



The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype

Sara Mack^a, Benjamin Munson^{b,*}

^a University of Minnesota, Department of Spanish and Portuguese, Folwell Hall, 9 Pleasant Street SE, Minneapolis, MN 55455, USA

^b University of Minnesota, Department of Speech-Language-Hearing Sciences, 115 Shevlin Hall, 164 Pillsbury Drive SE, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:

Received 1 February 2010

Received in revised form

28 September 2011

Accepted 1 October 2011

Available online 3 November 2011

ABSTRACT

Two experiments examined whether listeners associate frontally normal and misarticulated /s/ with gay-sounding voices, as is suggested by the popular culture stereotype that gay men "lisp". The first experiment showed that talkers were rated as younger-sounding and gayer-sounding when their speech included tokens with non-canonical variants of /s/ (i.e., a frontally misarticulated token of /s/, a dentalized /s/, or an /s/ produced with an especially high-frequency, compact spectrum). The second experiment showed that listeners recognize voices more quickly when they contain canonical /s/ variants than when they contain non-canonical /s/. Critically, these patterns were robust across different priming conditions in which listeners were presented with either a gay- or a heterosexual-sounding talker prior to the voice-recognition task. Together, these findings confirm experimentally that listeners make the association between non-canonical /s/ variants and male sexual orientation when asked to do so explicitly. However, though gay-sounding voices elicit longer reaction times in a voice-recognition task, we found no evidence that stereotypes about sexual orientation and /s/ production affect implicit processing of talkers' voices.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Sexual orientation and phonetic variation

This investigation examines how listeners associate phonetic variation with speaker sexual orientation, and in particular how listeners associate /s/ variation with sexual orientation. It is part of a broader program of research that is motivated in part by the existence of a widespread popular culture belief that talkers' sexual orientation can be identified accurately by simply listening to a person's voice. While there is a paucity of laboratory research studies on this topic, some data do suggest that people associate sexual orientation with distinctive pronunciation. For example, one of the strongest stereotypes about gay men documented in Madon's (1997) study of the content of stereotypes about gay men is the belief that they speak in a 'soft voice'. One of the attributes that gay men were seen to *reject* in that study is the propensity to speak in a 'deep voice'. Stereotypes of gay men's speech in Puerto Rican Spanish are documented in Mack (2010a), whose sociolinguistic interviews uncovered evidence of different stereotypes for different perceived sub-groupings of gay men, as well as stereotypes related to the realization of specific phonemes. Anecdotal evidence for

stereotypes about gay men's speech is revealed in a variety of popular culture sources, many of which suggest an association between sexual orientation and distinctive ways of speaking (see Russo, 1987, *inter alia*).

Previous studies of sexual orientation and speech documented that naive listeners are more likely to evaluate self-reported gay, lesbian, and bisexual talkers as more gay-sounding than self-reported heterosexual talkers when listening to audio-only samples of content neutral speech (Gaudio, 1994; Linville, 1998; Munson, 2007; Munson, McDonald, DeBoe, & White, 2006; Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004; Rendall, Vasey, & McKenzie, 2008; Smyth, Jacobs, & Rogers, 2003). However, no previous investigation of sexual orientation and speech has found a one-to-one correspondence between talkers' self-identified sexual orientation and specific acoustic measures or listeners' perceptual ratings. The speech production characteristics that listeners associate with sexual orientation are thus not an inevitable consequence of a person's stated sexual orientation. Rather, they are a set of phonetic practices that listeners associate with talkers' sexual orientation and which have been noted more often in the speech of people who self-report to be gay, lesbian, or bisexual people than in the speech of people who report being heterosexual.¹ These listener evaluations,

* Corresponding author. Tel.: +1 612 624 3322; fax: +1 612 624 7586.
E-mail address: munso005@umn.edu (B. Munson).

¹ We use these terms for convenience's sake. For a full discussion of the controversy surrounding naming and terminology, please see Kulick (2000).

or perceptions, of sexual orientation are the focus of this paper, and the term *perceived sexual orientation* (PSO) will be used throughout this paper to refer to listeners' judgments of the extent to which a talker's voice sounds gay, lesbian, bisexual, or heterosexual. Readers should keep in mind that these evaluations do not correlate perfectly with talkers' self-stated sexual orientation.

1.2. *Explicit measures of perceived sexual orientation*

Most previous research approaching this topic from listener perspectives has employed a seemingly straightforward model: listeners are presented with speech stimuli and then are asked to give ratings for different parameters. This method is a first step for establishing baseline understanding of the relationships between phonetic variables and social categories. For example, Gaudio (1994) investigated the stereotypical notion that self-identified gay speakers and self-identified straight speakers have differences in pitch range and variability, asking listeners to rate stimuli on four measures: straight/gay, effeminate/masculine, reserved/emotional, and affected/ordinary. He then investigated the relationship between the ratings of PSO and speakers' pitch range and variability, but did not find a significant relationship. Using somewhat similar methods, Levon (2006) examined the effect of sibilant duration and pitch range on the perception of sexuality and other personal characteristics. Listeners rated stimuli on ten affective scales. Stimuli were created from read speech recorded by a single male speaker; the original speech sample was rated by a group of listeners as both "extremely gay" and "extremely effeminate." The stimuli were digitally manipulated to yield four stimuli types using the combinations of wide and narrow pitch range and short and long sibilant duration. While the results did not reveal a significant relationship between perceptions of speaker sexuality and pitch range or sibilant duration, significant correlations between several of the affective scales highlighted the links between perception of sexuality and other personal characteristics. An expansion of the study (Levon, 2007) included stimuli from a second speaker whose speech sample was evaluated as "straight" and "masculine." Using the four combinations of pitch range and sibilant duration, listeners rated speakers on the same affective scales as in the previous study. In this case, there was evidence that pitch range and sibilant duration affected listener perceptions of sexuality, and that these effects varied as a function of the talker whose speech was manipulated and with the order of the speech samples in the experiment.

1.3. *Implicit measures of perceived sexual orientation*

While explicit reference to social characteristics is an important method for investigating individuals' associations between social categories and phonetic variation, this type of experimental design is not without weaknesses. To begin with, the design depends on labels that likely activate a set of social stereotypes that can influence participants' performance. For example, as Strand (2000) showed, the terms "masculine" or "feminine" are assumed to carry with them stereotypical ideals of how a prototypical masculine voice should sound and how a prototypical feminine voice should sound. Hence, listeners' responses may reflect the influence of stereotypical ideals in addition to listeners' more implicit knowledge of variation based on their own experiences. Conversely, labels could also introduce issues of social desirability that may cause participants to suppress stereotypes (Fazio & Olson, 2003). For example, on an individual level, listener responses may vary according to the persona they want to display to the researcher or for the study, or to correspond with their ideal persona (as politically liberal or conservative, for example).

On a societal level, the status and acceptance of social identities may intervene. For example, listeners native to a culture where there is less discrimination based on sexual orientation may feel more comfortable judging someone's voice as gay-sounding, while listeners native to a culture where homosexuality is more stigmatized may be less comfortable doing so.

Methods using explicit reference to social categories presuppose that listener perception functions in a very straightforward way: we hear a voice, and we make a judgment. However, it is increasingly clear that speech processing is impacted not only by acoustic cues, but also by the interaction of those cues with listener ideals (and ideas) of social identities. In the simplest case, this phenomenon is illustrated by studies showing that listeners recognize words more quickly when spoken by familiar talkers (Nygaard, Sommers, & Pisoni, 1995), or by talkers whose voices have a prototypical social attribute, such as being prototypically masculine or feminine (Strand, 2000). The phenomenon is also illustrated by studies suggesting that listeners calibrate their perception of sounds and words based on presumed attributes of the speakers they are listening to (Johnson, Strand, & D'Imperio, 1999; Munson, 2011; Strand & Johnson, 1996). This suggests that the notion that one hears a voice and then makes a judgment about social identity is, at best, an imprecise characterization of speech processing. The process is better characterized as an interactive process, with multi-directional flow of influence among judgments, acoustic cues, and social information. As such, it is beneficial to explore questions of social identity and phonology using methods that don't make explicit references to social categories alongside ones that do.

A growing body of research in social psychology and cognitive linguistics has documented and dealt with these methodological challenges by using priming techniques that circumvent this problem by suggesting social categories through experimental manipulations, rather than mentioning them explicitly. For example, Niedzielski (1999), Hay, Nolan, and Drager (2006) utilized visual stimuli to suggest to listeners what the age, sex, and socioeconomic status were for the talkers who produced the speech they were hearing, while Hay et al. (2006) used experimental materials to suggest national origin. Priming techniques target implicit processing to explore the strength of associations between variables. In general, priming experiments set up an expectation or judgment via a priming phase or section, and measure how a given variable is affected by that expectation in a test phase.

Implicit processing experiments have been successfully used in examining how tacit beliefs about gender, regional origin, social class, and age affect speech perception. For example, Strand (2000) showed that the gender-typicality of adults' voices affects the speed and accuracy with which listeners identify words presented in an acoustically degraded gated-word recognition task. That is, listeners' knowledge of variation socially stratified by gender affected performance on the task, even though gender was not mentioned explicitly.

1.4. *The gay 'lisp'*

The current investigation is inspired both by these studies and by the existence of a widespread popular culture belief that gay men speak with a lisp. This belief can be seen in discussions of gay men's speech styles in the mainstream media and in nontraditional media sources like internet postings. The experiments in this paper examine listener behavior in a series of perception tasks to validate experimentally whether listeners indeed associate lisping with gay-sounding speech. An immediate challenge to evaluating this hypothesis comes when considering what it actually means to lisp. Technical definitions of the term 'lisp' are elusive. One prominent early

textbook (Van Riper, 1937) described lisping as a variety of different misarticulations of both /s/ and /z/, including frontal misarticulation, lateral misarticulation, and audible nasal emission during fricative production. Contemporary texts in speech-language pathology have largely abandoned the use of the term *lisping*, opting instead to describe speech-production errors for /s/ and /z/ with terms that can be generalized across classes of sounds, such as *deletions*, *substitutions*, and *distortions*.

A starting point to understanding what the popularly accepted notion of lisping might encompass is to consider the acoustic characteristics of fricatives. The fricative /s/ has a relatively long interval of aperiodic noise, with a concentration of high-frequency energy. This sound is differentiated from the nonsibilant fricatives /f/ and /θ/ by being longer, having a more negatively skewed spectrum, and having higher amplitude relative to that of the following vowel. It is differentiated from the other voiceless sibilant fricative, /ʃ/, by having a higher peak frequency and more negatively skewed spectrum (Jongman, Wayland, & Wong, 2000). The acoustic characteristics of /s/ differ between men and women. Women produce this sound with a higher peak frequency and a more negatively skewed spectrum than men (Munson, McDonald et al., 2006). Sex differences in fricative production may be mediated by sociocultural variables as well. Stuart-Smith (2007) found that sex differences in /s/ in Glaswegian English were mediated by age and social class. While the effect of listeners' knowledge of socially stratified variation is observable in experiments such as these, knowledge of socially stratified variation affects performance in tasks that do not call explicit attention to that variation as well, as in the example of Strand (2000) mentioned above. Further evidence that sex differences in fricative characteristics are not reducible to sex differences in anatomy is provided by Fuchs and Toda (2009), whose results suggest that active articulatory manipulations, not simply anatomic differences, must be responsible for sex differences in fricative production.

A second element central to understanding what lisping might encompass is to consider the most common errors for /s/ in children acquiring English. Smit (1993a, 1993b) reports the results of a large-scale study on normal phonological development. The most common fricative substitution error for /s/ in children acquiring American English aged 2–9 years is a dentalized production, defined as “both dental and interdental variants as well as substitutions transcribed as [θ] and [ð]” (p. 540). If the description of gay men's speech as ‘lisped’ reflects a belief that gay men produce errors in /s/, then a reasonable first hypothesis would be that the stereotypically gay-sounding /s/ would be frontally misarticulated, given that this is the most common misarticulation of /s/ across the lifespan. Baum and McNutt (1990) found that /s/ tokens for children who produced frontally misarticulated /s/ nonetheless maintained a difference between target /s/ and /θ/ acoustically. Other studies have shown that even when children produce /s/ accurately, their tokens of this sound have a higher peak frequency than adults (Pentz, Gilbert, & Zawadzki, 1979).

Previous studies of gay men's speech styles have found that, while some gay men produce /s/ differently from their heterosexual peers, the acoustic characteristics of these sounds are not consistent with those of frontally misarticulated /s/. Indeed, Linville (1998), Munson, Jefferson, and McDonald (2006), and Rogers and Smyth (2003) all report acoustic characteristics of /s/ in self identified gay men and men whose speech is rated to sound gay that are in the *opposite* direction of what we would expect if these talkers were producing a frontally misarticulated /s/. These men are more likely to produce /s/ with a *higher* peak frequency (Linville, 1998; Rogers & Smyth, 2003) and a *more negatively skewed* spectrum (Munson, Jefferson et al., 2006) than

the /s/ of heterosexual and heterosexual-sounding men. Because none of these studies used population-based sampling methods, no definitive statement about differences between gay and heterosexual men in the incidence of lisping in the general population can be made.² However, the results of these studies using convenience samples suggest strongly there are no obvious group differences in the incidence of misarticulation of /s/. Munson, Jefferson et al. (2006) postulated that tokens of /s/ with a high peak frequency and an extremely negatively skewed spectrum were especially hyper-articulate variants. By producing a token of /s/ with these characteristics, listeners exaggerate the acoustic contrast between this sound and the sounds with which it is acoustically most similar, /f/ and /θ/. Moreover, these tokens of /s/ are more likely to be perceptible in the presence of most environmental noise, which typically has a concentration of energy in the lower frequencies. Put differently, the distinctive /s/ associated with gay speech styles appears to be a hyperarticulated /s/ rather than an error. This is consistent with Maniwa, Jongman, and Wade's (2009) finding that intentionally hyperarticulated /s/ tokens have a higher peak frequency than conversational-speech /s/ tokens. The finding that gay speech styles are associated with a hyperarticulated /s/ is consistent with the results of a second perception experiment in which Munson, Jefferson et al. (2006) found that men who were rated by one group of listeners to sound gay were rated by an independent group of listeners to speak more clearly than men who were rated to sound heterosexual. This finding has since been replicated with an independent group of talkers who had produced conversational and intentionally clear speech. Again, the clear-speech tokens were rated as gayer-sounding than the conversational-speech ones (Munson, Ferguson, & Connealy, 2009).

The findings detailed in the previous paragraphs suggest a dissociation between empirical findings and popular culture stereotypes about the relationship between /s/ production and sexual orientation. One purpose of this investigation, then, is to examine whether different types of /s/ affect implicit and explicit percepts of talkers' sexual orientation. The stimuli in this study include both natural productions of words beginning with /s/ and words created by editing off the natural /s/ and replacing it with different types of /s/ produced by a trained talker. This *matched guise* technique (Lambert, Hodgson, Gardner, & Fillenbaum, 1960) allows us to examine the influence of /s/ on perceived sexual orientation while keeping other parameters known to be associated with such judgments (i.e., vowels' formant frequencies) consistent. A finding that highly negatively skewed tokens of /s/ elicit judgments of gayness in speech would suggest that listeners are sensitive to the actual association between these variables in the population. An additional finding that misarticulated tokens of /s/ elicit judgments of gayness in speech would suggest that listeners' percepts are affected by the popular culture stereotype that gay men lisp.

1.5. Research questions

The first research question this study investigates is whether the acoustic characteristics of /s/ alone are sufficient to cue judgments of listeners' sexual orientation. The three tasks of the first experiment in this study examine listeners' associations between perceived sexual orientation and /s/ variation using methods that mention social categories explicitly. This work differs from other previous

² More recently, Van Borsel et al. (2009) claim to have found evidence for a higher incidence of misarticulations in /s/ in gay men than in heterosexual men in Belgium. However, a number of methodological issues limit how readily their findings can be generalized to the broader population of gay men. For a critique of that study, please see Munson (2010).

research on the topic (e.g., Levon, 2007; Munson, Jefferson et al., 2006) in two ways. Unlike Munson et al., we use stimuli both with natural productions of /s/, and stimuli created by combining tokens of /s/ produced by a trained phonetician with natural vocalic tokens. Unlike Levon (2007), we use a larger set of talkers ($n=16$), which allows us to examine whether the effect of /s/ variation on perceived sexual orientation interacts with phonetic characteristics of the men's voices that are evident in the vocalic portion of the stimuli.

A second purpose of this study, addressed in the second experiment, is to examine whether listeners associate talkers' sexual orientation with variation in /s/ in tasks that do not explicitly mention sexual orientation. By combining explicit measures of the association between phonetic variation and sexual orientation (i.e., tasks in which listeners are instructed to evaluate a talker's voice in terms of perceived sexual orientation or /s/ quality) with implicit measures (i.e., tasks in which neither /s/ quality nor perceived sexual orientation is mentioned overtly) we hope to gain a richer understanding of sexual orientation and speech, as well as a deeper understanding of the relationship between implicit and explicit perceptual measures of social categories more generally.

2. Experiment 1: Explicit measures

2.1. Introduction

The purpose of Experiment 1 was to examine the relationship between /s/ quality and perceived sexual orientation in explicit-measures tasks. It included three tasks completed by independent groups of listeners. The first task examined whether listeners' ratings of talkers' sexual orientation change when the quality of the /s/ in a monosyllabic /s/-vowel-consonant (sVC) words varies. The second task asked the inverse question, namely, whether judgments of the accuracy of /s/ are equivalent when the /s/ in the same sVC words is combined with VC sequences produced by talkers whose voices were previously rated to sound gay or heterosexual. The third task examines the specificity of the relationship between /s/ variation and judgments of the age of the talker who produced them.

Two types of stimuli were used in these tasks. The first of these were natural tokens of monosyllabic words beginning with /s/ produced by a variety of male speakers whose PSO had been measured and reported previously by Munson, McDonald et al. (2006). The second set of stimuli was created by editing off the naturally produced /s/ from tokens produced by men in Munson, Jefferson et al. (2006) and replacing them with /s/ tokens that had been produced by a trained phonetician. These variants of /s/ exemplified different variants of /s/, including both correctly articulated and misarticulated tokens. We refer to these /s/ tokens henceforth as the *matched-guise* /s/ tokens. As described below, one of the matched-guise /s/ productions had a high peak frequency and highly negative skewed spectrum, similar to the /s/ that the gay men in Munson, Jefferson et al. (2006) produced. Two others represented more-anterior productions of /s/, including both a frontally misarticulated /s/ and a token of /s/ produced with the tongue lamina approximating the back of the upper-front teeth. These two productions are referred to as misarticulated tokens, as they resemble the misarticulations often made by children. However, we note that the latter production is sometimes observed in normal adults' productions, and arguably serve a social-indexing function (e.g., Dart, 1991). If the influence of /s/ on listeners' perception of sexual orientation is guided by their experience listening to gay sounding speech, then they should judge talkers to sound gayer when a highly negatively skewed /s/ is present in the word than when an /s/ with typical acoustic characteristics is present. If judgments are guided by

the stereotype that gay men speak with a lisp, then listeners should judge a talker to sound gayer when a misarticulated /s/ is present in their production than when a typical /s/ is present.

The experiment also examined whether the influence of /s/ type on judgments of sexual orientation was equivalent for all talkers, or whether /s/ type interacts with talkers' perceived sexual orientation (PSO) as previously established by an independent group of listeners in Munson, Jefferson et al. (2006). This analysis allows us to gauge the relative influence of /s/ acoustics and other acoustic cues to sexual orientation. One possibility is that the influence of /s/ type on perceived sexual orientation would be so strong that evaluations made when listening to words containing two /s/ types (negatively skewed /s/ or misarticulated /s/) would be statistically equivalent for talkers previously evaluated as more stereotypically heterosexual sounding as well as those previously evaluated as stereotypically gay sounding. Another possibility is that listeners' PSO judgments will be more strongly affected by the acoustic cues present in the vocalic portion of the stimuli than by those in the fricatives; if this is the case, we would expect the PSO differences between the previously-established gay-sounding and heterosexual-sounding talkers to be similar in size across the different stimulus types. That is, this experiment collects measures of PSO and examines how they are affected by previously made measures of PSO. To minimize confusion, we refer to the measures in the current study as *measured PSO*, and those from previous studies as *previously established PSO*. As explained below, previously established PSO is treated as a categorical factor in the analysis of measured PSO.

The second task examined judgments of accuracy of /s/. This allows us to examine whether listeners perceive highly negatively spectrally skewed tokens of /s/ as less accurate than matched-guise /s/ tokens that mimicked those of a typical /s/, and similar in accuracy to frontally misarticulated /s/. Again, we explore whether listeners' judgments of the accuracy of matched-guise tokens of /s/ are mediated by the previously established PSO of the talker to which these sounds are appended. That is, we examine whether there is an interaction between previously established PSO and accuracy of different /s/ types. Such a finding would be predicted if listeners were judging the quality of /s/ relative to acoustic-perceptual characteristics of the rime portions of the stimuli. Such perceptual biases have been found previously in studies of the influence of talker characteristics on the perception of /s/-/ʃ/ continua (Munson, Jefferson et al., 2006; Strand & Johnson, 1996) and /s/-/θ/ (Munson, 2011; Munson & Coyne, 2010) continua. If listeners' behavior in the current experiment were to be influenced tacitly by the belief that gay-sounding men are more likely to misarticulate /s/ than heterosexual-sounding men, then we might predict that the /s/ accuracy judgments would differ as a function of the previously established PSO of the person to which the /s/ is appended.

The third experimental task examined the specificity of the relationship between /s/ quality and perceived talker age. Previous studies by Munson, McDonald et al. (2006) and Babel and Johnson (2006) found that judgments of PSO were significantly correlated with judgments of other perceptual parameters, such as speech clarity (Munson, Jefferson et al., 2006) and reading fluency (Babel, 2007). As discussed earlier, dental and frontal variants of /s/ are relatively common errors in children. Moreover, children produce /s/ with a higher peak frequency than adults (Pentz et al., 1979). Thus, all three of the variants that are potentially associated with actual gay speech or with stereotypically gay speech are, in a way, associated with children's speech. It is well established that listeners can gauge a talker's age from content-neutral speech, but that these judgments are not perfectly correlated with actual age (Linville & Fisher, 1985). Moreover, Nagao (2006) showed that there are culturally specific ways of conveying and perceiving age, suggesting that age perception

is not a simple result of perception of the acoustic consequences of the effect of aging on the speech-production mechanism. Listeners' behavior in the experiments in this paper could have stemmed from a broad tendency to label 'younger-sounding' variants of /s/ as less accurate and gayer. The conjecture that a sociophonetic variant would be associated with younger talkers is not unprecedented. As Foulkes and Docherty (2000) discuss, labiodentalized variants of /r/ in the United Kingdom are becoming increasingly more common in different socioeconomic strata and geographic regions. These variants are similar to the /r/ misarticulations that are made by many young children; hence, speakers who produce them are frequently characterized as sounding 'infantile' or having a habitual speech error. It is possible that the unifying feature of the variants that fall under the 'gay lisp' label is that listeners tacitly associate all of them with the speech of younger people. This may reflect a broader stereotype that gay culture is heavily youth oriented (see Harry & DeVall (1978), for a review and critique of this hypothesis). To examine this possibility, a third perception task was conducted, examining naïve listeners' percepts of the age of the same sVC stimuli used in tasks 1 and 2.

2.2. Methods

2.2.1. Participants

Participants for all three tasks of Experiment 1 were recruited from the University of Minnesota community using recruitment materials that did not make reference to sexual orientation. We refer to them as *listeners* henceforth, to distinguish them from the participants in earlier studies whose productions were used as stimuli in this experiment. All of the listeners were native speakers of North American English, and none of the listeners reported a history of speech, language, or hearing impairment. No attempt was made to stratify the listener group by sexual orientation or familiarity with GLB speech styles. The listeners are presumed to represent an average cohort of middle-SES traditional college-aged adults. Listeners' regional dialect was not controlled, beyond the requirement for participation that all listeners be native speakers of a North American dialect of English. The listeners were from the University of Minnesota student body, which largely comprises students from Minnesota, Wisconsin, North Dakota, South Dakota, and Iowa. All of the talkers who contributed stimuli for the experiment were from the same area. Thus, it is likely that, as a group, the listeners' regional dialect was similar to that of the talkers. However, we cannot rule out the possibility that mismatches between listeners' and talkers' dialects affected some listeners' responses.

2.2.1.1. Listeners, task 1, perceived sexual orientation task. Fifteen listeners participated in this experiment. Their median age was 22 years (Range=[19–32]).

2.2.1.2. Listeners, task 2, perception of /s/ accuracy task. Fifteen listeners participated in this experiment. Their median age was 23 years (Range=[19–45]).

2.2.1.3. Listeners, task 3, perceived age task. Sixteen listeners participated in this experiment. Their median age was 22 years (Range=[19–37]).

There was no overlap among the listener groups for tasks 1–3.

2.2.2. Stimuli

As mentioned previously, the same stimuli were used in all three tasks. Two types of stimuli were used in this experiment: unedited digitized tokens, and tokens made by concatenating naturally produced bases with digitized tokens of /s/ produced by a trained speaker. Stimuli consisted of tokens of the words *sack*,

sad, and *soon*. These words were chosen because they contained two vowels, /æ/ and /u/, whose acoustic characteristics Munson, McDonald et al. (2006) found to differ between gay and heterosexual men, and to be predictive of listeners' judgments of perceived sexual orientation. All three words were likely to be familiar to an average listener, as shown by their having average familiarity ratings of 7 on a 7-point scale, as reported by Pisoni, Nusbaum, Luce, and Slowiack (1985).

2.2.2.1. Natural unedited tokens. The naturally produced tokens were taken from the corpus of single-word productions analyzed in Munson, Jefferson et al. (2006). Each talker produced three tokens each of *sack*, *sad*, and *soon*. These were recorded with an AKG C420 head-mounted microphone attached to a Marantz CDR330 CD Recorder. These had a 22.05 kHz sampling rate with 16-bit quantization, and had been processed through an 11.025 kHz anti-aliasing filter. Each of these was used as a stimulus in the perception experiment in unedited form. For each word, one token was chosen randomly. The /s/ was edited off of this token, and it was used as the base for the stimuli containing the matched-guise fricatives, as described below.

Talkers were selected from the larger set of 22 male talkers in Munson, Jefferson et al. (2006). Because the second goal of this experiment was to examine whether the influence of /s/ type on measured PSO judgments in the current experiment interacted with the previously established PSO, we divided the talkers into gay- and heterosexual-sounding groups. This division was based on the ratings that were made in Munson, McDonald et al. (2006). These ratings were made on a five-point scale, where 5 indicated *definitely sounds GLB*, 1 indicated *definitely sounds heterosexual*, and 3 indicated *sounds neither GLB nor heterosexual*. Ratings were made by 40 naïve listeners. Each listener rated each talker four times. One of these ratings was based on three words that contained front vowels and no sibilant fricatives, one was based on three words that contained front vowels and sibilant fricatives, one was based on three words with back vowels and no sibilant fricatives, and one was based on three words with back vowels and sibilant fricatives.

For this experiment, we divided talkers into equal-sized gay- and heterosexual-sounding groups based on listeners' ratings averaged across the two conditions in which they were not presented with any tokens of /s/. The average ratings for the 22 talkers on this measure were unimodally distributed, with a cluster of talkers around the mean value of 2.72. Because we wanted well-defined groups of gay- and heterosexual-sounding talkers, we eliminated the six talkers in the middle of the distribution. We then defined the gay-sounding talkers as the men with the eight highest ratings, and the heterosexual-sounding talkers as the men with the eight lowest ratings. The mean rating for gay-sounding talkers was 3.47 (SD=0.42, range=[2.85–4.3]). This group included six self-identified gay men and two self-identified heterosexual men. The mean rating for heterosexual-sounding talkers was 2.25 (SD=0.26, range=[1.88–2.63]). This group included seven self-identified heterosexual men and one self-identified gay man. A Mann–Whitney *U* test showed the difference in ratings between the two groups to be significant, (Wilcoxon $W=36$, $z=-3.36$, $p=0.001$). It is important to emphasize that these ratings were made by listeners in conditions where /s/ was not present in the stimuli. Thus, we can be confident that phonetic cues to talkers' PSO would be present in other phonemes, such as in the formant frequencies and voicing-source characteristics of vowels. The average age for the heterosexual-sounding talkers was 23.1 years (SD=2.8), and the average for the gay-sounding talkers was 24.1 years (SD=4.1). These did not differ significantly in a Mann–Whitney *U* Test (Mann–Whitney $U=28.5$, $z=-0.373$, $p>0.05$).

Table 1
Acoustic characteristics of fricatives in the naturally produced tokens.

Group		Duration (ms)	Intensity (dB IL)	Fricative/vowel relative intensity ^a (dB)	Center of gravity (Hz)	Skewness
Gay-sounding talkers	Mean	196	69.4	12.3	6370	−0.68
	(SD)	(32)	(4)	(3.8)	(788)	(1)
Heterosexual-sounding talkers	Mean	200	70.4	11.9	6440	−0.53
	(SD)	(46)	(4)	(4.8)	(712)	(0.9)

^a The RMS amplitude for the vowel minus the RMS amplitude for the fricative.

Table 2
Acoustic characteristics of the vocalic portions of the naturally produced tokens.

Word	Group	Mean F0 (ERB)		Duration (ms)		F1 (Bark)		F2 (Bark)	
		Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
<i>Sack</i>	Gay	3.04	(0.25)	220	(20)	6.20	(1.33)	11.32	(0.82)
	Het.	3.04	(0.35)	240	(50)	4.89	(1.44)	10.44	(1.19)
<i>Sad</i>	Gay	3.13	(0.23)	268	(44)	6.03	(0.96)	11.56	(0.71)
	Het.	2.95	(0.32)	240	(22)	5.64	(0.58)	11.08	(0.59)
<i>Soon</i>	Gay	3.33	(0.35)	283	(32)	4.22	(1.88)	10.27	(1.72)
	Het.	2.92	(0.38)	255	(27)	5.19	(1.84)	10.77	(1.60)

The acoustic characteristics of the fricatives in the natural, unedited tokens are shown in Table 1. This table reports the center of gravity and skewness (calculated over the frequency range 0–11,025 Hz for a 40 ms interval of frication noise centered at fricative midpoint, using the methods in Forrest, Weismer, Milenkovic, and Dougall, 1988), duration, RMS amplitude of the fricative, and the relative amplitude between the vowel and fricative, as this variable has been shown to affect fricative identification (Hedrick & Ohde, 1993). As Table 1 shows, the groups' fricative productions differed primarily in spectral skewness. This is similar to how self-identified gay and heterosexual men's fricatives differed in Munson, McDonald et al.'s (2006) earlier study, which analyzed a larger corpus of fricatives.

Table 2 shows the average duration of the vocalic portions of the stimuli, as well as average F1, F2, and f0, taken at vowel midpoint. F1 and F2 are shown in the Bark scale (Zwicker & Ternhardt, 1980); f0 is shown in the ERB scale (Hermes & van Gestel, 1991). These show that the stimuli differed in a manner similar to how the larger group of gay- and heterosexual-sounding men in Munson, Jefferson et al. (2006) differed. Readers should note the data in Tables 1 and 2 are a subset of the larger set of tokens analyzed in Munson, Jefferson et al. (2006).

2.2.2.2. Matched-guise fricatives. Four fricatives were produced by the second author, a trained phonetician. These were produced in the word *sad*. One of these tokens was intended to mimic the average acoustic characteristics of the fricatives produced by the entire group of 16 talkers (henceforth *neutral /s/*). The second was produced with the tongue tip contacting the back of the lower incisors and the tongue lamina contacting the back of the upper incisors (henceforth *dental /s/*). The third was produced with the tongue tip and blade protruding between the upper and lower incisors (henceforth *frontal /s/*). The dental and frontal */s/* tokens were intended to be representative of common production errors for */s/*, such as those described in Smit (1993a, 1993b), which naïve listeners might identify as consistent with the label 'lisp'. In the remainder of this paper, they are referred to collectively as *misarticulated /s/* tokens. The last matched-guise fricative was

Table 3
Acoustic characteristics of matched-guise fricatives.

<i>/s/</i> Type ^a	Intensity (dB IL)	Duration (ms)	Center of gravity (Hz) ^b	Skewness ^b
Neutral <i>/s/</i>	65.0	208	6677	−0.31
Negatively skewed <i>/s/</i>	60.0	242	8470	−2.84
Dentalized <i>/s/</i>	57.5	226	6312	−1.08
Frontally misarticulated <i>/s/</i>	55.0	236	3068	0.60

^a See text for details.

^b Based on the method in Forrest, Weismer, Milenkovic, and Dougall (1988).

intended to be an exaggerated version of the */s/* that listeners in Munson, Jefferson et al. (2006) identified as gay-sounding. This */s/* had a very high peak frequency and an extremely negatively skewed spectrum (henceforth *negatively skewed /s/*).

The acoustic characteristics of these fricatives are shown in Table 3. These fricatives were normalized to the amplitudes listed in Table 3. They were then concatenated with the naturally produced bases, which were normalized to have an RMS amplitude of 75 dB SPL. The relative amplitudes were chosen to reflect the relative amplitudes in the token of *sad* from which they had been excised. The stimuli were edited at points of zero crossing, eliminating transients in the concatenated stimuli. The authors judged that the resulting stimuli did not have any artifacts (i.e., transients, or abrupt changes in amplitude) that would cue listeners that they were listening to acoustically modified stimuli. A small panel of listeners judged them to sound highly natural. By cross-splicing these stimuli with naturally produced vocalic portions, we were able to examine how the PSO of the vocalic portion influences how the talkers' sexuality and age are perceived, and how the quality of */s/* is perceived.

2.2.3. Procedures, task 1, perceived sexual orientation task

The experimental task was generated and run using the E-Prime software package (Schneider, Eschman, & Zuccolotto, 2002). On each trial, listeners were presented with a single spoken word concurrent with an orthographic transcription of the word on a video monitor located 0° azimuth from the listener's head. Listeners responded by pressing buttons on a keypad. Responses ranged from 1 (which corresponded to *this man's voice is definitely heterosexual sounding*) to 9 (*this man's voice is definitely gay sounding*). The instructions from this experiment emphasized that listeners should rate the perceived sexual orientation of the talker's voice, not whether they could guess the talker's actual sexual orientation. This scale differed from the five-point scale used in Munson, McDonald et al. (2006). Technical limitations prevented the use of a scale greater than five in that study. The nine-point scale in this study was intended to increase the range of responses relative to those in the earlier study.

The order of items was fully randomized across listeners. Each experimental block was preceded by five practice words, using talkers who were not in the experiment. The total set of stimuli included 3 word types \times 5 /s/ types \times 16 talkers (including 8 talkers who had previously been rated as sounding gay, and 8 who had previously been rated as sounding heterosexual), for a total of 240 stimuli. Listeners were compensated \$5.00 for participating.

2.2.4. Procedures, task 2, perceived /s/ accuracy task

Procedures for Task 2 were similar to the perceived sexual orientation task. For this task, responses ranged from 1 (which corresponded to a *very unclear and inaccurate production of “s”*) to 9 (a *very clear and accurate production of “s”*). The order of items was fully randomized across listeners. Matched-guise and unmodified stimuli were presented together in a single block. The experimental block was preceded by five practice words, using talkers who were not in the experiment. None of the instructions made reference to sexual orientation, nor did they use the word *lisp*. The decision not to draw listeners' attention to sexual orientation was based on the finding that listeners may attenuate stereotypes about GLB people when they are made explicitly aware of sexual orientation (e.g., Marsh, Cook, & Hicks, 2006). The entire experimental task took approximately 20 min. Listeners were compensated \$5.00 for participating.

2.2.5. Procedures, task 3, perceived age task

Procedures were similar to the two other tasks. In this task, listeners were told that they were listening to stimuli that had been produced by a variety of talkers who ranged in age from 12 to 60 years, and that they should guess the talker's age to the nearest year. Responses were logged automatically. The order of items was fully randomized across listeners. Matched-guise and unmodified stimuli were presented together in a single block. The experimental block was preceded by five practice words, using talkers who were not in the experiment. None of the instructions made reference to sexual orientation, nor did they call attention to the variation in /s/. The entire task took approximately 20 min, and was included in a session in which they completed a second, unrelated experiment. They were compensated \$10.00 for participating in the entire experimental protocol.

2.3. Analysis

2.3.1. Tasks one and two, perceived sexual orientation and perceived /s/ accuracy tasks

Responses occurring greater than 5 s after the word were presumed to reflect listener inattention, and were excluded from further analysis; these accounted for less than 0.5% of the total responses. For each listener, mean PSO ratings and mean ratings for /s/ were calculated separately for each /s/ type (natural, neutral, dental, frontal, and negatively skewed) for gay- and heterosexual-sounding talkers.

2.3.2. Task 3, perceived age task

The analysis of data from the third task was somewhat different. One of the goals of Experiment 1 was to examine ratings for individual stimuli averaged across listeners. Hence, we examined each listener's data were standardized relative to the overall mean and standard deviation of the ages they provided. For each listener, mean standardized ages for /s/ were calculated separately for each /s/ type (natural, neutral, dental, frontal, and negatively skewed) for gay- and heterosexual-sounding talkers.

2.4. Results

2.4.1. Task 1, perceived sexual orientation task

Fig. 1 shows the average measured PSO ratings for previously established gay- and heterosexual-sounding talkers for the five fricative types. Average perceived sexual orientation ratings were submitted to a two-factor within-subjects analysis of variance (ANOVA). The two factors were fricative type (5 levels: natural, neutral, dental, negatively skewed, and frontal), and previously established PSO (2 levels: gay-sounding and heterosexual-sounding). A significant main effect of fricative type was found, $F[4, 56]=39.03$, $p < 0.001$, $\eta_p^2=0.74$. Post-hoc Bonferroni-corrected paired comparisons showed significant group differences for all pairs except two. First, negatively skewed /s/ and frontal /s/ did not differ significantly. Second, the difference between negatively skewed /s/ and dental /s/ did not achieve statistical significance at the Bonferroni-corrected α level, but did approach significance ($p=0.078$). As Fig. 1 shows, measured PSO was highest (i.e., gayest-sounding) for stimuli with frontal and negatively skewed /s/, and lowest (i.e., most-heterosexual sounding) for natural and natural /s/ stimuli. Dental /s/ stimuli were intermediate.

Unremarkably, a significant main effect of previously established PSO on measured PSO was found, $F[1, 14]=33.37$, $p < 0.001$, $\eta_p^2=0.70$. Talkers who were previously established to sound gay elicited higher measured PSO values. This is not surprising, given that listeners were defined as 'gay-sounding' or 'heterosexual-sounding' based on how they were rated in Munson, Jefferson et al. (2006). This finding shows that the 9-point perceived sexual orientation ratings in the current experiment replicated the 5-point ratings made in Munson, Jefferson et al. (2006), and that the set of tokens in this experiment was representative of the larger group of stimuli examined by Munson et al.

More importantly, the interaction between previously established PSO and fricative type was significant, $F[4, 56]=4.71$, $p=0.004$, $\eta_p^2=0.25$. This interaction can be seen by comparing the bar heights in Fig. 1. As this figure shows, the differences between the two previously established PSO groups were smaller for the stimuli containing frontal /s/, dental /s/, or negatively skewed /s/ than for the naturally produced words, or the words containing the matched-guise neutral /s/. Post-hoc *t*-tests with a Bonferroni-corrected α level of 0.01 (0.05/5) showed significant differences between gay- and heterosexual-sounding talkers for natural /s/, neutral /s/, and frontal

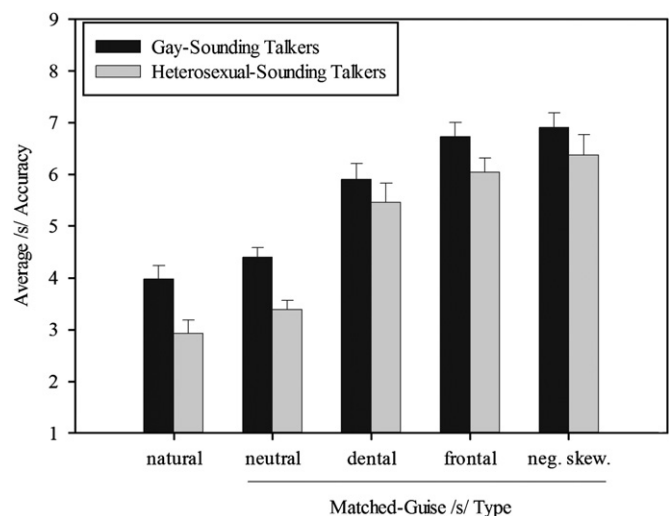


Fig. 1. Average talker perceived sexual orientation ratings for the four matched-guise fricatives used in Experiment 1 and the natural unedited stimuli, separated by previously established PSO. Error bars represent one standard error of measurement.

/s/ ($t[14]=4.05-7.09$), all p 's ≤ 0.01). The differences for dental /s/ and negatively skewed /s/ did not achieve statistical significance at the Bonferroni-corrected α level, but did approach significance ($t[14]=2.31$, $p=0.04$ and $t[14]=2.78$, $p=0.02$, respectively). This finding suggests that the influence of /s/ on measured PSO is not equivalent for talkers whose previously established PSO differs. As Fig. 1 shows, the effect of previously established PSO on measured PSO were smaller for dental /s/, frontal /s/, and negatively skewed /s/.

An additional ANOVA was conducted with just the four matched-guise fricatives, and the additional factor of word (3 levels: *sack*, *sad*, *soon*). Recall that the matched-guise fricative was taken from an /æ/ context. The purpose of this analysis was to examine whether the mismatch between the matched-guise context and the phonetic context of the words—particularly the /u/ context of *soon*—affected measured PSO. There was no significant main effect of the word factor, nor did any of the interactions involving this factor reach significance at the $\alpha < 0.05$ level. However, the same pattern of significance in the ANOVA without word remained in this ANOVA.

2.4.2. Task 2, perceived /s/ accuracy task

Ratings average averaged across the three words are shown in Fig. 2, which presents average ratings for the natural fricatives, and for the four matched-guise fricatives appended to gay-sounding and heterosexual-sounding voices. As this Figure shows, natural /s/ stimuli were perceived to be the most accurate, followed by neutral /s/, followed by dental /s/, followed by frontal and negatively skewed /s/. These mean ratings were submitted to a two-factor fully within-subjects analysis of variance (ANOVA). The within-subject factors were fricative type (5 levels: natural, neutral, dental, frontal, and negatively skewed) and talker's previously established PSO (2 levels: gay sounding and heterosexual sounding). The main effect of fricative type was significant, $F[4, 56]=122.9$, $p < 0.001$, $\eta_p^2=0.90$. Post-hoc Bonferroni-corrected pairwise comparisons revealed significant difference in all pairs except two. Neutral /s/ and natural /s/ did not differ, and the frontally misarticulated /s/ and the negatively skewed /s/ did not differ. This finding supports the notion that there is a gradient relationship between previously established PSO and perceived /s/ accuracy. They also support the conjecture that listeners perceive negatively skewed /s/ to be as inaccurate as frontally misarticulated /s/. There was no significant effect of previously established PSO, $F[1, 14]=0.05$, $p > 0.05$. Average ratings for gay- and heterosexual-sounding men

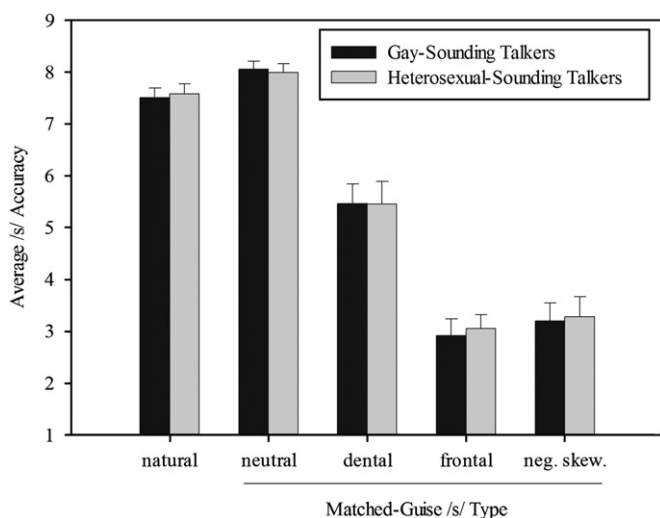


Fig. 2. Average /s/ accuracy ratings for the four matched-guise fricatives used in Experiment 1 and the natural unedited stimuli, separated by previously established PSO. Error bars represent one standard error of measurement.

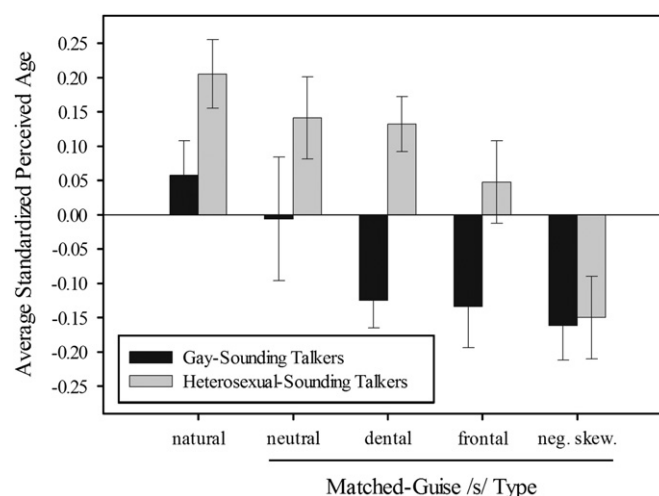


Fig. 3. Average perceived age ratings for the four matched-guise fricatives used in Experiment 1 and the natural unedited stimuli, separated by previously established PSO. Error bars represent one standard error of measurement.

were almost identical. Contrary to expectations, the interaction between fricative and previously established PSO was not significant, $F[1, 14]=0.71$, $p > 0.05$. As Fig. 3 shows, the bar heights for gay- and heterosexual-sounding men were nearly identical across the five fricative types.

A second ANOVA was run on just the four matched-guise fricatives, with word as a factor. In addition to there being a significant effect of fricative type, word was a significant main effect ($F[2, 28]=5.66$, $p=0.009$, $\eta_p^2=0.29$), and interacted significantly with fricative type ($F[6, 84]=3.498$, $p=0.004$, $\eta_p^2=0.20$). This interaction was explored by examining the influence of Word on accuracy ratings for the four different fricative types in four separate ANOVAs. For the neutral /s/, there was no effect of Word on accuracy ratings, $F[2, 28]=1.98$, $p > 0.05$. This indicates that the choice to use a matched-guise /s/ from an /æ/ context did not affect the acceptability of tokens when that was paired with the /u/ context in the *soon* tokens. In contrast, there was a significant main effect of Word on ratings for the dental /s/, the frontal /s/, and the negatively skewed /s/ (respectively: $F[2, 28]=5.277$, $p=0.011$, $\eta_p^2=0.27$; $F[2, 28]=3.378$, $p=0.048$, $\eta_p^2=0.19$; $F[2, 28]=6.057$, $p=0.007$, $\eta_p^2=0.30$). For the dental /s/, this interaction occurred because the ratings for *soon* were significantly lower (i.e., less accurate) than those for *sack* or *sad*. For the frontal /s/, this occurred because the ratings for *sack* were significantly lower than those for *sad*; those for *soon* were intermediate and not significantly different from either *sack* or *sad*. For the negatively skewed /s/, this occurred because the ratings for *soon* were significantly less accurate than *sad*; those for *sack* were intermediate and not significantly different from either *sad* or *soon*. This pattern does not lend itself to an explanation based solely on the mismatch between the phonetic context of the matched-guise fricatives and the phonetic context of the stimulus, which would suggest that ratings would consistently be lower for the word *soon*. Rather, this may be due to something iconically gay sounding about the word *sad*.

2.4.3. Task 3, perceived age task

A two-factor fully within-subjects ANOVA examined the influence of /s/ type (5 levels: natural, neutral, dental, negatively skewed, and frontal) and previously established PSO (2 levels: gay-sounding, heterosexual-sounding) on normalized age judgments. A significant main effect of previously established PSO was found, $F[1, 15]=5.813$, $p=0.029$, $\eta_p^2=0.28$. There was also a significant main effect of

/s/ type, $F[2.6, 39.7]=4.695$, $p=0.003$, $\eta_p^2=0.24$. Finally, these factors interacted significantly, $F[4, 60]=3.476$, $p=0.042$, $\eta_p^2=0.11$.³ The main effects and their interaction can be seen by comparing the bar heights in Fig. 3, which shows listeners' average normalized perceived ages for the five /s/ types, with averages for the gay- and heterosexual-sounding men plotted separately. The significant main effect of age arose because the gay-sounding men were rated to sound younger than the heterosexual-sounding men. The main effect of /s/ type on ratings arose principally because the talkers' naturally produced tokens were rated as older-sounding than the tokens containing negatively skewed /s/ were. Though the ratings for the three other /s/ types were intermediate between the other conditions, they did not differ statistically significantly from the naturally produced or negatively skewed /s/, or from each other.

The interaction between previously established PSO and /s/ type was explored two ways. First, the two groups were compared separately for each of the five stimulus types using post-hoc paired-sample *t*-tests with a Bonferroni-corrected α level of 0.01 (0.05/5). For three of the five stimulus types (natural, dental, and frontal), the heterosexual-sounding talkers were rated to sound significantly older than the gay-sounding talkers. Second, two one-factor within-subjects ANOVAs were calculated to assess the influence of /s/ type on ratings of gay- and heterosexual-sounding talkers separately. In both ANOVAs, the main effect of /s/ type was statistically significant. However, the pattern of post-hoc Bonferroni-corrected comparisons differed for the two groups. For the heterosexual-sounding talkers, perceived ages for natural, neutral, and dental /s/ words differed significantly from ratings for negatively skewed /s/ words. Moreover, natural /s/ words differed significantly from dental /s/ words, which differed significantly from negatively skewed /s/ words. No other comparisons were significant. In contrast, for the gay-sounding talkers dental, frontal, and negatively skewed /s/ words elicited younger ratings than natural /s/ words. No other comparisons were significant. Again, a second ANOVA was run on the measures for the four matched-guise stimuli, with the additional factor of Word. There was no significant main effect of word, nor did it interact with /s/ type or PSO. As with task 1, the same significant main effects and interactions as in the ANOVA without Word were present.

2.4.4. Relation between perceived sexual orientation, perceived /s/ accuracy measures, and perceived age

The final analyses compared the measures across the three tasks. For talker/fricative combination, average standardized perceived ages from Task 3, perceived clarity from Task 2, and PSO from Task 1, for a total of 80 data points. The associations among these measures are shown in Table 4. This table shows partial correlations between two measures when the third is controlled. As this table shows, the strongest pairwise partial correlation was between measured PSO and perceived accuracy.

2.5. Discussion

The results of this experiment confirm that the two groups of talkers were well-separated in their previously established PSO. The judgments of natural, unedited stimuli made by the listeners in this experiment were similar to those in the earlier Munson, Jefferson et al. (2006) study, even though they were based on a smaller set of stimuli. Moreover, statistically significant group differences in measured PSO were found to be present for two of the four matched-guise conditions, with the

Table 4

Partial correlations of the measures of PSO from Task 1, perceived clarity from Task 2, and average standardized perceived ages from Task 3. Each correlation controls for the third variable.

Word	Accuracy × Age	Accuracy × PSO	Age × PSO
Sack	−0.21 ^(*)	−0.83 ^{***}	−0.28 [*]
Sad	−0.06	−0.78 ^{***}	−0.30 [*]
Soon	−0.30 ^{***}	−0.83 ^{***}	−0.51 ^{***}

(*) $p < 0.10$.

* $p < 0.05$.

*** $p < 0.01$.

differences for the other two conditions approaching statistical significance.

More importantly, the results show an effect of fricative type on measured PSO, as well as an interaction between fricative type and previously established PSO. As mentioned previously, the four fricative types were neutral /s/, which was intended to mimic the average acoustic characteristics of the fricatives produced by the entire group of 16 talkers; dental /s/ and frontal /s/ (labeled collectively as misarticulated /s/ tokens), which were intended to be representative of common production errors for /s/ that naïve listeners might identify as consistent with the label 'lisp'; and negatively skewed /s/, an exaggerated version of the /s/ with a high peak frequency and an extremely negatively skewed spectrum that listeners in Munson, Jefferson et al. (2006) identified as gay-sounding. Talkers were rated to sound gayer when their productions contained misarticulated tokens of /s/ or negatively skewed /s/ than when they contained neutral /s/, or were natural and unedited. This tendency was so strong that it attenuated the PSO differences between the men who had been rated previously to sound gay and those previously rated to sound heterosexual. This finding suggests that listeners' association between the /s/ quality and perceived sexual orientation reflect both the social stereotype that gay men produce /s/ less accurately than heterosexual men, and the experience of hearing gay-sounding men producing /s/ with higher peak frequencies than heterosexual-sounding men.

The results of Task 2 show that negatively skewed matched-guise /s/ is perceived to be as inaccurate and as unclear as frontal /s/. Given this sound's acoustic characteristics, as well as the previous finding that men who produce this type of /s/ are rated as speaking clearly, we would have predicted that it would be rated as even more clear and accurate than natural /s/ or the neutral matched-guise /s/. Indeed, this finding extended to naturally produced tokens of /s/. Moreover, the ratings in Task 2 verified that the tokens of dental and frontal /s/ were indeed perceived as less accurate and less clear than natural tokens and the neutral matched-guise /s/. Finally, the results of Task 2 suggest that previously established PSO does not affect their ratings of /s/ accuracy. Listeners did not rate the accuracy of /s/ differently for natural, unedited words produced by men who were previously established to sound gay, as compared to ones who had been rated to sound heterosexual. Moreover, identical tokens of correct and misarticulated /s/ were not rated differently when appended to gay- and heterosexual-sounding men's voices.

As hypothesized, different variants of /s/ systematically affected judgments of perceived age. Listeners judged words containing negatively skewed /s/ to have been produced by younger talkers than words containing natural tokens. In general, the magnitude of this effect was smaller than the effect that /s/ had on the PSO judgments elicited in Task 1, and the interaction between /s/ type and previously established PSO was qualitatively different from that seen in Task 1. Specifically, the interaction suggested that only frontal /s/ and negatively skewed /s/ cue

³ An ANOVA on raw ages yielded similar results.

younger-sounding speech in both gay- and heterosexual-sounding talkers, while dental /s/ was sufficient to cue younger-sounding speech in talkers who had previously been established to sound gay. This suggests that age judgments are related to a combination of characteristics of /s/ and characteristics of the vocal portion of the stimuli that cued the previously established PSO judgments.

3. Experiment 2: Implicit measures

3.1. Introduction

The results of Experiment 1 show that individuals evaluate a talker's sexual orientation differently when they are presented with stimuli that contain different types of /s/ sounds. The purpose of the second experiment is to see whether these associations affect performance on a task that does not call attention explicitly to sexual orientation or any other social variables. The method used in this experiment is a primed voice recognition task. As detailed below, in this task listeners were first presented with tokens of words that did not contain instances of /s/ (priming phase). As mentioned in the earlier discussion of priming techniques, it is assumed that exposure to this stimulus activates a set of expectations in the mind of the listener. Much research in priming spoken language has examined semantic and phonological associations between lexical items, and showed that the presentation of a word automatically activates not just the strict semantic meaning of the word, but also related concepts and related forms (Fazio, 2001). More recently, this method has been used to investigate the automatic activation of sociocultural identity and sociocultural value judgments, such as racial and gender stereotypes (Bargh, 2006). In the current study, our goal was to activate a set of associations about the talker's social identity (including sexual orientation) in the priming phase.

In this experiment, the priming phase is followed by a brief distracter phase. After this, listeners were presented with a single word containing /s/ and were asked to judge whether it was the same talker that was presented previously. The speed of listeners' responses in the correct trials (i.e., the trials where the prime speaker and the test speaker were the same, and where the listener responded as such) is presumed to reflect the strength of the association between social information and pronunciation variation. Longer response times are assumed to occur in cases of mismatch, and faster responses are assumed to occur in the case of congruence. In other words, listeners will respond more quickly to voices that contain expected pronunciation variation than to ones that do not, even when this pronunciation was never explicitly presented in the study stimuli. That is, talkers' voices that are markedly gay-sounding should activate a set of stereotypes about gay men, including the stereotype that they produce frontal /s/. Hence, the person should recognize this voice more quickly if presented with a token that exemplifies this stereotype (i.e., a frontal /s/) than if presented with one that does not (i.e., a correctly articulated /s/).

In this experiment, test tokens were primed by productions by the same talkers. These were the same set of talkers from Experiment 1. The PSO of their voices had been measured in multiple previous studies. If gay-sounding voices activate a set of expectations that talkers should lisp, then listeners who are primed with a gay-sounding voice should respond more quickly if the test phase tokens contained a frontally misarticulated /s/ or a dental /s/ than a correctly articulated /s/. Similarly, if straight-sounding talkers activate no such stereotype, then listeners who are primed with a straight-sounding voice should respond less quickly if the test phase tokens contained a frontally misarticulated /s/ than a correctly articulated /s/.

3.2. Methods

3.2.1. Listeners

Twenty listeners participated in the experiment. Their median age was 29.1 years. They were recruited via word of mouth and fliers posted on the University of Minnesota campus. The recruiting materials did not mention sexual orientation, but referred to the experiment as a voice recognition task. The listeners were recruited from the same community as the listeners for Experiment 1, and thus share broadly the same characteristics as the listeners in Experiment 1.

3.2.2. Stimuli

Stimuli for the study were taken from the same two sets of stimuli used in Experiment 1. Stimuli for the priming phase of Experiment 2 included six different monosyllabic words for each of the 16 talkers whose productions were used in Experiment 1. The first set of three included front vowels and no sibilant fricatives: *bell*, *fade*, and *path*. The other three included back-round vowels and no sibilant fricatives: *hoop*, *note*, and *tooth*. Including all of the stimuli from Experiment 1 in the test phase of Experiment 2 would have made it prohibitively long. Hence, stimuli for the test phase of Experiment 2 were comprised of only two of the words from Experiment 2, *soon* and *sad*. Moreover, including all five fricative types from Experiment 1 also would have rendered Experiment 2 prohibitively long. Because the focus of this experiment was the stereotype that gay men produce frontally misarticulated tokens, and because of the need to keep the experiment length manageable, only four of the /s/ types were included: frontal /s/, dental /s/, neutral /s/, and natural /s/ (the unedited token as produced by the individual talker).

3.2.3. Procedures

Sixteen of the 20 listeners completed the task in individual carrels in a large computer lab designed to support social science research, and four listeners completed the task in a soundproof booth at a different location. The listeners were paid \$10.00 for their participation. The experiment was designed and carried out using E-Prime experiment management software. All listeners wore headphones and gave their responses via a serial response button box and numeric keyboard.

Each experimental trial was made up of a three-part sequence: the priming phase, the distracter task, and the test phase (shown in Fig. 4). In the priming phase, listeners heard the six priming words, always in the same order: *bell*, *fade*, *path*, *hoop*, *note*, and *tooth*. An orthographic display of the words appeared on the computer monitor at the same time as the word was heard. Again, the goal of this phase was to establish an idea of the social identity of the speaker in the mind of the listener; the words were the same for each talker. After the priming phase, listeners completed a distracter task of a short mental math problem. After the distracter task was completed, the listeners were instructed to press the middle button as soon as the screen turned from red to green; this action activated a short tone that corresponded to the start of the stimulus recording for the test phase. The stimulus for the test phase was auditory only and was

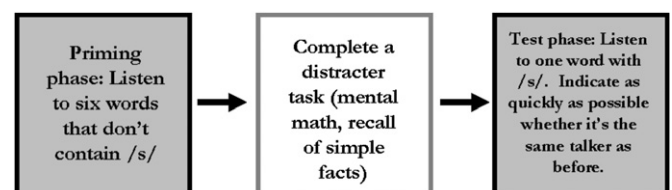


Fig. 4. Sequence of experimental trial.

not accompanied by an orthographic display of the word. The listeners then pressed the leftmost button if the talker was the same talker as in the prime phase, and pressed the rightmost button if the talker was a different talker than in the prime phase. Paper stickers on the button box marked the appropriate buttons as *same*, *middle*, and *different*. The listeners were instructed to press the appropriate button as quickly as possible without compromising accuracy, and these speeded response times (RTs) were logged. There were 10 practice items presented before the main experiment. The order of the experimental stimuli was fully randomized. The experiment protocol took listener approximately 60 min to complete. At no time was sexual orientation mentioned in the oral or written communication with the subjects.

One third of the trials were filler trials in which the prime talker and test talkers did not match, and included productions by talkers other than the sixteen talkers of interest. These were other talkers from the cohort of 44 men and women described by Munson, Jefferson et al. (2006). We reasoned that some of the cases in which the prime talker and the test talker matched would be responded to in error; hence, using two-thirds of trials where the voices matched and one-third where they did not would result in approximately equal numbers of 'same' and 'different' responses for individual listeners. There were 128 trials in which the prime talker and the test talker matched (2 levels of PSO \times 8 talkers per group \times 2 target words \times 4 fricative types in the target words). There were 64 trials in which the test talker and the target talker did not match, for a total of 192 trials.

3.3. Results

The primary focus of this experiment was whether RTs to stimuli with different types of /s/ varied as a function of the talker's perceived sexual orientation, as established in the test phase. Prior to analyzing RTs, we examined the influence of PSO proportion of correct responses for the trials in which the prime talker and the test talker matched. A two-factor (2 PSO \times 4 fricative type) within-factor ANOVA was performed with these as the dependent measure. A significant effect of fricative type was found, $F[3, 57]=13.684$, $p < 0.001$, $\eta_p^2=0.419$. There was no significant main effect of PSO, nor did the two factors interact. Bonferroni-corrected pairwise differences were significant except for two: the difference between natural /s/ and neutral /s/, and the difference between natural /s/ and dental /s/ (though this difference did approach significance, $p=0.076$). Response accuracy for 19 of the 20 listeners was very high, and resulted in no more than 10% of data-points lost. One listener had very low accuracy. That listener's RTs were eliminated from further analyses.

The next analysis examined RTs for correct responses to the trials in which the prime talker and test talker matched. RTs that were longer than 3000 ms removed from the analysis. Individual listeners' individual RTs were then standardized relative to that talker's overall mean and standard deviation. Responses greater than 3 standard deviations away from the individual subjects' mean RT were removed. For each listener, mean response times were calculated separately for each /s/ type by PSO combination.

A two-way within-subjects ANOVA was used to analyze the effect of the type of /s/ (four levels: natural, neutral, frontal, and dental) and previously established perceived sexual orientation (two levels: PSO gay and PSO non-gay) on response times. Mauchly's test of sphericity was violated for /s/ type, so a Greenhouse–Geisser correction was applied for that factor. There was a significant main effect of the s-type on the response times, $F[2.2, 37.2]=13.3$, $p < 0.001$, $\eta_p^2=0.44$. The longest responses were for the frontal and dental /s/ variants. There was also a significant main effect of perceived sexual orientation on

response times, $F[1, 14]=6.7$, $p=0.019$, $\eta_p^2=0.28$. Stimuli produced by talkers who had been previously rated as heterosexual sounding showed the quickest responses, while stimuli from those who had been previously rated as gay sounding had longer responses. There was not a significant interaction effect between perceived sexual orientation and the type of /s/ in the test phase $F[3, 51] < 1$, $p > 0.05$. Bonferroni-corrected post-hoc tests showed that RTs for the natural and neutral words did not differ from one another, nor did the RTs for the dental and frontal variants. These are shown graphically in Fig. 5, which also shows the percent of accurate responses in the eight conditions. As with the data from Experiment 1, a second ANOVA to examine whether the RTs were affected by the mismatch between the /æ/ context in which the matched-guise /s/ was produced and the phonetic context present in the stimulus word. This was a three-factor ANOVA with Word (2 levels: sad, soon), matched-guise /s/ type (3 levels: neutral /s/, dental /s/, frontal /s/), and previously established PSO (2 levels: gay-sounding, heterosexual-sounding). There was not a significant main effect of word on RTs, nor did word interact with any of the other factors.

The duration of the stimuli was not equivalent across the eight stimulus categories. To determine the possible influence that this might have had on the findings, a second ANOVA was performed using stimulus duration-normalized response times. To derive these, a series of regression analyses were conducted predicting individual subjects' response times from the duration of the stimuli. The standardized residuals from this regression were used as the dependent measure in this ANOVA, averaged by fricative type and previously established PSO. The results of this analysis paralleled those of the ANOVA on raw RTs: there was a significant main effect of previously established PSO, $F[1, 18]=7.423$, $p=0.014$, $\eta_{partial}^2=0.29$, and of fricative type, $F[3, 54]=1.777$, $p > 0.10$. These did not interact. The main effect of PSO occurred because the RTs for gay-sounding talkers were longer than predicted based on the duration of the stimuli.

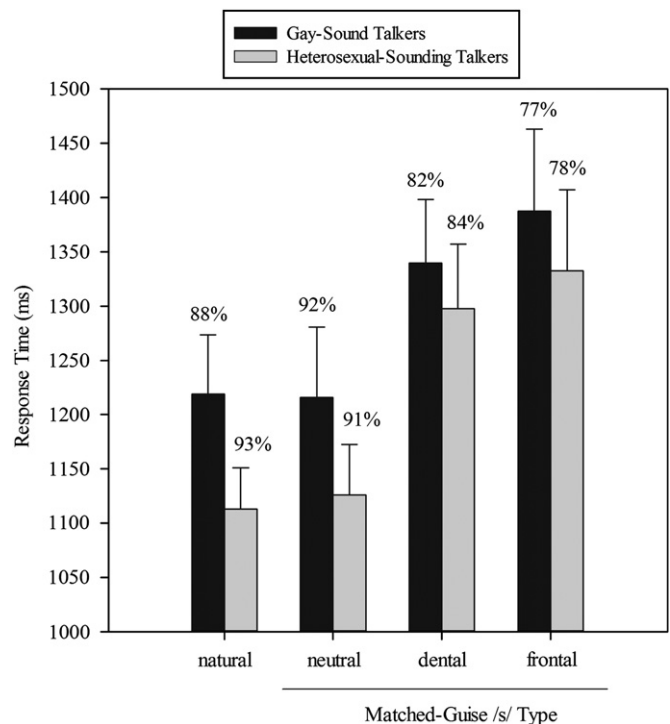


Fig. 5. Average voice-recognition response times for the three matched-guise stimuli used in Experiment 2, and the natural unedited stimuli, separated by previously established PSO. Error bars represent one standard error of measurement. Percentages refer to the percentage of correct trials per condition.

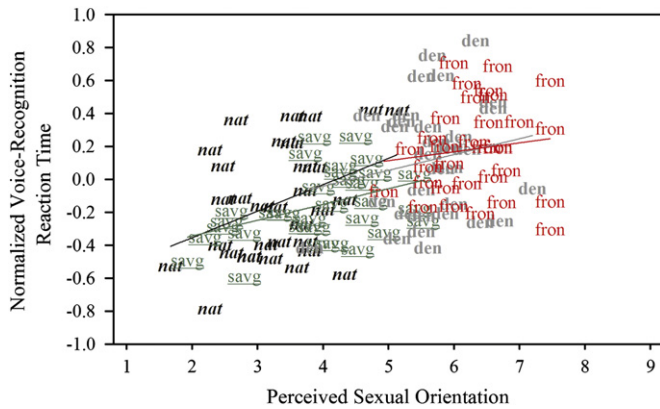


Fig. 6. Scatterplot of normalized voice-recognition response times by average PSO ratings from task 1 of Experiment 1. Natural, unedited tokens (**nat**, bold italic), tokens with neutral /s/ (savg, underlined), dental /s/ (**den**, bold), and frontal /s/ (fron, plain font) are plotted separately. Linear regression lines are fitted separately for each stimulus type.

Fig. 6 is a scatterplot of measured PSO values (from Task 1 of Experiment 1) by average normalized voice-recognition time. Each stimulus type is plotted separately, and separate regression lines are fitted for each stimulus type. As this figure shows, there is overall a relationship between these variables. Moreover, the regression lines are steeper for the natural, unedited stimuli and the stimuli containing the neutral /s/ than for the dental /s/ and frontal /s/ stimuli. The final analysis of these data examined whether the relationship between perceived sexual orientation and response times held when the association between these variables and the other two measures from Experiment 1—perceived age and perceived accuracy—were controlled statistically. This was examined by performing a series of partial correlations. The first partial correlation examined the partial correlation between measured PSO and response times when stimulus duration, perceived age, and perceived accuracy/clarity was controlled. This correlation was significant, $r=0.234$, $p=0.009$. The partial correlation between response time and perceived age with perceived accuracy, measured PSO, and stimulus duration controlled for was also significant, $r=0.229$, $p=0.010$. In contrast, the partial correlation between response time and perceived accuracy when the other three variables were controlled was not significant, $r=-0.110$, $p=0.221$. The result of this analysis shows RTs are associated equally strongly with PSO and perceived age, and not with perceived accuracy or stimulus duration.

3.4. Discussion

The results of the study show that the quality of the fricative in an sVC word influences listeners' RTs when recognizing the talker who produced that word. Moreover, it showed an influence of the PSO of a talker's voice on response times. Longer responses were observed for the dental and frontal /s/ variants (the /s/ types that were intended to be representative of common production errors for /s/ that naïve listeners might identify as consistent with the label 'lisp') as compared to the natural /s/ (the unedited token as produced by the individual talker) and neutral /s/ (intended to mimic the average acoustic characteristics of the fricatives produced by the entire group of 16 talkers). Longer responses were also observed for talkers who had been rated previously as gay sounding as compared to those who had been rated previously as straight sounding. However, these factors did not interact; the influence of talkers' perceived sexual orientation on listeners' response times was similar for test stimuli with different /s/ variants, and vice versa. We failed to find the predicted pattern of

short response times for gay-sounding voices combined with frontal or dental /s/, and heterosexual-sounding voices combined with neutral /s/.

A second ANOVA using normalized response times verified that stimuli duration does not drive the differences in response times among different categories of stimuli. In other words, differences in response times are not an artifact of stimuli durations. However, it did show that the effect of fricative type on RTs could not be disentangled from the duration of the stimuli.

The results presented here do not confirm the hypothesis that there is an observable relationship between /s/ variation and perceived sexual orientation in a task that does not call explicit attention to sexual orientation. Specifically, the RTs were not faster when there was a congruence between the PSO of the talker in the prime phase and the variant of /s/ that is associated with stereotypes about that PSO. Rather, the RTs were predicted by independently made measures of how gay-sounding the talkers were. The observable effect of perceived sexual orientation on response time suggests that social information influences low-level speech perception tasks to a certain degree, and that social information plays some role in the storage of voice information in memory, in that gay-sounding voices elicited longer overall voice-recognition times than heterosexual-sounding ones. We interpret this as an effect of how canonically heterosexual the talkers sound. Speech tokens produced by someone whose voice is easy to parse as heterosexual are responded to more quickly than ones that are less canonical. This interpretation is based on the well-documented finding that tokens whose acoustic characteristics are not easy to parse as members of a regular phonological category are responded to more slowly than canonical ones (e.g., Fox, 1984). Related to this, the results can also be interpreted in terms of Exemplar Theory. The assumption of the theory would be that straight-sounding voices are more strongly represented in the mental lexicon due to their relative prevalence as compared to gay-sounding ones; therefore, we would predict that the congruence between the acoustic character of stored heterosexual voices and those heterosexual voices presented in the study would facilitate processing.

4. General discussion

4.1. Summary

The two experiments in this study investigated whether listeners' judgments of a talker's sexual orientation are influenced by the acoustic and perceptual characteristics of their productions of /s/. The results of Experiment 1, which included three explicit-measures tasks, confirm that there is an association between the quality of /s/ and judgments of sexual orientation. First, consistent with popular culture stereotypes, listeners rated men as gayer sounding when dental and frontal tokens of /s/ (tokens representative of common production errors for /s/ that naïve listeners might identify as consistent with the label 'lisp') were combined with vowel-consonant sequences they had produced. Consistent with previous empirical findings, listeners rated talkers to sound gay when their vowel-consonant productions were combined with tokens of /s/ with a high peak frequency and a highly negatively skewed spectrum. Previous research (Munson, Jefferson et al., 2006) had established that this type of /s/ is associated with gay- and gay-sounding men's speech. The effect of /s/ type on ratings of perceived sexual orientation was stronger for men who had previously been rated as heterosexual-sounding. That is, misarticulated tokens of /s/ (i.e., frontal tokens of /s/ and dentalized tokens of /s/) and negatively skewed /s/ were a more powerful perceptual cue to sexual orientation than were the

acoustic characteristics of the vocalic portion of stimuli, which presumably cued the previously established judgments. The second task of Experiment 1 found that listener judge dental and frontal variants of /s/ to be less accurate than tokens of /s/ that matched the averaged acoustic characteristics of /s/ for the sample of talkers being studied. In addition, listeners judged that extremely negatively skewed /s/—a variant of /s/ that mimicked the acoustic characteristics of /s/ that listeners in a previous study (Munson, Jefferson et al., 2006) had associated with gay-sounding speech—as just as unclear and inaccurate as frontally misarticulated /s/. The negatively skewed /s/ used in Tasks 1 and 2 was intended to be an exaggerated version of the tokens of /s/ that listeners in Munson, Jefferson et al. (2006) associated with gay-sounding men's voices. An objective analysis of the acoustic characteristics of negatively skewed /s/ would lead us to believe that it should have been rated as extremely clear and accurate. In contrast, an objective acoustic analysis of frontal /s/ supports listeners' judgments that it was not produced accurately. The acoustic characteristics of this variant of /s/ make it less differentiable from other voiceless fricatives, particularly /θ/. Finally, Task 3 of Experiment 1 showed that men who had previously been rated to sound gay were perceived by naïve listeners to sound younger than men who were previously rated as heterosexual-sounding. Moreover, stimuli containing frontal, negatively skewed, and dental /s/ elicited younger age judgments than natural, unmodified stimuli.

Together, the results of Experiment 1 shed light on the nature of social stereotypes about sexual orientation and speech. Task 1 provides a controlled, laboratory-based validation of the stereotype that gay-sounding speech is associated with distinctive productions of /s/. Indeed, it is the first study to show that /s/ quality contributes independently to judgments of sexual orientation in speech. Previous research on this topic (e.g., Linville, 1998; Munson, McDonald et al., 2006) used stimuli in which /s/ characteristics may have co-varied with other parameters that may have cued judgments of sexual orientation. The results of Tasks 2 and 3 suggest that the types of /s/ that elicit judgments of gay-sounding speech are also perceived as being less accurate, and that they cause talkers to be judged to sound younger than in stimuli containing natural /s/ tokens.

The results of Experiment 2, the implicit measures task, found that response times in a voice-recognition task using a subset of stimuli from Experiment one were affected by the quality of /s/. Faster voice-recognition times were found for stimuli with the natural and neutral /s/ variants and slower for stimuli with the dental and frontal /s/ variants. This effect appeared to be related to the duration of the stimuli. The results also show a relationship between response times and the speaker's perceived sexual orientation, with faster responses for less-gay sounding talkers and slower responses to gayer-sounding talkers. This effect continued to be present when stimulus duration was controlled statistically. However, the lack of a significant interaction between PSO and /s/ type no relationship between /s/ variation and perceptions of sexual orientation observable in a task that does not explicitly call attention to sexual orientation. One factor that limits the interpretation of this finding is that, in the absence of an interaction between PSO and fricative type, we have no clear evidence that any social information was indeed primed. Rather, the response times appeared to be affected solely by the PSO of the test-phase stimulus. Expanding the methods used to induce priming is a key challenge to future work on this topic.

4.2. Implications and future research

The results of the experiments in this investigation set the stage for future studies of on the perception of sexual orientation

in speech. The first is to better understand why listeners responded similarly to the tokens of /s/ that were similar to those in gay men's speech, the negatively skewed /s/, and those that are not attested in gay men's speech but which are a component of stereotypes about gay men, the dental and frontal /s/. One possibility is that listeners' perception is guided by both experience perceiving gay men's speech, and by adherence to stereotypes about how gay men's speech should sound. Another possibility that we cannot rule out from these data is that listeners used a simple strategy of labeling talkers whose productions of /s/ were not canonical as gay. This may have been due to the arguably artificial experimental task of judging sexuality from content-neutral samples of speech, or because the listeners did not have much experience interacting with gay men. This possibility could be tested by including stimuli in future research that is neither part of the gay-speech stereotype nor attested in gay men's productions, such as the lower peak-frequency /s/ tokens produced by young working-class Glaswegian girls in Stuart-Smith (2007). If stimuli with those tokens of /s/ were rated as gay sounding, then we could conclude that the data in this paper reflect listener uncertainty. The potential role of listener uncertainty on the findings of could also be shown if listeners who differed in their experience interacting with gay people also differed in the influence of /s/ type on their ratings of sexual orientation. Such a finding would be consistent with Carahly's (2000) study of gay, lesbian, and heterosexual listeners' perception of men and women's sexual orientation. Carahly found that gay and lesbian listeners perceived women's sexual orientation better than did heterosexual listeners, presumably because they had more experience interacting with both lesbian and heterosexual women.

It is also possible that systematic differences among listeners assess listeners' beliefs about sociophonetic variation and about gay men resulted in the lack of a predicted interaction between /s/ quality and previously established PSO on voice-recognition response times. As in many previous studies (e.g., Gaudio, 1994; Munson, Jefferson et al., 2006) that group-level judgments of sexual orientation were correlated with group-level judgments of other attributes about talkers. It is possible that individual listeners in Experiment 2 experienced the activation of different aspects of that correlated stereotype in the priming phase. While most explicit measures tasks, including those in Experiment 1 of this protocol, pool listener groups and analyze group-wise trends in evaluations, at least one study (Mack, 2010b) documents that smaller clusters of listeners, rather than the listener group as a whole, show similar patterns of evaluations of perceived sexual orientation. Furthermore, Mack shows that within the cluster groups, there is variation according to which acoustic cues are significantly correlated to evaluations of PSO. In terms of Experiment 2 of the current study, this individual variation may have introduced a level of variability in activation (and therefore response times) that obscured an observable combined effect of /s/ type and PSO.

A related question concerns the measurement of PSO. This and other studies on PSO (e.g., Gaudio, 1994; Levon, 2006) have elicited judgments of PSO using interval-level scales. The implicit assumption behind this choice is that PSO is gradient, rather than categorical. Other studies (Linville, 1998; Smyth et al., 2003) have examined gradience in PSO by eliciting categorical gay/straight judgments from a cohort of listeners, and using the proportion of listeners who judge a talker to sound gay as a gradient measure of the talker's PSO. The presumption that PSO is gradient follows from contemporary research on language and gender, which has argued that gender, rather than sex, is a better construct for understanding linguistic variation among and between groups of men and women (see discussion in Milroy & Gordon, 2003).

Gender differs from biological sex in that it is seen as a learned, socially constructed, culturally specific, gradient variable, which presumably accounts for the variation among men and women in the extent to which their speech conforms to the group norms for their biological sex. Correspondingly, we presume that PSO is gradient because it reflects the extent to which people engage in the communities of practice in which distinctively GLB and heterosexual speech styles are constructed. The presumption that PSO is gradient, however, remains a presumption, which should be subjected to rigorous empirical tests. Moreover, the use of equally appearing interval scales implies that listeners can divide PSO into equal intervals. If, however, PSO is continuous but not divisible into equal-sized units, then direct magnitude estimates of PSO would be required to obtain a reliable estimate of this variable. It is possible to test the appropriateness of each of these measurement scales; indeed, the speech–language pathology literature is rife with examples of such work (e.g., Eadie & Doyle, 2002). Our ongoing research examines the psychometric properties of PSO, in an attempt to obtain more reliable quantification of this parameter.

More generally, work on this topic should integrate methods and findings from experimental sociolinguistics to arrive at a more-complete picture of the nature of the ‘meaning’ associated with gay speech styles is, and how it affects implicit processing. The past decade has seen numerous studies examining the nature of sociolinguistic meaning. These studies have shown fluid associations between many variables and potential meanings. For example, Campbell-Kibler (2007) showed that the interpretations of the *-ing/-in* alternation vary widely as a function of other things that they believe about the talker who produced the sample. Podesva, Jamsu, Callier, and Heitman (2008) showed a similarly fluid relationship between medial consonant release and attributes of national politicians. In that study, the influence of the phonological variable was strongly mediated by prior beliefs about the politicians. Eckert (2008) formalizes this using the notion of an ‘indexical field’. Eckert defined indexical fields as “constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable.” A better understanding of PSO could also help interpret differences among listeners of the type described in the previous paragraph. It may be that different stereotypes about PSO—i.e., stereotypes about gay men in casual interactions with peers, or about professional gay men—are activated in implicit perception tasks like our Experiment 2. Each of these might lead to a different set of expectations about the quality of the /s/ that the talker is likely to produce. More generally, though, the indexical fields hypothesis provides a formalism for rigorously testing the nature of sociolinguistic meaning and its potential influence on low-level speech processing.

Acknowledgments

The order of authors in this paper is alphabetical, and reflects equal contributions on both authors’ parts. This research was supported by a Grant-in-Aid of Research, Scholarship, and Artistry from the University of Minnesota, Graduate School to Benjamin Munson. We express our strong gratitude to Lindsey Kuntz for her help with the perceived /s/ accuracy task. We thank Molly Babel, Adriane Baylis, Nancy DeBoe, Sarah Jefferson, Eden Kaiser, Elizabeth McDonald, Marie Meyer, and Aubrey White for their work on various aspects of this project. This study was inspired, in part, by the comments of an anonymous reviewer of an earlier study (Munson, Jefferson et al., 2006), and an earlier version of this manuscript. We also thank audiences at the 2006 meeting of the American Speech–Language–Hearing Association (where there results of tasks 1 and 2 were presented) and the 2008

Linguistic Society of America (where preliminary results for a subset of listeners in Experiment 2 were presented) for useful comments on this work as it was in progress.

References

- Babel, M. E. (2007). Social judgments and their acoustic cues in read speech. *UC Berkeley Phonology Lab Annual Report*, 226–264.
- Babel, M. E., & Johnson, K. (2006). Global talker characteristics and judgments of gay-sounding speech. *Oral presentation given at the conference on new ways of analyzing variation in language*. Columbus, OH, 9–12 November 2006.
- Bargh, J. A. (2006). What have we been priming all of these years? *European Journal of Social Psychology*, 36, 147–168.
- Baum, S., & McNutt, J. (1990). An acoustic analysis of frontal misarticulation of /s/ in children. *Journal of Phonetics*, 18, 51–63.
- Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, 82, 32–64.
- Carahaly, L. (2000). *Listener accuracy in identifying the sexual orientation of male and female speakers*. Unpublished MA thesis, The Ohio State University.
- Dart, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, 79.
- Eadie, T., & Doyle, P. (2002). Direct magnitude estimation and interval scaling of pleasantness and severity in dysphonic and normal speakers. *Journal of the Acoustical Society of America*, 112, 3014–3021.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12, 453–476.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, 15, 115–141.
- Fazio, R. H., & Olson, M. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297–327.
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, P. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 115–123.
- Foulkes, P., & Docherty, G. (2000). Another chapter in the story of /r/: ‘Labiodental’ variants in British English. *Journal of Sociolinguistics*, 4, 30–59.
- Fox, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 526–540.
- Fuchs, S., & Toda, M. (2009). Do differences in male versus female /s/ reflect biological factors or sociophonetic ones? In: S. Fuchs, M. Toda, & M. Zygis (Eds.), *An interdisciplinary guide to turbulent sounds*. Berlin: Mouton de Gruyter.
- Gaudio, R. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, 69, 30–57.
- Harry, J., & DeVall, W. (1978). Age and sexual culture among homosexually oriented males. *Archives of Sexual Behavior*, 7, 199–209.
- Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, 23, 351–379.
- Hedrick, M., & Ohde, R. (1993). Effect of relative amplitude of frication on perception of place of articulation. *Journal of the Acoustical Society of America*, 94, 2005–2026.
- Hermes, D. J., & van Gestel, J. C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America*, 90, 97–102.
- Johnson, K., Strand, E., & D’Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, 27, 359–384.
- Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108, 1252–1263.
- Kulick, D. (2000). Gay and lesbian language. *Annual Review of Anthropology*, 29, 243–285.
- Lambert, W., Hodgson, C., Gardner, R., & Fillenbaum, S. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology*, 60, 44–51.
- Levon, E. (2006). Hearing “gay”: Prosody, interpretation, and the affective judgments of men’s speech. *American Speech*, 81, 56–78.
- Levon, E. (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society*, 36, 533–554.
- Linville, S. (1998). Acoustic correlates of perceived versus actual sexual orientation in men’s speech. *Pholia Phoniatica et Logopaedica*, 50, 35–48.
- Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *Journal of the Acoustical Society of America*, 78, 40–48.
- Mack, S. (2010a). Perception and identity: Stereotypes of speech and sexual orientation in Puerto Rican Spanish. In: C. Borgonovo, M. Español-Echevarría, & P. Prevoste (Eds.), *Selected proceedings of the 12th Hispanic linguistics symposium* (pp. 136–147). Somerville, MA: Cascadilla Press.
- Mack, S. (2010b). *A sociophonetic analysis of perception of sexual orientation in Puerto Rican Spanish. Laboratory phonology*, 11. New York: Mouton de Gruyter.
- Madon, S. (1997). What do people believe about gay males? A study of stereotype content and strength. *Sex Roles*, 37, 663–685.
- Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken fricatives. *Journal of the Acoustical Society of America*, 125, 3962–3973.
- Marsh, R., Cook, G., & Hicks, J. (2006). Gender and orientation stereotypes bias source-monitoring attributes. *Memory*, 14, 148–160.
- Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Malden, MA: Blackwell.

- Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech*, 50, 125–142.
- Munson, B. (2010). Variation, implied pathology, social meaning, and the 'gay lisp': A response to Van Borsel et al. (2009). *Journal of Communication Disorders*, 43, 1–5.
- Munson, B. (2011). The influence of actual and imputer talker gender on fricative perception, revisited. *Journal of the Acoustical Society of America*, 130. doi:10.1121/1.364140.
- Munson, B., & Coyne, A. (2010). The influence of apparent vocal-tract size, contrast type, and implied sources of variation on the perception of American English voiceless lingual fricatives. *Journal of the Phonetic Society of Japan*, 14, 48–59.
- Munson, B., Ferguson, S. H., & Connealy, C. (2009). Perceived sexual orientation and speech style: A perceptual and acoustic analysis of intentionally clear and conversational speech. *Poster presented at the spring 2009 meeting of the Acoustical Society of America*. Also in *Journal of the Acoustical Society of America*, 125, 2574. Downloaded on January 18, 2010 from <http://www.tc.umn.edu/~munso005/MunsonFergusonConnealy_May13_Final.pdf>.
- Munson, B., Jefferson, S. V., & McDonald, E. C. (2006). The influence of perceived sexual orientation on fricative identification. *Journal of the Acoustical Society of America*, 119, 2427–2437.
- Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34, 202–240.
- Nagao, K. (2006). *Cross-language study of age perception*. Unpublished Ph.D. Dissertation. Bloomington, IN: Department of Linguistics, Indiana University.
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18, 62–85.
- Nygaard, L., Sommers, M., & Pisoni, D. B. (1995). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception and Psychophysics*, 57, 989–1001.
- Pentz, A., Gilbert, H., & Zawadzki, P. (1979). Spectral properties of fricative consonants in children. *Journal of the Acoustical Society of America*, 66, 1891–1892.
- Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America*, 116, 1905–1908.
- Pisoni, D., Nusbaum, H., Luce, P., & Slowiack, L. (1985). Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, 4, 75–95.
- Podesva, R., Jamsu, J., Callier, P., & Heitman, J. (2008). The social meaning of released /t/ among U.S. politicians: Insights from production and perception. *Talk presented at the conference new ways of analyzing variation*. Houston, TX, November 8, 2008.
- Rendall, D., Vasey, P., & McKenzie, J. (2008). The queen's English: An alternative, biosocial hypothesis for the distinctive features of 'gay speech. *Archives of Sexual Behavior*, 37, 188–204.
- Rogers, H., & Smyth, R. (2003). Phonetic differences between gay- and straight-sounding male speakers of North American English. *Paper presented at the 15th international congress of phonetic sciences*. Barcelona, Spain, August 3–9, 2003.
- Russo, V. (1987). *The celluloid closet: Homosexuality in the movies (Revised Edition)*. New York: Harper and Row.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.
- Smit, A. (1993a). Phonologic error distributions in the Iowa–Nebraska articulation norms project: Consonant singletons. *Journal of Speech and Hearing Research*, 36, 533–547.
- Smit, A. (1993b). Phonologic error distributions in the Iowa–Nebraska articulation norms project: Consonant clusters. *Journal of Speech and Hearing Research*, 36, 931–947.
- Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, 32, 329–350.
- Strand, E. (2000). *Gender stereotype effects in speech processing*. Ph.D. Dissertation. Department of Linguistics. Columbus, OH: The Ohio State University.
- Strand, E., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In: D. Gibbon (Ed.), *Natural language processing and speech technology. Results of the third KONVENS conference, Bielfelt, October 1996* (pp. 14–26). Berlin: Mouton de Gruyter.
- Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. In: J. Cole, & J. Hualde (Eds.), *Laboratory phonology 9*. Berlin: Mouton de Gruyter.
- Van Borsel, J., De Bruyn, E., Lefebvre, E., Sokoloff, A., De Ley, S., & Baudonck, N. (2009). The prevalence of lisping in gay men. *Journal of Communication Disorders*, 42, 100–106.
- Van Riper, C. (1937). *Speech correction*. New York: Prentice-Hall.
- Zwicker, E., & Ternhardt, E. (1980). Analytical expression for critical-band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America*, 68, 1523–1525.