Title
Social Judgments and Their Acoustic Cues in Read Speech

Journal
UC Berkeley PhonLab Annual Report, 3(3)

ISSN
2768-5047

Author
Babel, Molly

Publication Date
2007

DOI
10.5070/P765x7n2tt

Social judgments and their acoustic cues in read speech*

Molly Babel
University of California, Berkeley

1 Introduction

In perceiving speech, listeners understand the talker's intended message by extracting phonetic segments and suprasegmental information from the acoustic signal and mapping the information onto meaningful chunks in the lexicon. The acoustic signal, however, contains talker-specific characteristics as well, including phonetic cues for typical sociolinguistic indexical variables such as gender, regional identity, and age. Listeners process this type of type of information, storing it in memory (Palmeri et al. 1993).

To some extent, these talker-specific characteristics can be discussed as single isolated social categories, and gender may be the most well-studied of these characteristics. Speakers encode their gender in speech production from early childhood before any of the sexual dimorphic physiological changes take place that allow for men, on average, to be larger than women (Sachs et al. 1973, Perry et al. 2001). This gender information is used by listeners in processing speech sounds. Strand and Johnson (1996) synthesized the American English fricatives /s/ and /ʃ/ on a seven point continuum and appended the sounds to naturally produced vowel-consonant closures from both male and female talkers that ranged from *sod* to *shod*. Male talkers elicited more *sod* responses than females, presumably due to listeners' expectations that men, with typically larger vocal tracts, should produce fricatives with lower spectral center of energy than women (p. 18). In a second experiment, Strand

*Thanks very much to Ron Smyth and Henry Rogers for providing me with their database of recordings. I would also like to extend my gratitude to Keith Johnson and John Ohala for providing me with insightful comments that improved this paper immeasurably. This work could not have been done without research participants; thank you all for volunteering your time.

and Johnson visually presented a face along with the audio stimuli. The results show that listeners integrate the visual image with the acoustic signal to perceive the identity of the fricative. More recently, Munson et al. (2006b) investigated how knowledge about a talker's sexual orientation is used to interpret fricatives on an /s/-/ʃ/ continuum. In this experiment, listeners judged lesbian-sounding women as producing more [ʃ] tokens than straight-sounding women. There was no difference in fricative identification for men's voices.

The use of gender information in speech processing has also been extended to vowels. Johnson et al. (1999) investigated the effects of visual stimuli on the perception of an ambiguous vowel continuum. Continua were constructed from /ʊ/ to /ʌ/ as in the words *hood* and *hud* for stereotypical- and non-stereotypical-sounding male and female voices. In the first experiment two faces were used in the visual stimuli: a female face rated as the most female-looking and a male face rated as the most male-looking (these ratings were from a group of participants that simply judged the gender typicality of the face and were not exposed to any speech stimuli). There were also four talkers used in this experiment: the most and least female-sounding female voices and most and least male-sounding male voices (again, an independent group of participants judged the perceived gender typicality of these voices) were used as the stereotypical and non-stereotypical sounding voices for the audio stimuli. These two faces and four voices were incorporated into audio-visual clips in all possible combinations. Listeners judged the audio-visual tokens as either instances of *hood* or *hud*. The results show that listeners chose more *hood* responses for female faces than male faces, there were also more *hood* responses when the original voice was female, the more stereotypical sounding female voice elicited more *hood* responses, and the more non-stereotypical sounding male voices received more *hood* responses. These results are surprising because the most stereotypical female voice had a lower fundamental frequency than the non-stereotypical female voice. Johnson and his colleagues then presented two groups of listeners the *hood* to *hud* vowel continuum with a voice that had been synthesized to sound ambiguously male or female. The first group of listeners was told that the talker was female and the second group was told the talker was male. Listeners were more likely to label a stimulus as *hood* if they had been told the talker was female.

In addition to gender, speakers convey their regional identity through speech. Clopper et al. (2006) found that listeners can classify a talker's regional dialect in a four-alternative (North, General American, Mid-Atlantic, and South) forced-choice task from sentential stimuli. In this experiment listeners also rated the perceptual similarity of voices differing in region and gender. Listeners were divided into four listener groups based on their own regional affiliation and whether they had lived in more than one region. All listener groups rated same-dialect voice pairs as more similar than different-dialect voice pairs. Multidimensional scaling of listener responses revealed that listeners were using two dimensions to classify talkers: geography and markedness. Further analysis by Clopper and colleagues found that listeners who had lived in only one region placed more emphasis on the geography dimension while the mobile group used both dimensions equally.

The effect of regional identity on speech perception was investigated in Niedzielski (1999). Niedzielski presented Detroit-based listeners (n=41) with 50 sentences spoken by a single female talker and instructed them to match the vowel of the target word in the sentence with one of six synthesized vowels. The listeners were divided into two groups: half of the listeners were told the talker was from Canada, while the others were told she was from Michigan. Listener responses greatly depended on the talker's purported regional identity. Listeners perceived the target vowels differently based upon knowledge of the talker's regional identity.

Building on these studies, this paper investigates how judgments of talkers' sexual orientation correlate with other judgments about talkers through two perception experiments. In an additional experiment judgments of sexual orientation are compared across dialects of English. Sexual orientation and speech is a burgeoning field in linguistics (Munson and Babel, to appear), and Section 2 provides a review of pertinent studies of gay and gay-sounding speech. The three experiments that investigate social category perception are discussed in Sections 3, 4, and 5. In Section 6 the listener judgments from the current experiments and perception data from Smyth et al. (2003) are modeled with stepwise linear regressions to determine the acoustic information on which listeners based their judgments. I argue that while social talker judgments may overlap with one another as abstract social identities that coalesce to create a whole individual's identity, the acoustic cues on which listeners base each individual judgment differ.

2 Speech and sexual orientation

In addition to region and gender, sexual orientation is another indexical feature that can be encoded in the speech signal. It has become well established in the literature that naive listeners can accurately judge the sexual orientation of talkers from spoken stimuli. In an early study Gaudio (1994) presented thirteen listeners excerpts of two passages spoken by eight talkers (4 gay, 4 straight). Listeners were asked to rate the talkers on 7-point scales for four adjective pairs: straight/gay, effeminate/masculine, reserved/emotional, and affected/ordinary. Gaudio found that listeners were largely accurate (88% accuracy) in judging talkers' sexual orientation and "straight" and "gay" judgments were strongly correlated with "masculine" and "effeminate" ratings, respectively. Various measurements of pitch variability and range were taken from the talkers. Pitch range was calculated as the difference between maximum F0/minimum F0, maximum F0/mean F0, and maximum F0/median F0. These values were measured on 15-second segments. Variance and standard deviation were calculated for each segment. To capture dynamic pitch movement, Gaudio also considered the average change in F0 by calculating the absolute value of each pitch contour, the number of pitch changes from a positive to negative slope (and the reverse), the number of pitch excursions of 15 Hz or more, and the number of such excursions per second. None of these measurements correlated significantly with listeners' judgments. A study by Linville (1998) illustrated that listeners ($n = 25$) were able to correctly judge the sexual orientation of nine talkers (5 gay, 4 straight) 79.6% of the time when listening to read passages. Linville investigated a number of acoustic cues including aspects of /s/, F0, speech rate, and formants

1-3; she found that talkers' average /s/ duration and average /s/ peak frequency accounted for 88% of the variance of listeners' responses in her experiment.

Smyth et al. (2003) investigated the speech of gay and straight men in Toronto. In their 2003 paper they describe the creation of a database of 25 (17 gay, 8 straight) male voices and the results of an acoustic and perceptual study investigating the relationship between average F0 and judgments of sexual orientation and masculinity/femininity. This study is critical to the current investigation as the database developed by Smyth and colleagues is used in the experiments discussed below. They recorded spontaneous speech, the reading of a scientific text, and the reading of a dramatic text from their 25 talkers and presented thirty seconds of each type of speech to 46 Toronto-based listeners who judged each talker as gay or straight. Ratings of sexual orientation interacted with the type of listener stimuli. That is, while gay men were rated as more gay-sounding than straight men across all speech styles (mean "sounds gay" value of 0.56 for gay talkers and 0.28 for straight talkers), straight talkers were rated as more gay-sounding for the scientific passage than for the other two passages (mean "sounds gay" values of 0.35, 0.26, and 0.22 for straight talkers for the scientific passage, dramatic passage, and spontaneous speech, respectively). Such a finding suggests that more formal speech styles are associated with gay-sounding speech. The stimuli were again presented to 16 listeners who judged the voices as masculine- or feminine-sounding. These judgments were highly correlated with judgments of sexual orientation across all discourse types. The mean ratings, however, were lower for masculine/feminine ratings, demonstrating the labels 'gay' and 'feminine' are not synonymous. Smyth et al. also extracted mean pitch from a single sentence from each discourse type and, like previous investigations, the relationships between F0 and gay/straight and feminine/masculine were not significant.

Recent investigations of the acoustics and perceptions of the speech of gay/bisexual men and lesbian/bisexual women suggest that these social communities share their own particular region-specific dialects and that the acoustic patterns of the gay/bisexual and the lesbian/bisexual communities are not imitations of women and men, respectively. Pierrehumbert et al. (2004) examined read speech from a large group (n=103) of gay and straight male and female talkers from the Chicago area. F1, F2, F3, and duration were measured for the vowels /i/, /eɪ/, /æ/, /a/, and /u/. Results show that gay men and lesbian women had more dispersed vowel spaces than straight men and women, respectively. In this case, this was due to lesbian women having lower F2 in back vowels and gay men having more extreme formants for the front vowels. Pierrehumbert et al.'s results suggest that the speech patterns of the gay community are not mimicked speech patterns of the opposite sex.

A recent investigation on gay and lesbian speech patterns is a study conducted by Munson et al. (2006a) through three experiments. Munson et al. (2006a) collected read speech from 44 gay, lesbian, bisexual, and straight talkers. Contrary to the findings in Pierrehumbert et al. (2004), they did not find any significant effect of vowel dispersion; this result is likely because Munson and colleagues were investigating a different community of talkers. Lesbian and straight women, however, did differ significantly in F1 of /ɛ/ and F2 of /oʊ/, while the groups of men differed significantly in the F1 of /ɛ/ and /æ/, with both vowels having lower F1 in the speech of gay men. Gay and straight men were also found to differ in the skewness of /s/ spectra. In a perception experiment, Munson et al. (2006a) presented listeners with single word tokens from the read speech of the 44 talkers. Listeners rated self-identified gay and lesbian talkers as more gay-sounding than straight talkers. After a main effect had been found for sexual orientation, men and women were analyzed independently. There was a main effect of sexual orientation for men [F (1,20) = 10.5, $p < 0.01$]; gay talkers were rated as more gay-sounding than straight talkers. A main effect was also revealed for women [F (1,20) = 5.1, $p < 0.05$] with lesbian women being rated as more lesbian-sounding than straight women. The acoustic differences noted in gay and straight men's voices predicted the listener ratings. For women, regression analysis revealed that women's voices would more likely be rated as lesbian-sounding with lower F1 in low vowels, lower F2 in back vowels, and an overall less-expanded vowel space. The third experiment was a perception task which illustrated how perceptions of sexual orientation are related to judgments of perceived clarity and perceived height. Gay men were rated as speaking more clearly than straight men, while straight women were rated to sound more clear than lesbian women. Judgments of perceived height and perceived clarity accounted for 81% and 73% of the variance in judgments of sexual orienation for women and 48% and 53% of the variance for male judgments of sexual orientation. Munson et al. suggest that speech style (i.e., speech clarity) may be an indicator of membership in the gay community. They also conclude by cautioning that "judgments of sexual orienation from single words may not be judgments of sexual orienation per se, but judgments of another parameter that is robustly encoded" in the speech signal as well (p. 233).

In summary, listeners are able to judge the actual sexual orientation of talkers from both sentential and single word stimuli. Moreover, the literature reveals that stereotypical features of gay-sounding speech such as higher pitch are not actual attributes of gay-sounding speech. The findings also suggest that gay-sounding speech is more salient in formal speech styles. The results of Pierrehumbert et al. and Munson et al. propose that gay-sounding speech is dialect specific with gay-sounding acoustic features being limited to small-scale communities of practice. The relationship between judgments of gay-sounding speech and other talker characteristics suggest that the perception of gay speech may be mediated by other social stereotypes. These conjectures are examined in the experiments presented below.

3 Experiment 1: Gay-sounding speech across dialects

The purpose of Experiment 1 was to examine how judgments of sexual orientation made by a group of listeners from one dialect group compare to those of another dialect. To date, no experiment on sexual orientation and speech has presented the same stimuli to two distinct populations and given listeners the same task. This experiment directly tests the predictions made by Pierrehumbert et al. and Munson et al. by testing gay-sounding speech cross-dialectally. If judgments of sexual orientation are dialectally conditioned, we predict that listeners with different dialect backgrounds will show different patterns of response.

3.1 Stimuli

The stimuli used in this experiment were 25 readings of the Rainbow Passage, recorded by Smyth et al. (2003). The Rainbow Passage is a phonetically balanced description of past and present scientific explanations for rainbows that "evokes little emotional involvement in the speaker" (Smyth et al. 2003: 333). The selection of the Rainbow Passage presented to the listeners in the current study was the same selection played to Toronto listeners in Smyth et al. The Rainbow Passage text used in these experiments is provided in Appendix A.

The stimuli were presented to listeners as .wav files at a sampling rate of 22 kHz. The duration of each stimulus was 25-30 seconds. All 25 talkers studied by Smyth et al. were used.

3.2 Participants

Listeners were recruited from a list of undergraduate students enrolled in linguistics courses at the University of California, Berkeley. Subjects were asked to participate in an experiment investigating social judgments listeners make about talkers. There was no mention of sexual orientation in the recruitment process.

Eighteen native speakers (10 females, 8 males) of American English aged 18 to 33 participated as listeners in this experiment. An attempt was made to recruit subjects who had spent the majority of their lifetimes in California. Subjects reported no speech, hearing, or language disorders. All participants were compensated $5 for participating in the experiment. The data from one listener was removed due to technical difficulties.

3.3 Procedure

Listeners were asked to listen to the passages and judge the sexual orientation of the talkers. They were informed that there were 25 talkers each reading the same passage. The task of the listener was to rate gay- and straight-soundingness on a 5-point scale. Listeners were instructed to use the scale as follows: a definitely gay-sounding talker received a score of [1], a talker who was neither gay- nor straight-sounding was to be given [3], and a talker who was definitely straight-sounding was to be considered a [5]. The values [2] and [4] were used to note intermediate levels of gay- and straight-soundingness.

Responses were logged automatically on a equal-interval 5 point button-box. Listeners were told to wait until the talker finished reading the passage before responding on the button-box. Buttons [1], [3], and [5] were labeled with the descriptions given above as a reminder for the listeners'. The stimuli were presented to listeners over headphones at a comfortable level at a workstation. The sound files were randomized separately for each participant and presented using E-prime experiment software (Schneider et al. 2002). The total duration of Experiment 1 was less than 15 minutes.

3.4 Analysis and Results

Prior to the analysis, data from two listeners were removed. The first was a non-native speaker of English and the second listener was not a Californian. This left 15 listeners for the analysis.

Listeners' responses were standardized to z-scores before being submitted to the analysis. This factor will be called California gay-sounding judgment (henceforth, CGJ) for the duration of the paper. After the experiment, listeners were asked if they were aware that the talkers were from Canada and how that affected their responses. Seven listeners commented that they guessed (at least some) of the talkers were from Canada; all of these listeners explicitly stated that the more Canadian-sounding the talker was the more likely they were to judge them as more gay-sounding.[1] Listeners' explicit judgments about knowledge of Canadian talkers was coded as a binary 'yes' or 'no' variable.

[1] These statements contrast with the judgment made by the listener whose data was removed for not being Californian. This listener had recently moved to California from the Northern Cities dialect area. After the experiment, she stated that the Canadian-sounding talkers sounded as if they could be residents of her dialect area. She said she was less likely to judge these talkers as gay-sounding.

The data was submitted to a stepwise hierarchical regression model with the standardized CGJ as the dependent measure and percentage of gay-sounding judgments across the three passage styles (PSG) from Smyth et al. (2003) and actual sexual orientation of the talkers (SO) predictors. The model excluded SO as predictor, using only PSG values. PSG accounted for 36% of the variance in CGJ [F (1, 373) = 215.6, $p > 0.001$]. The correlation between CGJ and PSG had a *Pearson's R* value of -0.61, indicating that talkers with a high PSG rating by Smyth et al.'s Canadian listeners were given a more gay-sounding rating by Californians. Figure 1 illustrates this relationship with PSG and average CGJ ratings for each talker.

A post-hoc test showed that listeners who explicitly stated they heard Canadian talkers gave more gay-sounding judgments (M = -0.14) than listeners who did not state they heard Canadian-sounding talkers (M = 0.09) [$t(373) = 2.24$, $p < 0.05$]. A second post-hoc t-test examined how well California listeners were able to rate self-identified gay talkers as more gay-sounding than straight talkers. Listeners did indeed label gay talkers as more gay-sounding (M = -0.17) than straight talkers (M = 0.36) [$t(373) = -5.03$, $p < 0.001$].

3.5 Discussion

Experiment 1 compared ratings of sexual orientation of Canadian male talkers from Canadian listeners and those with California listeners. While there was a significant relationship between these two sets of judgments, the ratings given by the groups were not identical. California listeners labeled talkers' sexual orientation accurately, as did the Canadian listeners; gay talkers were more likely to be rated as gay-sounding than straight talkers. Listeners who had knowledge of the regional identities of the talkers were more likely to judge talkers as more gay-sounding, suggesting that explicit stereotypes influence rating scores.

The disparity between the gayness ratings provided by the two groups lend credence to the notion that gay-sounding speech is dialectically conditioned and not a judgment that is equally generalizable across listener groups. The next experiments will examine how

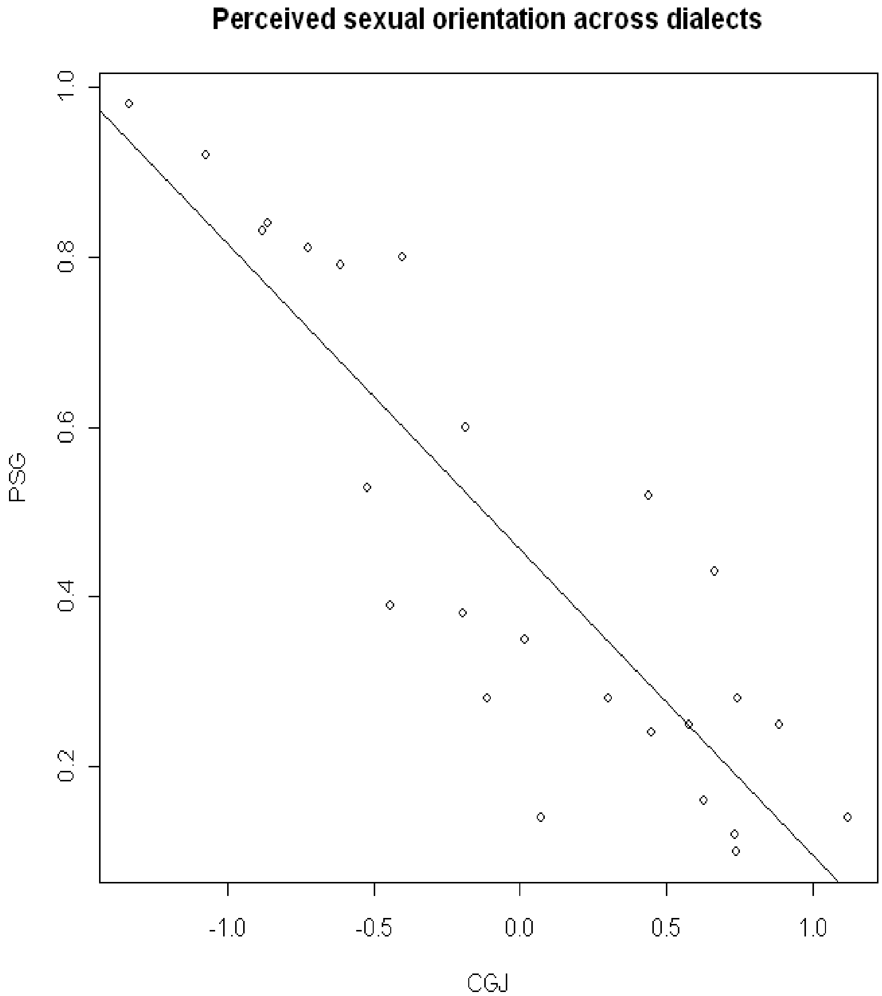

Figure 1: Averaged PSG values from Smyth et al. (2003) plotted against averaged CGJ from California listeners. A negative CGJ value is more gay-sounding and a positive number is more straight-sounding. A PSG score approaching 1 is more gay-sounding while a value near 0 is more straight-sounding.

judgments of sexual orientation correlate with other social judgments.

4 Experiment 2: Reading ability

The intention of Experiment 2 was to examine the relationship between a general judgment about a person's reading ability and judgments about gay-sounding speech. It was assumed that if judgments about gay-sounding speech relate to formal speech styles and speech clarity, then ratings of reading ability should access the same social category and produce parallel judgments.

4.1 Stimuli

The Smyth et al. recordings of the Rainbow Passage were used again in Experiment 2.

4.2 Participants

Listeners were recruited from a list of undergraduate students enrolled in linguistics courses at the University of California, Berkeley. Subjects were asked to participate in an experiment investigating what type of information people convey about themselves when they read aloud. There was no mention of sexual orientation in the recruitment process.

Fifteen native speakers (9 females, 6 males) of American English aged 18 to 28 from the University of California, Berkeley participated as listeners for this experiment. Subjects reported no speech, hearing, or language disorders. All participants were compensated $10 for participating in the experiment.

4.3 Procedure

Listeners were asked to listen to the passages and determine the talkers' reading ability. They were informed that there were 25 talkers each reading the passage and they were asked to rate how well the speakers read the passage on scale from [1] to [5]. Listeners were instructed to use the scale as follows: a good reader — an individual whose reading was clear, expressive, and pleasing — was considered a [1], while a poor reader — an individual whose reading was mumbled and filled with mispronunciations, hesitations, and false starts — was given a rating of [5]. Ratings [2] through [4] were used to note intermediate reading levels.

Responses were logged automatically on a equal-interval 5 point button-box. Listeners were told to wait until the talker finished reading the passage before responding on the button-box. The stimuli were presented to listeners over headphones at a comfortable level. The sound files were randomized for each participant and presented using E-prime experiment software. The total duration of Experiment 2 was less than 15 minutes.

This set of subjects also participated in Experiment 3, which was run immediately after Experiment 2. After Experiment 3, participants were asked about their impression of the data. Specifically, listener were asked:

1. Did it sound like the talkers were from a distinct dialectal region? Could you tell where they were from?
2. Do you have any judgments about the sexual orientation of the talkers? Did the majority of them sound gay or straight?

Participants' responses to these questions are shown in Appendix B.

4.4 Analysis and Results

Prior to the analysis, the perceived reading ability responses were standardized into z-scores. Participants' responses from the post-experiment questionnaire were coded and added into the data set in the following ways: One common response was that the talkers sounded British. It is a popular culture stereotype in American society that British accents are particularly gay-sounding. For this reason, any participant who mentioned British accents was coded 'yes' on the variable "sounded British" column. Similarly, because talkers with a Canadian accent may sound gay to these listeners, listeners who noted Canadian accents were coded 'yes' on the variable "sounded Canadian". Recall that the majority of talkers self-identified as gay. Despite this, most listeners commented that only a few talkers had stereotypically gay-sounding voices. Since this question did not reveal which talkers listeners felt were the most gay-sounding, listeners who said there were some gay-sounding voices in the talker set were coded 'yes' on the variable "Sexual Orientation (SO)."

A logistic regression with standardized z-score of reading ability as the dependent measures with CGJ, sounded British, sounded Canadian, and SO as the independent measures. There was a single interaction between CGJ and SO [$F(1,\ 366) = 349.26$, $p < 0.05$]. This interaction is shown in Figure 2. In this figure the black data points represent data from listeners who were aware of gay-sounding talkers in the data set. The regression line for these listeners begins at a negative point on the Reading ability axis and progresses with a positive slant through the figure. This regression line illustrates that listeners who were aware of gay-sounding talkers in the data set judged the more gay-sounding talkers as better readers and more straight-sounding talkers as worse readers. Listeners who did not feel there were gay-sounding talkers in the data set are shown with light gray data points. The regression line for this subset of listeners begins with a positive integer and has a negative slope; these listeners judged gay-sounding talkers as worse readers than straight-sounding talkers.

4.5 Discussion

Experiment 2 revealed no relationship between judgments of reading ability and judgments of sexual orientation by Californians. This finding does not provide support for claims in the literature suggesting that gay-sounding speech is a formal speech style salient in read speech, but rather does suggest that associating gay-sounding speech with speech style and speech clarity is limited to certain dialect communities. This experiment did, however, illustrate that awareness of talkers' social identities influences other types of judgments. Listeners who were sensitive to the presence of gay talkers in the recordings were more likely to judge gay-sounding talkers as good readers.

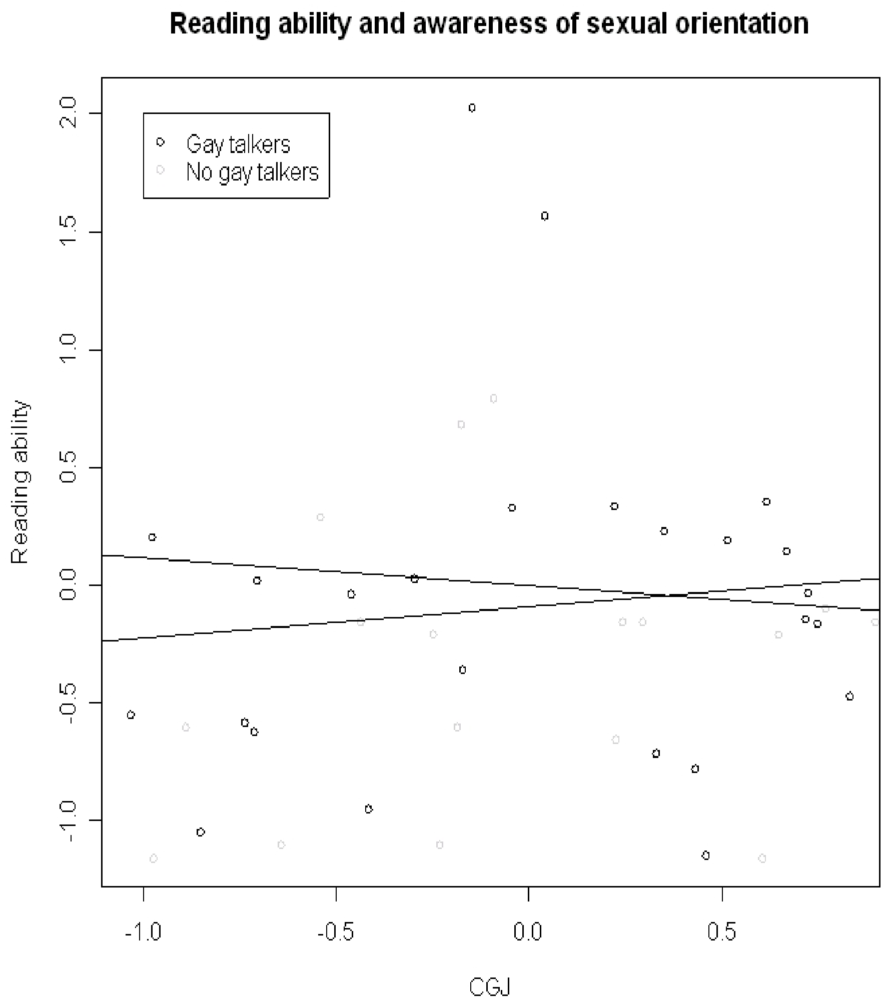

Figure 2: Averaged reading ability judgments plotted against averaged CGJ from two groups of listeners. Negative CGJ and reading ability values indicate more gay-sounding talkers and better readers and positive numbers show more straight-sounding talkers and worse readers. The black data points represent data from listeners who were aware of gay-sounding talkers in the data set. The gray data points represent listeners who did not hear gay-sounding talkers. The regression line that begins at a negative point on the Reading ability axis is that of the subset of listeners who heard gay-sounding talkers. The regression line starting at a positive point on the Reading ability axis represents listeners who did not hear gay-sounding talkers.

5 Experiment 3: Listener Presence

The rationale for Experiment 3 was to determine whether asking participants to make judgments of a listener's presence would be a more sensitive question, perhaps revealing judgments of speech clarity and it's relationship to sexual orientation that reading ability did not capture.

5.1 Stimuli

The Smyth et al. recordings of the Rainbow Passage were used again in Experiment 3.

5.2 Participants

The same 15 subjects who volunteered for the second experiment participated in this experiment as well.

5.3 Procedure

Experiment 3 had two parts. First, listeners were asked to determine whether the talker was reading the passage to someone. That is, listeners were forced to judge whether there was an individual in the room listening to the passage at the time it was being read by the talker. Participants were told to press button [1] if they thought there was a listener present at the moment of recording, and button [5] if they felt no one was present. It was predicted that talkers with clearer speech would be judged as more likely to be reading to someone, as the clearer speech would be indicative of an attempt to communicate effectively and read expressively. Again, listeners were told to wait until the talker had finished reading the passage before responding on the button-box.

The second part of Experiment 3 was dependent upon the answer to the first question. If listeners responded affirmatively that, *yes*, there was an individual listening to the reading of the passage at the moment of recording, they were directed to a second question asking whether the identity of the individual was a child or an adult. It was predicted that talkers who receivded more *child* responses would represent the more extreme examples of speech clarity.

Listeners were informed that they would be hearing the same set of talkers read the same passages as they did in Experiment 2 in a different random order.

5.4 Analysis and Results

In the first analysis, the listener presence variable – the response of 'listener' or 'no listener' – was the dependent factor in a logistic regression. CGJ, sounded Canadian, sounded British, and SO were the independent measures. There was a main effect of sounded Canadian [$F(1, 371) = 505.76$, $p < 0.05$] and an interaction between sounded British and sounded Canadian [$F(1, 367) = 498.21$, $p < 0.05$]. Listeners who reported hearing Canadians in the talker set were more likely to judge talkers as reading to someone (63% 'yes' listener responses) than listeners who did not hear Canadians in the data set (51% 'yes' lsitener responses). There was little difference in the percentage of 'yes' listener responses between the groups of listeners who mentioned British accents; 57% 'yes' responses from listeners who reported British talkers and 54% 'yes' responses from listeners who did not say anything about British talkers. The main effect of CGJ was slightly beyond the level of significance [$F(1, 373) = 512.11$, $p = 0.07$]. This trend is depicted in Figure 3; listeners who were rated as gay-sounding tended to be judged as reading to someone more often that straight-sounding talkers.

The second analysis examined the subset of responses with affirmative 'listener present' responses again using a logistic regression. The listener identity variable – child or adult – was the dependent measure while CGJ, sounded Canadian, sounded British, and SO were the independent factors. CGJ returned as a main effect [$F(1, 206) = 13.41$, $p < 0.001$]. Figure 4 shows this result; Talkers who were rated by Californians as gay-sounding were also rated as reading to a child more than straight-sounding talkers. There was an interaction between CGJ and sounded British [$F(1, 202) = 4.85$, $p < 0.05$], as shown in Figure 5. This illustrates that listeners who guessed British talkers were present in the data set were more likely to rate straight-sounding talkers as reading to an adult and gay-sounding talkers as reading to a child than listeners who did not perceive British talkers.

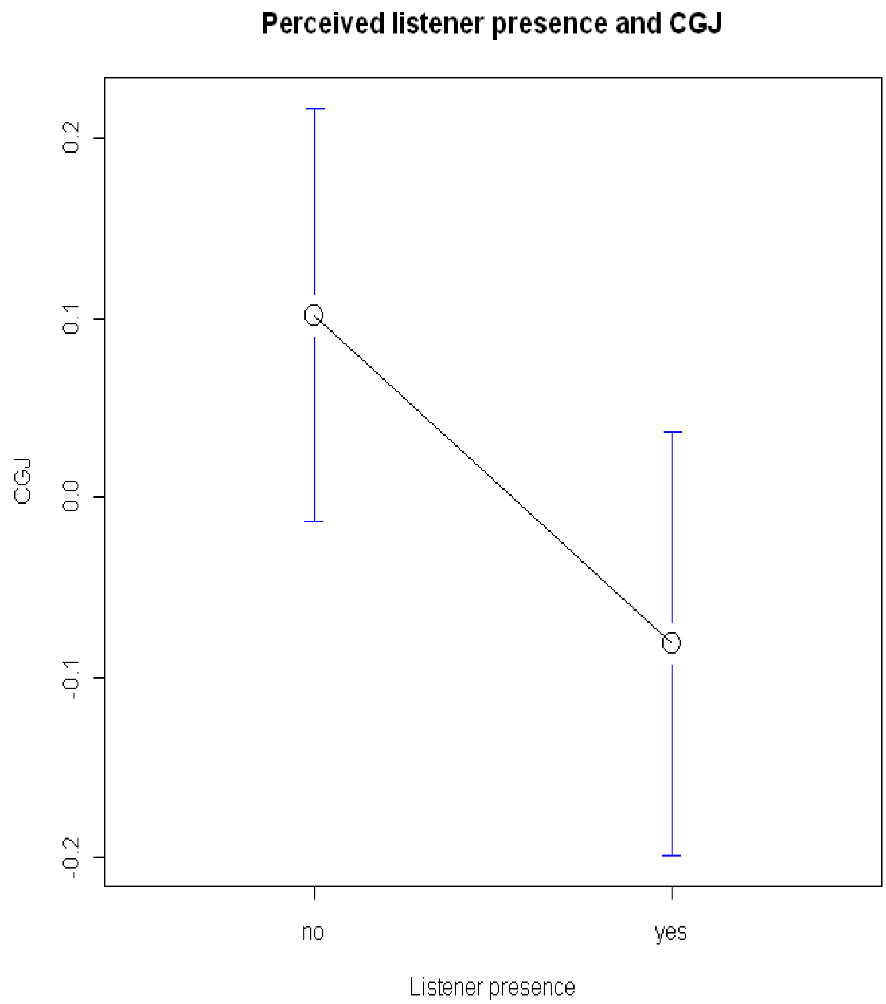

Figure 3: CGJ and perceived listener presence. 'Yes' and 'no' refer to listeners' responses regarding whether a listener was present in the room at the time of recording the Rainbow Passage. Note that a negative CGJ value indicates a more gay-sounding talker while a positive number indicates a straight-sounding talker. This trend was not statistically significant.

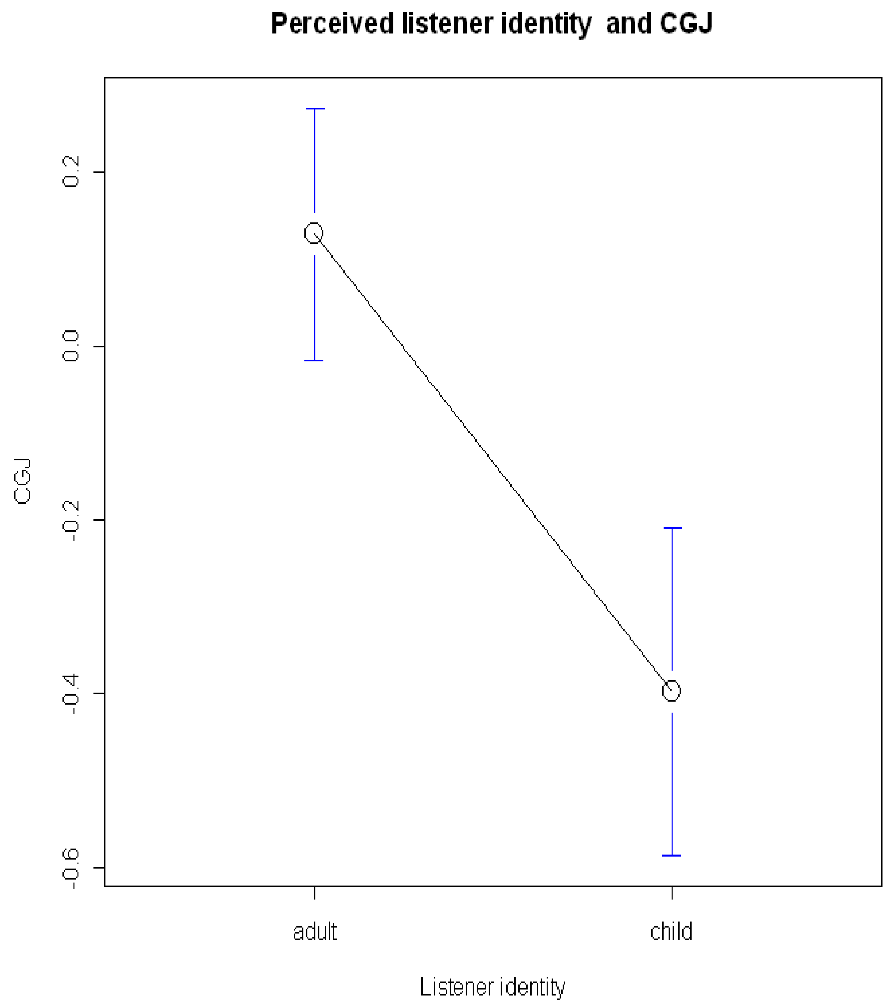

Figure 4: CGJ and perceived listener identity. 'Adult' and 'child' refer to listeners' responses regarding the identity of the listener present in the room at the time of recording the Rainbow Passage. Note that a negative CGJ value indicates a more gay-sounding talker while a positive number indicates a straight-sounding talker.

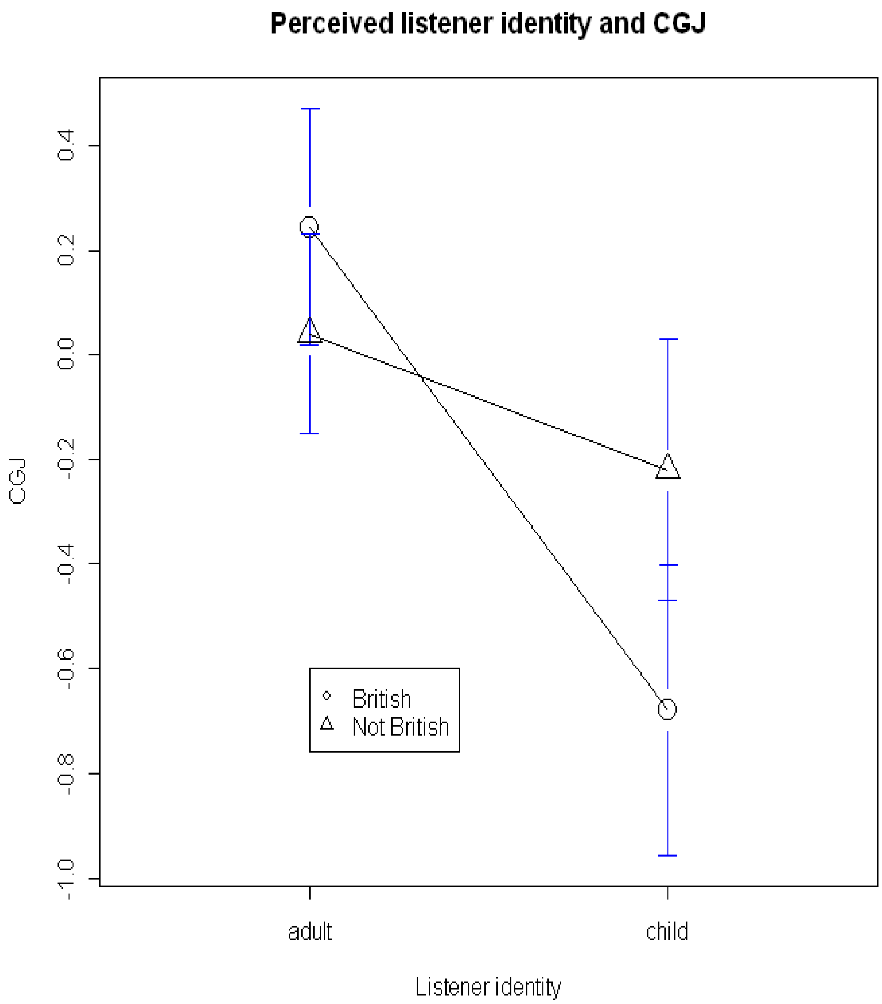

Figure 5: CGJ and perceived listener identity. 'Adult' and 'child' refer to listeners' responses regarding the identity of the listener present in the room at the time of recording the Rainbow Passage. Note that a negative CGJ value indicates a more gay-sounding talker while a positive number indicates a straight-sounding talker. 'British' indicates a subset of listeners who thought a few of the talkers sounded British while 'Not British' indicates the subset of listeners who did not feel British talkers were represented in the recordings.

5.5 Discussion

Experiment 3 demonstrated that asking participants to listen for cues associated with reading *to* someone was not tightly linked to perceived sexual orientation judgments made by California listeners. Asking listeners to determine the identity of the individual being read to was, however, strongly related to judgments of perceived sexual orientation. Gay-sounding talkers are perceived to be reading to children rather than adults. The second half of this experiment established the connection between judgments of gay-sounding speech and judgments of speech clarity found in previous research. There was a pattern in the data suggesting that expressive, clear speech produced for a child's benefit is associated with a British identity and that Canadian talkers are perceived as more likely to be reading to someone. How these connections are related to acoustic correlates of clear speech is a topic that will be explored in the acoustic investigations in Section 6 below.

6 Experiment 4: The acoustic cues for social judgments

The purpose of Experiment 4 was to determine the acoustic cues on which listeners based their judgments of sexual orientation, reading ability, listener presence, and listener identity. Such judgments are likely not to be based on arbitrary selections made by individual listeners, but rather pre-existing social stereotypes or judgments about how people sound based on linguistic experiences.

6.1 Procedure

Each phone was labeled in Wavesurfer (Sjölander and Beskow 2005) by the author. Phones were labeled with the Arpabet symbols; this is a broad transcription system that, for example, encapsulates closure duration and VOT into a single consonant's label. Diphthongs were also marked with one label. A script extracted means and standard deviations of formant values and F0. LPC formant tracking was used to calculate formant values at 2.5 ms intervals in 49 ms windows. These values among others were used in the linear models described below.

6.2 Choosing acoustic parameters

It was assumed that judgments of reading ability and listener presence would be related to acoustic measurements associated with talker intelligibility. This assumption was supported by the correlation between judgments of speech clarity and perceived sexual orientation reported in Munson et al. (2006a). Therefore, acoustic factors deemed significant for talker intelligibility in Bradlow et al. (1996) were used in the linear models. The overall range of F1, /a/ F2-F1, /i/ F2-F1, average vowel duration, and F0 variability were included as predictors in the linear regressions. Recall that all the speakers in the data set are male, eliminating any need to normalize across gender before extracting frequency values. F1 range was determined by subtracting the lowest average F1 from the highest average F1 in the vowel space of a talker. Because of the large number of vowels distinguished by height in the vowel space, F1 range is a good indicator of how well separated the vowels are, and, hence how intelligible speakers are. Bradlow et al. found the difference between F1 and F2 of the point vowels /a/ and /i/ to be negatively and positively correlated, respectively, with ratings of talker intelligibility. While Bradlow et al. found no significant relationship between speech rate and talker intelligibility, average vowel duration was included in the model. Intuitively, a talker with longer average vowel durations would be more pleasing to listen to than a talker that spoke excessively fast-paced. Variation in F0, grossly calculated as the standard deviation of F0, was also included in the model; it was predicted that a talker with more variability in F0 would have sounded like a more expressive reader, resulting in better perceived reading ability and positive judgments of listener presence. Following Munson et al., these parameters measuring talker intelligibility should also account for judgments regarding gay-sounding speech. This set of acoustic parameters will be referred to as the intelligibility parameters.

A second set of acoustic parameters associated with particular dialects were selected to be used in an independent set of linear regressions. These dialect parameters were comprised of mean F1 and F2 of /æ/, /a/, and /au/. Recall that the talkers are from Toronto and the listeners are from California, two dramatically different dialects of North American English. Notable features of California English relate to the position of /æ/ and /a/. Toronto is part of the region undergoing the Northern Cities Vowel Shift. Therefore, as compared to the variety of English spoken in the Toronto, in California English /æ/ and /a/ are produced with lower F2 and F1 (Labov 1994, Clopper and Pisoni 2004). The diphthong /au/ was also included in the set of dialect parameters because of the Canadian raising pattern of words like *about* being produced with a vowel with a nucleus raised to /u/.[2]

A third and final set of reading parameters were compiled based on the author's impressionistic judgments of reading ability after listening to the passages a number of times. The number of disfluencies and pauses, the duration of words and pauses, and the F1 range were the factors included in the set of reading parameters. Stuttering and word repetitions were the most common types of disfluencies in the data; these speech errors were simply counted for each talker. Pauses made at expected clausal boundaries were counted in addition to pauses that occurred in unexpected locations; the durations of the pauses were averaged for each talker. The mean duration of each word was calculated for talkers. F1 range was also included in this model because a good reading voice was expected to produce words clearly using a large vowel space, leading to an increased F1 range.

[2]The California and Toronto English dialects also pattern distinctly with respect the vowel /ɔ/. Specifically, the variety of English spoken in Toronto has this phoneme, while the speech of Californians does not. This also would have been an interesting vowel to consider as a factor in the set of dialect parameters, but it did not occur in the data set.

6.3 Analysis and Results

The sets of acoustic parameters were entered into stepwise linear regression models to predict listener responses for the talker judgments. The model fitting procedure determines which parameters are included as predictors and excludes those that do not account for a significant proportion of variance.

6.3.1 Canadian gay-sounding judgments: PSG

Munson et al. (2006a) found that 53% of the variance in judgments of perceived sexual orienation for males could be accounted for by independent judgments of speech clarity. The intelligibility parameters were, therefore, predicted to account for some of the PSG responses. As shown in Table 1, F0 variability and /i/ F2-F1 were the intelligibility factors included in the model, accounting for 25% of the variance. Gay-sounding talkers (M = 17 Hz, SD = 6 Hz) had more F0 variation than not gay-sounding talkers (M = 14 Hz, SD = 4 Hz). Gay-sounding talkers also had a larger mean difference in between F2 and F1 for the point vowel /i/ (M = 1670 Hz, SD = 222 Hz) than not gay-sounding talkers (M = 1602 Hz, SD = 119 Hz).[3]

Table 1: Predicting PSG judgments with intelligibility parameters.

Acoustic parameter	B	SE B	β
F0 variability	0.022	0.0097	2.261*
/i/ F2-F1	0.0004	0.0003	1.494
Multiple R^2=0.4178	Adjusted R^2=0.2528	$F(2, 21) = 4.89$	$p < 0.05$

Next, the dialect parameters were put into another model. Three of the six dialect features were included in the model, shown in Table 2; these acoustic measurements were

[3]The average values cited for the talkers is the result of a median split on the average values of the dependent variable. This method for determining average values for talker groups was conducted for all of the analyses. The fact that these values are reported as a median split explains why some of the differences are incredibly small.

able to account for 39% of the PSG responses. The formants of these vowels are shown in Table 3. The vowels /æ/ and /au/ of gay-sounding talkers have higher F2 values than those of straight-sounding talkers. The vowel /a/, on the other hand, has higher F2 for straight-sounding talkers.

Table 2: Predicting PSG with dialectical parameters.

Acoustic parameter	B	SE B	β
F1 /æ/	0.0033	0.0011	3.060**
F1 /au/	0.0018	0.0009	1.995
F1 /a/	-0.0017	0.0013	-1.357
Multiple R^2=0.4677	Adjusted R^2=0.3879	$F(3, 20) = 5.858$	$p < 0.01$

Table 3: Average formant values of vowels used in the dialect regression model for Canadian rated gay-sounding and straight-sounding talkers. Those marked with asterisks are those formants that were used in the hierarchical regression model.

Talker set	Vowel	F1	F2
Gay-sounding talkers	/a/	633	1146*
Straight-sounding talkers	/a/	635	1201*
Gay-sounding talkers	/æ/	628	1592*
Straight-sounding talkers	/æ/	576	1546*
Gay-sounding talkers	/au/	652	1325*
Straight-sounding talkers	/au/	604	1286*

The reading parameters were entered into the linear model to investigate how these parameters would predict the PSG values. None of the five of the parameters were included in the model. Clearly, the acoustic values in the reading parameter set were not those that listeners used to base their PSG judgments.

6.3.2 California gay-sounding judgments: CGJ

A series of stepwise regressions were carried out for the gay-sounding judgments made by California listeners. Since these judgments were not perfectly correlated with the ratings of sexual orientation from Toronto listeners, it was predicted that the regression models would capture this variation. Of the five intelligibility factors entered into the model, only F0 variability was used, as shown in Table 4. This factor was able to account for 24% of the variance in judgments of sexual orientation by California listeners. The set of gay-sounding talkers had more F0 fluctuation (M = 18 Hz, SD = 6 Hz), as compared with straight-sounding talkers (M = 13 Hz, SD = 3 Hz).

Table 4: Predicting CGJ with intelligibility parameters.

Acoustic parameter	B	SE B	β
F0 variability	-0.0664	0.0232	-2.861**
Multiple R^2=0.2711	Adjusted R^2=0.238	$F(1, 22) = 8.183$	$p < 0.01$

The dialect parameters were submitted to a second model. The two vowel formants used in the model seen in Table 5 – F1 of /æ/ and F2 of /a/ – account for 54% of the variance in the data. The values of these formants can be seen in Table 6. Gay-sounding talkers had higher F1 values for /æ/ and lower /a/ F2 than straight-sounding talkers.

Table 5: Predicting CGJ with dialect parameters.

Acoustic parameter	B	SE B	β
F1 /æ/	-0.0102	0.0021	-4.822***
F2 /a/	0.0015	0.001	1.489
Multiple R^2=0.5822	Adjusted R^2=0.5424	$F(2, 21) = 14.63$	$p < 0.001$

Table 6: Average formant values of vowels used in the dialect regression model for judgments of gay-sounding talkers by California listeners. Those marked with asterisks are those formants that were used in the hierarchical regression model.

Talker set	Vowel	F1	F2
Gay-sounding talkers	/a/	640	1161*
Straight-sounding talkers	/a/	628	1186*
Gay-sounding talkers	/æ/	637*	1594
Straight-sounding talkers	/æ/	567*	1545

The final regression model used to fit CGJ used the reading parameters. These parameters from Table 7 accounted for 12% of the data by using two of the parameters. Gay-sounding talkers had slightly fewer pauses (M = 7.25, SD = 1.86) and had fewer disfluencies (M = 0.25, SD = 0.62) than straight-sounding talkers (pause number M = 8.16, SD = 3.19; disfluencies M = 0.42, SD = 0.9).

Table 7: Predicting CGJ with reading parameters.

Acoustic parameter	B	SE B	β
Number of pauses	0.102	0.0531	1.920
Disfluencies	0.2951	0.1811	1.629
Multiple R^2=0.1944	Adjusted R^2=0.1177	$F(2, 21)$=2.534	$p = 0.1033$

6.3.3 Perceived reading ability

Intelligibility parameters were entered into a model to predict perceived reading ability. This model accounted for 24% of the variance in listener judgments of reading ability. F2-F1 of /a/ and F2-F1 of /i/ were the intelligibility parameters that contributed significantly to the model illustrated in Table 8. Good readers had a smaller F1 range for /a/ (M = 518 Hz, SD = 89 Hz) than bad readers (M = 560 Hz, SD = 119 Hz), following the predictions that good readers more intelligible. Strangely, good readers had less clear /i/ vowels than bad readers; the difference between F1 and F2 of /i/ for good readers was smaller (M = 1579 Hz, SD = 176 Hz) than that of bad readers (M = 1694 Hz, SD = 167 Hz).

Table 8: Predicting reading ability with intelligibility parameters.

Acoustic parameter	B	SE B	β
/a/ F2-F1	0.0064	0.0027	2.411*
/i/ F2-F1	0.0024	0.0015	1.543
Multiple R^2=0.3048	Adjusted R^2=0.2386	$F(2, 21) = 4.605$	$p < 0.05$

Another model was built to determine how well the dialect parameters fit the reading ability data. Achieving a model with a better fit from this type of vowel information that is unrelated to intelligibility would indicate that reading judgments were based on dialect perception. Table 9 shows that only one dialect parameters was included in the model, accounting for 14% of the listener judgments of reading ability. Table 10 shows that good readers produce /a/ with lower F2 than bad readers.

Table 9: Predicting reading ability with dialect parameters.

Acoustic parameter	B	SE B	β
F2 /a/	0.0065	0.003	2.194*
Multiple R^2=0.1796	Adjusted R^2=0.1423	$F(1, 22) = 4.815$	$p < 0.05$

Table 10: Average formant values of vowels used in the dialect regression model for perceived reading ability. Those marked with asterisks are those formants that were used in the hierarchical regression model.

Talker set	Vowel	F1	F2
Good readers	/a/	638	1156*
Bad readers	/a/	630	1191*

A third model was attempted with the perceived reading ability judgments using the reading parameters. Good readers made fewer pauses (M = 7.4, SD = 2.3) with shorter mean durations (M = 452 ms, SD = 152 ms) than bad readers (pause number M = 8, SD = 2.9; pause duration M = 582 ms, SD = 157 ms). As demonstrated in Table 11, only two of the factors were included in the model. These reading parameters accounted for 27% of the variance.

Table 11: Predicting reading ability with reading parameters.

Acoustic parameter	B	SE B	β
Number of pauses	0.3022	0.1062	2.847**
Pause duration	0.0036	0.0017	2.175*
Multiple R^2=0.3307	Adjusted R^2=0.2669	$F(2, 21) = 5.188$	$p < 0.05$

6.3.4 Perceived listener presence

For the linear regressions, the dependent factor must be a continuum. The perceived listener presence task, however, was categorical. A listener presence continuum was constructed by counting the number of listener responses for each talker. The range of listener responses was from 3 to 15. The intelligibility parameters were entered into the model to predict judgments of listener presence. Table 12 shows how F0 variability, F2-F1 of /a/, and average vowel duration were included in the model. These values alone were able to account for 60% of listener responses. Talkers judged to be reading to someone had more F0 variability (M = 17 Hz, SD = 6 Hz) and longer mean vowel durations (M = 97 ms, SD = 8 ms) than those judged to be reading alone (F0 variability M = 14 Hz, SD = 4 Hz; vowel duration M = 95 ms, SD = 9 ms). Talkers perceived to have an audience had lower difference of F2-F1 of /a/ (M = 516 Hz, SD = 84 Hz) than talkers perceived to not have listeners (M = 567 Hz, SD = 109 Hz), in accordance with the predicted values.

Table 12: Predicting judgments of listener presence with intelligibility parameters.

Acoustic parameter	B	SE B	β
F0 variability	0.3206	0.0884	3.626**
/a/ F2-F1	-0.0147	0.0046	-3.201**
Vowel duration	0.1429	0.04873	2.932**
Multiple R^2=0.6535	Adjusted R^2=0.6039	F (3, 21) = 13.2	$p < 0.001$

The dialect parameters were entered into a second linear regression to investigate whether these acoustic parameters could account for more of the listener responses than the intelligibility parameters. Not surprisingly, F1 and F2 of /a/ were included in the model; the difference between these values accounted for a significant portion of responses in the intelligibility model as well. As shown in Table 13, these formant values alone accounted for 30% of the data. Talkers perceived to be reading to someone have higher F1 and lower F2 for /a/, as shown in Table 14.

The third linear model for the listener presence data was conducted with the reading pa-

Table 13: Predicting judgments of listener presence with dialect parameters.

Acoustic parameter	B	SE B	β
F1 /a/	0.0334	0.0138	2.415*
F2 /a/	-0.0151	0.0059	-2.583*
Multiple R^2=0.3569	Adjusted R^2=0.2956	$F(2, 21) = 5.826$	$p < 0.01$

Table 14: Average formant values of vowels used in the dialect regression model for perceived listener presence. Those marked with asterisks are those formants that were used in the hierarchical regression model.

Talker set	Vowel	F1	F2
Listener present	/a/	641*	1157*
No listener present	/a/	628*	1159*

rameters. In this model, the number of pauses, mean word duration, and F1 range accounted for 20% of listeners' responses, as seen in Table 15. Talkers with listeners made fewer pauses ($M = 7$, $SD = 2$) than talkers without listeners ($M = 8$, $SD = 3$). Talkers perceived to be reading alone had shorter mean word durations ($M = 291$ ms, $SD = 26$ ms) and smaller F1 range ($M = 287$ Hz, $SD = 49$ Hz) than talkers who were perceived to be reading to someone (word duration $M = 295$ ms, $SD = 24$ ms; F1 range $M = 298$ Hz, $SD = 60$ Hz).

Table 15: Predicting judgments of listener presence with reading parameters.

Acoustic parameter	B	SE B	β
Number of pauses	-0.692	0.2634	-2.627*
Word duration	0.0381	0.0262	1.455
F1 range	0.0148	0.0112	1.322
Multiple R^2=0.2997	Adjusted R^2=0.1947	$F(3, 20) = 2.853$	$p > 0.05$

6.3.5 Perceived listener identity

A continuum was constructed from the categorical 'child' or 'adult' responses by counting the number of 'child' responses given for a particular talker. Values ranged from 0 to 13 'child' responses. The child continuum was entered into the linear regression model with the intelligibility parameters. The results are shown in Table 16. F0 variation and vowel duration were the only acoustic parameters included in the model, but these two parameters accounted for 46% of the variation in child responses. Talkers perceived to be reading to a child had more F0 variability (M = 18 Hz, SD = 8 Hz) and longer vowel durations (M = 98 ms, SD = 8 ms) than those percieved to be reading to an adult (F0 variability M = 16 Hz, SD = 4 Hz; vowel duration M = 96 ms, SD = 9 ms).

Table 16: Predicting identity of listener with intelligibility parameters.

Acoustic parameter	B	SE B	β
F0 variation	0.4041	0.0919	4.399***
Vowel duration	0.1315	0.05383	2.443*
Multiple R^2=0.5066	Adjusted R^2=0.4618	F(2, 22)=11.3	$p < 0.001$

In the next model, the dialect parameters were entered into the linear regression model. Table 17 shows /au/ F1 and /a/ F2 accounting for nearly 23% of the 'child' responses. The vowels in Table 18 illustrate that talkers perceived to be reading to a child have higher F1 values for /au/ and lower F2 values for /a/.

Table 17: Predicting identity of listener with dialect parameters.

Factor	B	SE B	β
F1 /au/	0.0246	0.0105	2.358*
F2 /a/	-0.0099	0.0061	-1.623
Multiple R^2=0.2966	Adjusted R^2=0.2296	F(2, 21)=4.427	$p < 0.05$

All five potential predictors were selected in the stepwise estimation of the model with the reading parameters, as presented in Table 19. Together, these parameters account for

Table 18: Average formant values of vowels used in the dialect regression model for perceived listener identity. Those marked with asterisks are those formants that were used in the hierarchical regression model.

Talker set	Vowel	F1	F2
Child present	/a/	639	1145*
Adult present	/a/	641	1165*
Child present	/au/	654*	1285
Adult present	/au/	628*	1306

34% of the variance in listener responses. Talkers perceived to be reading to a child paused slightly more often (M = 7.5, SD = 2) and longer (M = 566 ms, SD = 188 ms) than those percieved to be reading to adults (pause number M = 7.1, SD = 2; pause duration M = 503 ms, SD = 128 ms). Talkers reading to a child also had longer mean word durations (M = 302 ms, SD = 23 ms) than talkers reading to an adult (M = 0.89 ms, SD = 0.44 ms). Those reading to an adult listener had more disfluenices (M = 0.89, SD = 0.44) than those reading to a child (M = 0.07, SD = 0.38). F1 range was larger for talkers perceived to be reading to a child (M = 307 Hz, SD = 61 Hz) than those perceived to be reading to an adult (M = 291 Hz, SD = 59 Hz).

Table 19: Predicting identity of listener with reading parameters.

Factor	B	SE B	β
Word duration	0.0448	0.0229	1.958 .
Disfluencies	-1.7839	0.757	-2.357*
Pause duration	0.0057	0.0035	1.623
F1 range	0.0161	0.0097	1.659
Number of pauses	-0.3237	0.2417	-1.339
Multiple R^2=0.4802	Adjusted R^2=0.3358	$F(5, 18)$=3.326	$p > 0.05$

6.4 Discussion

The stepwise linear regression models using the intelligibility, dialect, and reading parameters for the talker judgments (PSG, CGJ, reading ability, listener presence, and listener identity) demonstrate that particular sets of acoustic features are better able to predict listener judgments than others. Table 20 summarizes the amount of variance accounted for by each set of parameters for the various talker judgments presented above.

Table 20: Variance accounted for across talker judgments with the different sets of acoustic parameters. The number of variables used in each stepwise regression model are shown in parentheses following the percentage of variance accounted for.

	Intelligibility	Dialect	Reading
PSG	25% (2)	39% (3)	0
CGJ	24% (1)	54% (2)	12% (2)
Reading ability	24% (2)	14% (1)	27% (2)
Listener presence	60% (3)	30% (2)	19% (3)
Listener identification (child)	46% (2)	23% (2)	34% (5)

The dialect parameters performed the best on the PSG and CGJ values. This is in accordance with previous findings; Munson et al. (2006a) and Pierrehumbert et al. (2004) report evidence that speech patterns particular to gay communities in Minneapolis, Minnesota and Chicago, Illinois, respectively, are dialect specific. The ability of the intelligibility parameters to account for large amounts of variance for the two sets of sexual orientation judgments illustrates the relatedness of the social judgments of speech style and sexual orientation.

Perhaps unsurprisingly, reading ability was best predicted by the reading parameters, but there was not much difference between the reading and intelligibility parameters. The fact that the dialect parameters, however, also predicted a significant portion of the reading ability judgments suggests that judgments of reading ability are related to dialect perception. Listeners tend to rate talkers of less standard dialects as less intelligent (see discussion in Chapter 1 of Clopper (2004)); the California listeners who rated the Toronto talkers on their reading ability may have been judging talkers with stronger Toronto accents as poorer readers. Of the vowels included in the analysis here, Clopper and Pisoni (2004) report that the backness of /æ/ is an acoustic measurement that distinguishes Western dialects (California) from Northern dialects (Toronto).

Both listener presence and listener identification judgments were strongly predicted by the intelligibility parameters. Increased pitch variability and slower speech rate are characteristic features of child-directed speech; both of these acoustic features were significant factors in the linear regressions of listener presence and listener identity. For both of these talker judgments, vowel duration was a significant factor in the intelligibility parameters and word duration was a significant factor in the reading parameters. Charles-Luce (1997) found longer word durations for words produced with a flap with an underlying /d/ than with a /t/ in the production of semantically biased and neutral sentences when a listener was present in the room at the time of recording. In a perception experiment, Charles-Luce found that listeners were sensitive to the increased duration as evidenced by improved performance in a lexical decision task. Her result combined with the findings here related to increased vowel and word durations suggest that increased duration as a means of communicating meaning to a listener is an aspect of inter-talker communication to which listeners are sensitive.

7 General Discussion

The experiments reported in this paper illustrate how listener judgments of talker characteristics overlap with judgments of other talker characteristics made with independent groups of listeners and several subtle differences in social judgments made by different listener populations. When making judgments about gay-sounding speech, California and Canadian listeners pattered differently, yet in the same direction. Both sets of listeners were making use of dialectically significant differences in vowel productions, in accord with generalizations made by Munson et al. (2006a), Pierrehumbert et al. (2004). Sexual orientation as perceived by Californian listeners was also related to perceived listener identity; gay-sounding talkers were more likely to sound as if they had been reading to a child than an adult, indicating some sort of increased level of speech clarity. The overlap of these social judgments echo the conclusion in Munson et al. (2006a), cautioning that judgments of sexual orientation may not be isolated judgments of sexual orientation, but, rather, other social judgments that are abstractly related to stereotypes regarding sexual orientation. While these significant relationships emerged in the analysis, examining the acoustic cues listeners used to based their judgments revealed that listeners used different sets of acoustic parameters to make their judgments about sexual orientation, reading ability, and listener presence and identity.

The stepwise linear regression models used to determine the acoustic features used by listeners found that while such social judgments as sexual orientation and listener identity overlap as social categories, listeners made these judgments on different sets of acoustic parameters. Dialectal information was crucial to rating sexual orientation for both Canadian and Californian listeners, while reading ability was determined by factors related to reading performance and talker intelligibility. Acoustic cues revealing listener presence and listener identity dealt with the intelligibility parameters. Listeners used intelligibility cues to categorize those talkers who had been perceived to be reading to an individual, particularly a child. The listener identity judgment had overlapped with judgments of sexual orientation despite the fact that the intelligibility parameters barely accounted for 24% of the variance in listener responses to perceived sexual orientation from Californians while having had accounted for 46% of 'child' responses in the listener identity task.

The social indexical properties encoded in speech are not isolated judgments; rather, they are subjective judgments about abstract social categories that overlap and relate to other social categories. Remarkably, despite the relationship between these judgments, listeners are using different sets of acoustic features to determine their social decisions. This finding underscores how adept listeners are at deciphering indexical information in the acoustic signal and possessing abstract knowledge about the meaning of such information.

A Appendix: Rainbow passage stimuli

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 increases.

B Appendix: Post-experiment questionnaire

The two questions asked of participants upon the completion of the two tasks were:

1. Did it sound like the talkers were from a distinct dialectal region? Could you tell where they were from?
2. Do you have any judgments about the sexual orientation of the talkers? Did the majority of them sound gay or straight?

Responses to question (1):

1. Californian, "regular", slight British accent and one bum.
2. South, but not deep south. Some Canadian/Midwestern, one gruff-sounding guy who sounded like everyone's grandpa. Maybe one from South Africa.
3. No, but a few sounded like they were from the north east or Canada.
4. Minnesota region and there was a slight Irish accent on one.
5. 1 British, 1 Southern, 1 Indian accent.
6. Two or three sounded British, 1 MN/Canadian/Northern accent, rest American.
7. Yes, a couple did sound distinct.
8. Kept saying *about*. Michiganish, but not to the extent where they would say *pop*.
9. Fairly local sounding, plus a handful of Canadians.
10. Some had accents. One from England, mostly normal American speakers.
11. Some English and Southern accents. Some definitely not CA, but still US.
12. Midwestern /o/s. Some west-coasty vowels. There was one NPR voice.
13. Some were British.
14. *Complicated* was a cue to tell who was from the East coast and who was from the West coast. *About* said who was from the north.

15. Canada.

Responses to question (2):

1. Lisping, more gay-sounding people than in a normal population; more were gay-sounding than not.
2. The southern-lilter was gay-sounding. Gruff guy was straight. The heterosexual people were poorer readers.
3. Stressed parts seemed to cue gayness; those that stressed differently from the majority sounded gay.
4. A couple sounded homosexual, 4 or 5 of the talkers stood out.
5. They all sounded male, and a few sounded stereotypically homosexual.
6. One or two sounded really gay, that's about it. All were college educated except for a couple of them; one was in his teens. There were no black people, 1 smoker, 1 homeless guy. Most of them were in their 20s or 30s.
7. Nobody sounded too strongly gay, but didn't really think about it. Only recognized one voice as sounding gay.
8. A couple (1-2) sounded gay, but there could be more. Some sounded like pretentious students, some sounded like slackers.
9. A couple sounded like they might be gay based on lisping and inflectional patterns.
10. Two or three sounded gay. One sounded like a woman, the rest were men.
11. Some I think were gay. Most sounded straight.
12. A couple of them had stereotypical gay-speech.
13. Some sounded gay, some sounded old.
14. Didn't even think of that. No, could tell age, but not sexual orientation. Could tell that some were high school students.
15. Figured they were mostly straight.

References

Bradlow, Ann R., Gina M. Torretta, and David B. Pisoni. 1996. Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics. *Speech Communiation* 20:255–272.

Charles-Luce, Jan. 1997. Cognitive factors involved in preserving a phonemic contrast. *Language and Speech* 40:229–248.

Clopper, Cynthia G. 2004. Linguistic experience and the perceptual classification of dialect variation. Doctoral Dissertation, Indiana University.

Clopper, Cynthia G., Susannah V. Levi, and David B. Pisoni. 2006. Perceptual similarity of regional dialects of American English. *Journal of the Acoustical Society of America* 119:566–574.

Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32:111–140.

Gaudio, Rudolf P. 1994. Sounding gay: pitch properties in the speech of gay and straight men. *American Speech* 69:30–57.

Johnson, Keith, Elizabeth A. Strand, and Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27:359–384.

Labov, William. 1994. *Principles of linguistic change: Internal Factors*. Blackwell Publishing.

Linville, Sue Ellen. 1998. Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Pholia Phoniatrica et Logopaedica* 50:35–58.

Munson, Benjamin, and Molly Babel. to appear. Loose Lips and Silver Tongues, or, Projecting Sexual Orienation through Speech. *Language and Linguistics Compass* 1:??–??

Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe, and Aubrey R. White. 2006a. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics* 34:202–240.

Munson, Benjamin, Sarah V. Jefferson, and Elizabeth C. McDonald. 2006b. The influence of perceived sexual orientation on fricative identification. *Journal of the Acoustical Society of America* 119:2427–2437.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85.

Palmeri, Thomas J., Stephen D. Goldinger, and David B. Pisoni. 1993. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:309–328.

Perry, Theodore L., Ralph N. Ohde, and Daniel H. Ashmead. 2001. The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America* 109:2988–2998.

Pierrehumbert, Janet, Tessa Bent, Benjamin Munson, Ann R. Bradlow, and J. Michael Bailey. 2004. The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America* 116:1905–1908.

Sachs, Jacqueline, Philip Lieberman, and Donna Erickson. 1973. Anatomical and cultural determinants of male and female speech. In *Language Attitudes: Current Trends and Prospects*, ed. Roger W. Shuy and Ralph W. Fasold, 74–84. Georgetown University Press.

Schneider, Walter, Amy Eschman, and Anthony Zuccolotto. 2002. E-Prime: User's Guide, version 1.0. Psychology Software Tools.

Sjölander, K., and J. Beskow. 2005. Wavesurfer: an open source speech tool. Version 1.8.5, http://www.speech.kth.se/snack/.

Smyth, Ron, Greg Jacobs, and Henry Rogers. 2003. Male voices and perceived sexual orientation: an experimental and theoretical approach. *Language and Society* 32:329–350.

Strand, Elizabeth A., and Keith Johnson. 1996. Gradient and visual speaker normalization in the perception of fricatives. In *Natural language processing and speech technology, results of the 3rd KONVENS conference, Bielefeld, October 1996*, ed. Dafydd Gibbon, 14–26. Mouton de Gruyter.