herdt (eds.), *Cultural psychology*, 287–309. Cambridge: Cambridge University Press.
- (1992). Indexing gender. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context*. Cambridge: Cambridge University Press.
- 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 (n.d.). The stylistic use of phonation type: Falsetto, fundamental frequency and the linguistic construction of personae. Ms., Stanford University.
- Rogers, Henry, & Smyth, Ron (2003). Phonetic differences between gay- and straight-sounding male speakers of North American English. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 1855–58. Barcelona: Universitat Autònoma de Barcelona.
- , ———, & Jacobs, Greg (2000). Vowel and sibilant duration in gay- and straight-sounding male speech. Paper presented at IGALA 1, Stanford University.
- Scherer, Klaus (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality* 40:191–210.
- 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 one of my own kind once’: The erasure of desire in the name of identity. *Language and Communication* 23:123–38.

(Received 1 August 2005; revision received 15 August 2006;
accepted 28 August 2006; final revision received 16 November 2006)