Comparing first spectral moment of Australian English /s/ between straight and gay voices using three analysis window sizes

Tünde Szalay[1], *John Holik*[1], *Duy Duong Nguyen*[1], *James Morandini*[1], *Catherine J. Madill*[1]

[1]The University of Sydney, Australia

tuende.szalay, john.holik, duong.nguyen, james.morandini, cate.madill {@sydney.edu.au}

Abstract

Men producing /s/ with higher first spectral moment (M1) are more likely to be perceived as gay, yet it is unclear if M1 differs in production. Inconsistent results might be caused by inherent change in M1 over time. Therefore, we explored M1 change over time and tested if the length and location of the analysis window affects results on gay-straight differences. 37 gay and 29 straight male speakers of Australian English produced two /s/ tokens in continuous speech. M1 was extracted in each quarter of /s/ to explore change over time and in three windows: the entire fricative, the mid-50 ms, and the third quarter. Gay and straight men produced lower M1 in the first and last quarters relative to the midpoint. Despite the M1 change over time, we found no effect of analysis window on gay-straight differences, as gay men consistently showed higher M1. The lack of effect of analysis window on M1 is attributed to the overlap between the analysis windows caused by the duration of /s/.

Index Terms: sexual orientation, fricatives, spectral change, data collection with smart phones, gay males

1. Introduction

Variation in speech sounds conveys socio-indexical information regarding speaker characteristics and group affiliations, such as ethnicity, social class, geographical origin, gender, and sexual orientation [1, 2]. Listeners rely on such variation to attribute group affiliations to speakers [3]. Listeners can reliably identify speakers' sexual orientation as gay or straight from speech samples alone in several languages, [4, 5, 6, 7], including Australian English (AusE) [8], potentially using acoustic differences such as fundamental frequency and spectral vowel quality showing systematic differences between gay and straight men [4, 9, 10].

In particular, /s/ is perceived by English listeners as an important cue to men's sexual orientation: male speakers producing /s/ with higher M1 and more negative skew are more likely to be perceived as gay [2, 11, 12, 6]. However, there is no clear evidence that M1 and skewness of /s/ correlate with actual sexual orientation [5, 2, 6]. In a sample of five gay and four straight male speakers of American English, higher M1 correlated with being gay as well as with being perceived as gay; data on skewness were not reported [5]. In a larger cohort of American English speakers, more negatively skewed /s/ was associated with being gay and with being perceived as gay [2]. This larger sample did not show evidence for higher M1 correlating with actual or perceived gay sexual orientation [2, 11]. In Italian, no significant difference between spectral peak and skewness of gay and straight men was found; however, M1 and negative skew correlated with being perceived as gay [6]. To the best of our knowledge, differences in /s/ production by gay and straight men have not been examined in AusE.

High-frequency spectral peaks and negatively-skewed noise are more likely to be present in the middle and end of /s/, showing inherent temporal changes in /s/ production [13, 14, 15]. The M1 of /s/ shows a gradual increase at fricative onset, stabilises at the midpoint [13] or the two-third mark [15] and falls during the offset. The gradual increase in spectral mean is attributed to the gradual raising of the jaw during /s/ onset and its lowering during offset [15]. As the vertical jaw movement is compensated for by a vertical movement of the tongue tip, the alveolar constriction location and fricative constriction degree remain constant [15].

The constant constriction degree and location of /s/ may suggest that its acoustic characteristics are best extracted from its entire length. The change in spectral mean suggests that M1 is best measured near the middle or the end and that location and duration of analysis window from which spectral moments are extracted may affect M1 measurements. Thus, choice of analysis window may contribute to the results on gay-straight differences being inconsistent. For example, [5] measured centre of gravity throughout the whole fricative, whereas [2] and [6] did not report if the spectral moments were measured using the whole or part of the fricative.

To better understand gay-straight differences in the dynamics of /s/, we measured M1 in the first, second, third, and last quarter of the fricative. To explore how choice of analysis window may affect the results, we created three analysis windows using the entire fricative [5], a 50 ms long window at the midpoint [13], and the third quartile [15].

2. Methods

The current sample was drawn from a larger cohort of Australian gay, bisexual, and heterosexual men, recruited between 2015 – 2016 for an online survey to examine sexual orientation, gender expression, and mental health in men [16].

2.1. Speakers

Sixty-four male [gay = 36, mean age = 33.2, straight = 28, mean age = 35.9] speakers of Australian English (AusE) producing acceptable audio quality were identified in a cohort of 171. Sexual orientation was determined based on self-identification as gay or straight. Participation was incentivised by the chance to win one of ten $50 gift vouchers.

2.2. Material and procedure

Speakers produced the first four lines of the Australian National Anthem '*'Australians all let us rejoice, for we are young and free. We've golden soil and wealth for toil, our home is girt by sea*". Two tokens of word-final /s/ were identified in ***us*** and *rejoice*. /s/ tokens in *us* were excluded for speakers who produced *Australians let us all rejoice*. Participants recorded their own voice using their smartphones. Expert listeners excluded all recordings with background noise or distortion in a consensus procedure, as well as non-native speakers, selecting the 64 analysed recordings.

2.3. Phonetic analysis

123 /s/ tokens were analysed (64 speakers x 2 target fricatives - 5 excluded tokens). Prior to analysis, audio was filtered using Audacity®[17] High-Pass Filter with a cut-off frequency of 500 Hz and a 6 dB roll-off. Segment boundaries were located automatically using the MAUS forced aligner with the AusE grapheme-to-phoneme converter [18, 19, 20] and corrected manually in a Praat interface [21]. Fricative onset was determined on the basis of high-intensity noise onset. Fricative offset was determined on the basis of pause before the following word or formant onset of the following phoneme.

To capture change over time, M1 was estimated in the first quarter (from onset to 25% of the fricative), second (25% to 50%), third (50% to 75%), and fourth (75% to offset) quarter (Figure 1). To capture potential window effects, three analysis windows were created using (1) the entire length of the fricative; (2) the mid-50ms of the fricative; and (3) the 50%-75% interval of the fricative relative to its duration (Figure 2). For fricatives with less than 50ms total duration, the mid-50ms window was not analysed. Each quartile and analysis window was extracted and converted into a frequency spectrum using fast Fourier transformation. M1 was estimated as the average frequency of the spectrum weighted by the power spectrum (i.e., the absolute value of the spectrum's frequency to the power of two) using Praat [21].

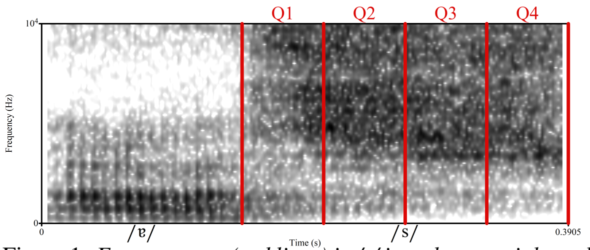

Figure 1: *Four quarters (red lines) in /s/ in* us *by a straight male.*

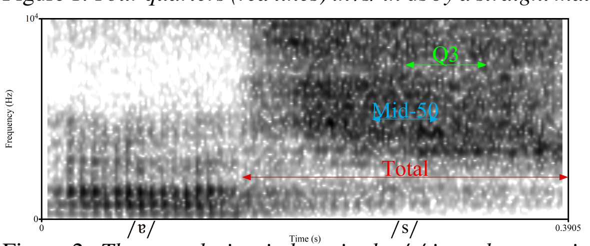

Figure 2: *Three analysis windows in the /s/ in* us *by a straight male. Red arrow (bottom): total duration. Blue arrow (middle): mid-50 ms. Green arrow (top): third quarter.*

2.4. Statistical analysis

To validate measurements extracted from smartphone recordings, M1 values in our dataset were compared to normative M1 values of /s/ produced by male speakers of SAusE using a one-sample t-test [22]. Because in the current dataset, M1 was measured adjacent to /ɹ/ in the phrase ***us** rejoice* and AusE /s/ is retracted (i.e., produced with a lower M1) in the /sCr/-context, reference values of /s/ produced in the /sCr/-context were selected [22]. M1 measurements in each quarter were compared to location-matched M1 values (Table 1). M1 values estimated in the three analysis windows (Total, mid-50ms, and third quarter) were compared to peak M1 of 5250 Hz measured at midpoint [22]. Gay and straight men's M1 values were compared to the same reference measurements separately.

Table 1: *Reference M1 values for the four quarters*

	Q1	Q2	Q3	Q4
Reference value	4800	5250	5100	4500

To explore dynamic properties of fricatives, we constructed a generalised linear mixed effect model (GLM) with the dependent variable M1 using the Gamma family with the identity linking function [23]. The independent variables were Sexual Orientation (dummy coded comparing gay to the baseline straight) and Quarter (sum-coded comparing each quarter to the grand mean). Speaker was added as a random intercept. Convergence was estimated using the BOBYQA (Bound Optimization BY Quadratic Approximation) optimizer and an increased number of maximum iterations [24]. p-values were calculated using Satterthwaite's degrees of freedom method [25].

To test if choice of analysis window affects spectral centre of gravity, we constructed another GLM with the dependent variable M1 using the Gamma family with the identity linking function [23]. The independent variables were Sexual Orientation (dummy coded comparing gay to the baseline straight) and Analysis Window (dummy coded with measurements taken in the whole fricative as a baseline). Speaker was added as a random intercept. Convergence was estimated using the Nelder-Mead optimizer and an increased number of maximum iterations [26]. p-values were calculated using Satterthwaite's degrees of freedom method [25]. Pairwise planned comparisons with Bonferroni correction were used to examine gay-straight differences in each analysis window [27]. Statistical analysis was carried out in R [28].

3. Results

3.1. Validity of M1 measurements

Straight men produced /s/ with a significantly lower M1 in the first quarter compared to the reference value in the /sCr/-context (95% CI = 3501 – 4474, mean = 3987, t = -3.3497, p = 0.0015) [22]. No significant difference was found between the M1 produced by straight men and the reference values [22] in the second, third, and fourth quarters (Second quarter: CI = 4969 – 5868, mean = 5414, t = 0.7245, p = 0.472; Third quarter: CI = 4893 – 5754, mean = 5323, t = 1.041, p = 0.3023; Fourth quarter: CI = 4368 – 5207, mean = 4788, t = 1.3766, p = 0.1745). No significant difference was found between gay men's M1 production and reference values in the first quarter (CI = 3967 – 4864, mean = 4415, t = -1.7123, p = 0.0913) [22]. Gay men produced /s/ with a significantly higher M1 than the reference values in all other quarters (Second quarter: CI = 5672 – 6565, mean = 6119, t = 3.882, p = 0.0002; Third quarter: CI = 5713 – 6639, mean = 6176, t = 4.6341, $p < 0.0001$; Fourth quarter: CI = 5318 – 6324, mean = 5821, t = 5.2388, $p < 0.0001$).

No significant difference was found between the M1 produced by straight men and /s/-produced in the /sCr/-context [22] using any of the analysis windows (Total: 95% CI = 4499 – 5424, mean = 4961, t = -1.2517, p = 0.2163, 50ms: 95% CI = 4891 – 5792, mean = 5341, t = 0.407, p = 0.6858; Third quarter: CI = 4893 – 5754, mean = 5323, t = 0.3422, p = 0.7336). Gay men produced /s/ with significantly higher M1 than the reference value using all analysis windows (Total: 95% CI = 5288 –

6235, mean = 5762, t = 2.1546, p = 0.0347; 50ms: 95% CI = 5698 – 6618, mean = 6158, t = 3.9407, p = 0.0002; Third quarter: CI = 5713 – 6639, mean = 6176, t = 3.9879, p = 0.0002).

3.2. Change over time

In the first GLM, the main effect of Gay Sexual Orientation shows that gay men produce /s/ with a higher mean M1 compared to straight men ($\beta = 764, t_{38.15} = 22.03, p < 0.0001$) (Figure 3). In straight men's speech, M1 in the first quarter of /s/ was significantly lower compared to their mean M1 ($\beta = -1048, t_{27.83} = -47.654, p < 0.0001$). M1 was significantly higher in straight men's speech in the second and third quarters compared to their mean M1 (Second quarter: $\beta = 484, t_{28.09} = 17.24, p < 0.0001$, Third quarter: $\beta = 475, t_{29.4} = 16.14, p < 0.0001$) (Figure 3). These results indicate that /s/ produced by straight men reaches its peak M1 during the second and the third quarters.

The significant negative interaction between the first quarter and Gay Sexual Orientation indicates that the M1 increase in gay men's speech is smaller in the first quarter than overall ($\beta = -324, t_{25.83} = -12.56, p < 0.0001$) (Figure 3). The significant positive interaction between gay sexual orientation and the second and third quarters indicate that the increase in gay men's M1 in those quarters relative to the mean is larger than the increase in straight men's M1 (Second quarter: $\beta = 80, t_{31.33} = 2.46, p = 0.014$ Third quarter: $\beta = 129, t_{25.15} = 5.12, p < 0.0001$). The interactions between Gay Sexual Orientation and the first three quarters collectively indicate that gay speakers produce a larger M1 increase from /s/ onset to the midpoint relative to straight men (Figure 3).

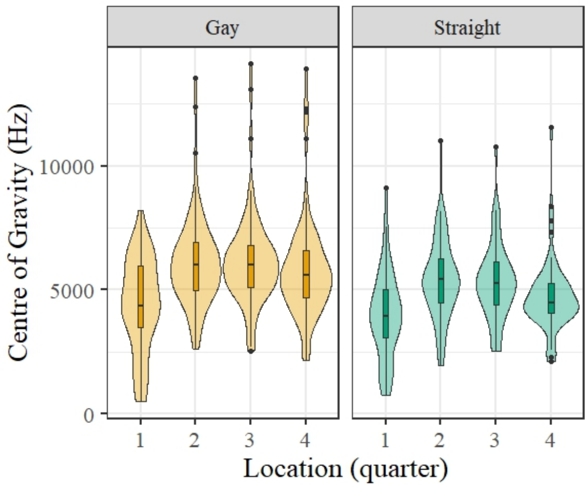

Figure 3: *Change in M1 over time.*

3.3. Effect of analysis window

In our second GLM, the main effect of Gay Sexual Orientation showed that gay men produced /s/ with a significantly higher M1 when M1 was measured in the entire duration of the fricative ($\beta = 737, t_{40.16} = 18.36, p < 0.0001$) (Figure 4). The main effect on Window Size showed that M1 of straight men was higher when M1 was measured in the mid-50 ms window ($\beta = 424, t_{39.22} = 10.82, p < 0.0001$) and in the third quarter of the fricative ($\beta = 417, t_{32.35} = 12.9, p < 0.0001$) compared to measuring M1 in the entire /s/ duration (Figure 4). There were no significant interactions between Gay Sexual Orientation and Window. That is, we found no evidence for choice of analysis window affecting measurements extracted from the speech of gay and straight men differently.

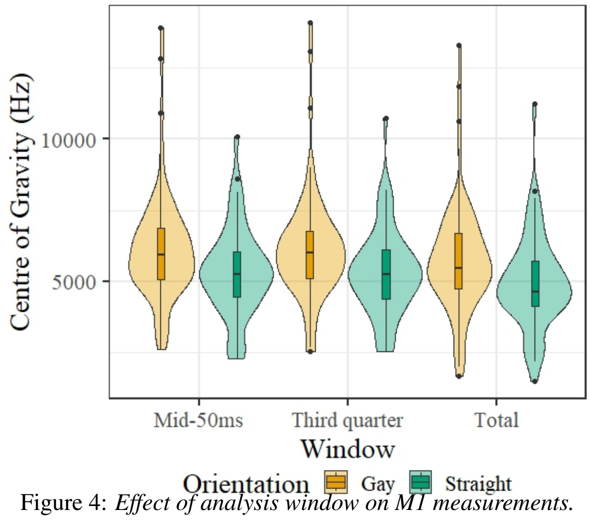

Figure 4: *Effect of analysis window on M1 measurements.*

Planned comparisons showed that gay men produced /s/ with a significantly higher M1 when M1 was measured in the total duration of the fricative ($\beta = 737, z_{40.2} = 19.351, p < 0.0001$), in the mid-50 ms window ($\beta = 780, z_{55.1} = 14.15, p < 0.0001$), and in the third quarter ($\beta = 794, z_{49.8} = 15.94, p < 0.0001$) (Figure 5).

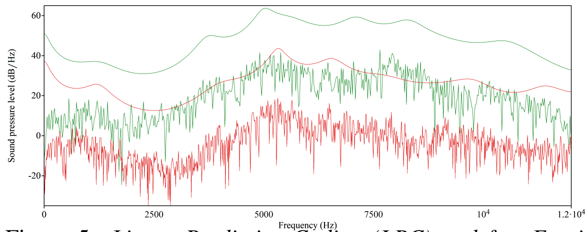

Figure 5: *Linear Predictive Coding (LPC) and fast Fourier-transformed (FFT) envelopes extracted from the third quarter of /s/ in* us *by age-matched gay (red) and straight (green) speaker.*

4. Discussion

4.1. M1 change over time

We explored inherent spectral change over time by extracting M1 in four quarters of /s/. M1 was lower near the fricative onset and offset and higher near the midpoint. Our results are consistent with the previous research on American and Australian English fricatives showing the same pattern [13, 15, 22]. The spectral change has been attributed to jaw raising during fricative production: the jaw is raised from fricative onset to midpoint, leading to an increase in M1, and as the jaw lowers from midpoint to offset, M1 decreases [15]. In our dataset, low M1 at fricative onset and offset can also be attributed to the coarticulatory influence of /ɹ/ [22]. In AusE, /s/ is produced with a lower M1 in the /sCr/ context relative to the pre-vocalic context (e.g, *strict* vs *stick*) [22]. The lower M1 is consistent with a more retracted place of /s/ articulation caused by the coarticulatory influence of tongue dorsum backing during /ɹ/ production [22]. In our dataset, M1 was estimated in the phrase *let us rejoice*, thus the proximity of /ɹ/ to the offset of /s/ in *us* and the onset of /s/ in *rejoice* may further reduce M1 at offset and onset.

When comparing M1 changes over time between gay and straight men, we found similar M1 trajectories. While overall

M1 was higher for gay men than straight men, both groups produced the peak M1 near the midpoint, in the second and third quarters (25%–75% of the total fricative). For both groups, M1 was lowest at the fricative onset. Gay men produced /s/ with a larger difference between the onset and the midpoint than straight men due to gay men's M1 being lower near the onset. The lower M1 near the onset did not compensate for the higher M1 during the rest of the fricative, as indicated by the overall higher M1 in gay men's production.

4.2. Effect of analysis window

We explored whether the location and duration of the analysis window affects M1 measurements and/or differences between M1 produced by gay and straight men. We found that M1 values were lower when estimated in the entire fricative duration compared to estimating M1 in the mid-50 ms or in the third quartile. M1 measured in the entire fricative is lower, as the total duration includes the fricative onset and offset in which M1 values are lower. In contrast, using the mid-50ms window and the third quartile excludes the lower M1 values measured near onset and offset, resulting in overall higher M1.

M1 measurements were consistently higher for gay men compared to straight men. No difference was found in window-size effects between gay and straight men, and gay men showed higher M1 than straight men using any of the three analysis windows. That is, the size and location of the analysis window affected absolute M1 values but no effect of M1 values relative to sexual orientation was found. The lack of difference of M1 values between analysis windows may be attributed to fricative duration (Figure 6). The shorter the duration, the larger the overlap between the analysis windows (e.g., for 50ms-long, there is no difference between measuring M1 in the entire fricative and measuring M1 in the mid-50ms). Therefore, the size and the location of the analysis window is less likely to affect results in shorter fricatives; however, differences may increase as fricative duration increases. Our results on gay men producing /s/ with a higher M1 are consistent with [5], where M1 was measured in the entire fricative. Other studies reporting neither window length nor duration values found no M1 difference [2, 6].

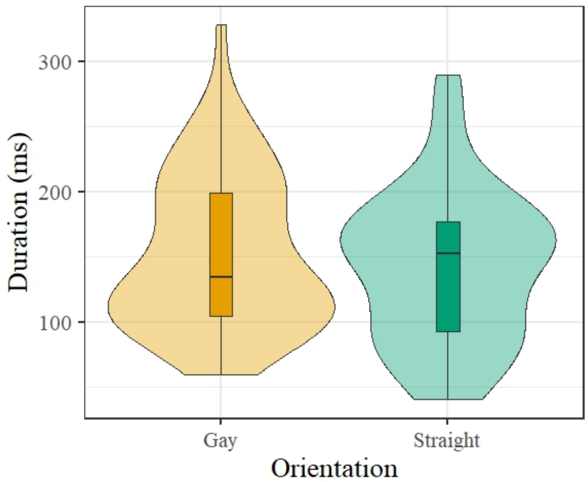

Figure 6: *Fricative duration.*

4.3. Implications for phonetic analysis of clinical data

The main limitation of our research was using smartphones to record audio. Online data collection was chosen to reach the largest number of participants. As it was not guaranteed that participants would be able to view text on their phone while recording their voice, the Australian National Anthem, a well-known text that could be recited, was chosen as the stimulus. Using audio recorded via smartphones is likely to have affected the M1 measurements: acoustic measurements vary between personal devices [29]. Variation caused by individuals' smartphones may have influenced between-group differences.

The two /s/ instances were produced in continuous speech, in different phonetic contexts. While continuous speech has high ecological validity, variation associated with coarticulation may have affected group-differences between gay and straight males, as speakers may have exhibited different coarticulatory patterns irrespective of their sexual orientation, leading to inter-speaker variation [30, 31]. Differences between anticipatory and carryover influence of /ɹ/ on /s/ may have influenced the extent of /s/-retraction, as AusE speakers show individual differences in the extent to which they retract /s/ in the /sCr/-context [30]. The M1 of /s/ varies with vowel-context [32, 15]; individual differences in vowel-/s/ coarticulation were found in Cantonese, as some show vowel-/s/ coarticulation based on vowel height, while others based on vowel-frontness [31] .

We validated our data by comparing our measurements to normative data [22]. Our M1 measurements for straight men were consistent with M1 measurements reported for AusE in the /sCɹ/-context in all quarters except for the first [22]. Therefore, the M1 measurements in /s/ produced by gay men exceeding both normative values and the M1 of /s/ produced by straight men can be attributed to differences in sexual orientation.

Extracting reliable acoustic characteristics from speech collected using smartphones with a phonetically not balanced stimulus is promising. It may open new research collaborations between sociophonetics and other fields. For instance, future studies may use clinical data collected through voice or speech therapy recorded in a quiet room using stimulus primarily aimed at assessment of speech disorder rather than sociophonetic analysis. However, normative data from a comparable speaker group (e.g., matched for gender, geographical origin, etc.) must be available when analysing data collected for other purposes.

5. Conclusion

Both gay and straight speakers of AusE produce a peak M1 near the midpoint of /s/, in the 25%-75% window of their total /s/ duration. Gay speakers of AusE produce /s/ with an overall higher M1 compared to straight men in all analysis windows. Gay men produce /s/ with a larger increase from onset to peak than straight men. Therefore, future sociophonetic studies on sexual orientation and fricative production should (1) measure dynamic properties of /s/; (2) measure M1 near the midpoint; or (3) empirically motivate and report details of the analysis window. Future studies may explore variation within gay speakers and if AusE listeners' perception of sexual orientation is influenced by /s/-production.

6. Acknowledgements

Human Research Ethics Committees approved data collection at Griffith University (PSY/13/15/HREC) and analysis at The University of Sydney (2019/287). This research was funded by the Dr Liang Voice Program, The University of Sydney. The authors thank Antonia Chacon for quality control of sound files.

7. References

[1] P. Foulkes and G. Docherty, "The social life of phonetics and phonology," *Journal of phonetics*, vol. 34, no. 4, pp. 409–438, 2006.

[2] B. Munson, E. C. McDonald, N. L. DeBoe, and A. R. White, "The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech," *Journal of Phonetics*, vol. 34, no. 2, pp. 202–240, 2006.

[3] K. Drager, "Sociophonetic variation in speech perception," *Language and Linguistics Compass*, vol. 4, no. 7, pp. 473–480, 2010.

[4] R. P. Gaudio, "Sounding gay: Pitch properties in the speech of gay and straight men," *American speech*, vol. 69, no. 1, pp. 30–57, 1994.

[5] S. E. Linville, "Acoustic correlates of perceived versus actual sexual orientation in men's speech," *Folia phoniatrica et logopaedica*, vol. 50, no. 1, pp. 35–48, 1998.

[6] S. Sulpizio, F. Fasoli, A. Maass, M. P. Paladino, F. Vespignani, F. Eyssel, and D. Bentler, "The sound of voice: Voice-based categorization of speakers' sexual orientation within and across languages," *PloS one*, vol. 10, no. 7, p. e0128882, 2015.

[7] J. V. Valentova and J. Havlíček, "Perceived sexual orientation based on vocal and facial stimuli is linked to self-rated sexual orientation in czech men," *PloS one*, vol. 8, no. 12, p. e82417, 2013.

[8] J. S. Morandini, D. Beckman-Scott, C. J. Madill, and I. Dar-Nimrod, "Bidar: Can listeners detect if a man is bisexual from his voice alone?" *The Journal of Sex Research*, pp. 1–13, 2023.

[9] J. B. Pierrehumbert, T. Bent, B. Munson, A. R. Bradlow, and J. M. Bailey, "The influence of sexual orientation on vowel production (l)," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 1905–1908, 2004.

[10] S. Kachel, A. Radtke, V. G. Skuk, R. Zäske, A. P. Simpson, and M. C. Steffens, "Investigating the common set of acoustic parameters in sexual orientation groups: A voice averaging approach," *PloS one*, vol. 13, no. 12, p. e0208686, 2018.

[11] B. Munson, "The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation," *Language and speech*, vol. 50, no. 1, pp. 125–142, 2007.

[12] S. Mack and B. Munson, "The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the "gay lisp" stereotype," *Journal of Phonetics*, vol. 40, no. 1, pp. 198–212, 2012.

[13] A. Jongman, R. Wayland, and S. Wong, "Acoustic characteristics of English fricatives," *The Journal of the Acoustical Society of America*, vol. 108, no. 3, pp. 1252–1263, 2000.

[14] S. J. Behrens and S. E. Blumstein, "Acoustic characteristics of English voiceless fricatives: A descriptive analysis," *Journal of Phonetics*, vol. 16, no. 3, pp. 295–298, 1988.

[15] K. Iskarous, C. H. Shadle, and M. I. Proctor, "Articulatory–acoustic kinematics: The production of American English /s/," vol. 129, no. 2, pp. 944–954.

[16] C. J. Hunt, J. Morandini, I. Dar-Nimrod, and F. K. Barlow, "Why do some gay men identify as "straight-acting" and how is it related to well-being?" *Archives of sexual behavior*, vol. 49, no. 5, pp. 1713–1723, 2020.

[17] Audacity Team, "Audacity®: Free audio editor and recorder. version 3.0.0." 2021. [Online]. Available: RetrievedMarch17th2021fromhttps://audacityteam.org/

[18] T. Kisler, U. Reichel, and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, p. 326–347, 2017.

[19] F. Schiel, "Automatic Phonetic Transcription of Non-Prompted Speech," in *Proceedings 14th International Congress of Phonetic Sciences*, J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey, Eds., San Francisco, CA, USA, August 1999, p. 607–610.

[20] ——, "A statistical model for predicting pronunciation." in *Proceedings 18th International Congress of Phonetic Sciences*, 2015.

[21] P. Boersma and D. Weenink, "Praat 6.1.41." 2021. [Online]. Available: http://www.fon.hum.uva.nl/praat/

[22] M. Stevens and J. Harrington, "The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English," *Journal of Phonetics*, vol. 58, pp. 118–134, 2016.

[23] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, p. 1–48, 2015.

[24] M. J. D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives," *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, p. 26–46, 2009.

[25] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, p. 1–26, 2017.

[26] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The computer journal*, vol. 7, no. 4, pp. 308–313, 1965.

[27] R. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means*, 2019, R package version 1.3.4. [Online]. Available: https://CRAN.R-project.org/package=emmeans

[28] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[29] J. Penney, A. Gibson, F. Cox, M. I. Proctor, and A. Szakay, "A comparison of acoustic correlates of voice quality across different recording devices: A cautionary tale." in *Interspeech*, 2021, pp. 1389–1393.

[30] M. Stevens and D. Loakes, "Individual differences and sound change actuation: evidence from imitation and perception of English /str/," in *Proceedings of the 19th International Congress of Phonetic Sciences*, 2019, pp. 3200–3204.

[31] A. C. L. Yu, "Vowel-dependent variation in Cantonese /s/ from an individual-difference perspective," *The Journal of the Acoustical Society of America*, vol. 139, no. 4, pp. 1672–1690, 2016.

[32] L. V. Bondarko, "The syllable structure of speech and distinctive features of phonemes," *Phonetica*, vol. 20, no. 1, pp. 1–40, 1969.