LARGE-SCALE ACOUSTIC ANALYSIS OF DIALECTAL AND SOCIAL FACTORS IN ENGLISH /S/-RETRACTION

Jane Stuart-Smith,[a] Morgan Sonderegger,[b] Rachel Macdonald,[a] Jeff Mielke,[c] Michael McAuliffe,[b] Erik Thomas[c]

[a]Glasgow University, UK; [b]McGill University, Canada; [c]North Carolina State University, US
Jane.Stuart-Smith@glasgow.ac.uk

ABSTRACT

The retraction of /s/ in /str/, eg *street*, is a sound change found in certain English dialects. Previous work suggests that /s/-retraction arises from lower spectral frequency /s/ in /str/. The extent to which /s/-retraction differs across English dialects is unclear. This paper presents results from a large-scale, acoustic phonetic study of sibilants in 420 speakers, from 6 spontaneous speech corpora (9 dialects) of North American and Scottish English. Spectral Centre of Gravity was modelled from automatic measures of word-initial sibilants. Female speakers show higher frequency sibilants than males, but more so for /s/ than /ʃ/; /s/ is also higher in American than Canadian/Scottish dialects; /ʃ/ is surprisingly variable. /s/-retraction, modelled as retraction ratios, is generally greater for /str/ than /spr skr/, but varies by dialect; females show more retraction in /str/ than males. Dialectal and social factors clearly influence /s/-retraction in English clusters /sp st sk/, /spr skr/, and /str/.

Keywords: sibilants, /s/-retraction, large-scale studes, sociophonetics, English

1. INTRODUCTION

The auditory retraction of /s/ to an [ʃ]-like quality in the cluster /str/, as in e.g. *street*, is a sound change found in some but not all dialects of English [31, 4]. /s/-retraction is typically associated with southern varieties of British English, such as London [1] and the South East [5]. However Glain [11] observed /s/-retraction in all British dialects including Scottish English, bar the North-East, in his auditory analysis of the IDEA corpus [23]. In North America, /s/-retraction has been noted in Philadelphia since the 1980s [17, 18, 31, 13]. It has also been studied in Columbus, Ohio [10], in General American [25], in Louisiana [30], in Raleigh, North Carolina [34], and was noted for African American Vernacular English by Glain [11]. /s/-retraction appears to be well-established in New Zealand [19], and is increasing in some [7], but not all [32], varieties of Australian English. Gender differences in /s/-retraction seem to vary by dialect: in British English, male speakers retract more [5, 11], in Raleigh, only female speakers born more recently show retraction [34]; and in Philadelphia no gender differences were found [13].

The phonetic bases of /s/-retraction likely relate to the lowering of spectral frequency of /s/ in clusters. Baker et al. [4], also [32], show progressive lowering of the spectral Centre of Gravity (COG). The COG of singleton /s/ is higher than in clusters without /r/, /sC = sp st sk/, which is in turn higher than in /sCr = spr skr/, and /str/ shows the lowest COG. The role of coarticulation with /r/ in /s/-retraction is not entirely clear. Mielke et al [24] demonstrate that while ‘retractors’ have a phonologised target closer to /ʃ/ and do not show coarticulation, in ‘non-retractors’ the degree of retraction correlates with decreased articulatory distance between /s/ and /r/. Individual speaker variation, in conjunction with coarticulatory bias, provides a source for /s/-retraction to take off in non-retracting dialects [4, 14].

There is also evidence that /str/ can be phonetically distinct from /sCr/. Spectral COG trajectories for a non-retracting Australian English dialect [32] show anticipatory assimilation for /spr skr/, but not /str/, whose overall shape is similar to /ʃ/ from /s/ onset. Current work on U.S. English also shows that /str/ is produced [25] and perceived [26] differently from /spr skr/, though with substantial individual variation.

Previous studies of /s/ retraction have tended to focus on a single dialect, using different methods to assess retraction. Here we seek to expand the view by using the same acoustic measurement strategy across a large number of speakers from spontaneous speech from several dialects.[1] Our research questions are: How does phonetic context affect English /s/? What is the evidence for /s/-retraction across English varieties? What role does gender play in English /s/-retraction?

2. METHODOLOGY

Our sample was spontaneous speech recordings from 420 speakers from nine broad dialect groups representing English dialects from North America and Scotland, from six corpora held by the SPADE project: (1) subsets of the Santa Barbara corpus [9], representing the **Northern Cities Shifted** (Inland North) region (19 speakers, 11 female) including Philadelphia and New York, and **Western** U.S. (43 speakers, 23 female); (2) the Buckeye corpus [27], representing U.S. North Midland, including **Columbus** (40 speakers, 20 female); (3) the **Raleigh** corpus [8], representing the U.S. urban South (101 speakers, 51 female); (4) the Canadian subset of the International Corpus of English [12], ICE-CAN, representing **Canada** (28 speakers, 10 female); (5) subsets of the Scottish Corpus of Texts and Speech (SCOTS) [2], representing East coast (22 speakers, 11 female), West coast (38 speakers, 19 female), and Highland/Island and North (54 speakers, 34 female) Scottish dialects; (6) the Sounds of the City corpus [33], representing **Glasgow** (70 speakers, 36 female). We removed recordings made in the 1970s from the Glasgow data, so that the year of recording for all data ranges from 1980s to the 2000s. In future work, we intend to consider, where possible, the impact of ethnicity, year of recording and speaker age on /s/-retraction across these dialects [34, 7].

We analysed all instances of stressed, word-initial, prevocalic /s/ e.g. *seat* and /ʃ/ e.g. *sheet*, plus /sp st sk/, e.g. *speech, skip, steep*, /spr skr/, e.g. *spree, scream*, and /str/ e.g. *street*. We did not code for following vowel quality, given inconsistent results in previous studies, e.g. [10, 4]. We predicted COG might lower by onset structure as follows: /s/ > /sp st sk/ > /spr skr/ > /str/ > /ʃ/.

In order to compare acoustic measures across corpora with different sampling rates, we downsampled all soundfiles to 16kHz (the rate of the Buckeye corpus), and high-pass filtered at 1kHz. We selected the central 50% of the sibilant interval [4] from force-aligned segment boundaries, and calculated the duration of the sibilant interval, and COG, spread, peak and slope from spectra taken using a 10ms Hamming window in Praat [6]. We report here measures for COG; note that the impact of downsampling was to depress the COG values to below 8KHz. This means that COGs shown here are relative indicators as opposed to actual values (the spectral peak for female English speakers can be as high as 12kHz).

The data analysis was carried out using new open-source software developed for the SPeech Across Dialects of English (SPADE) project. The Integrated Speech Corpus ANalysis (ISCAN; [22]) software enables automated acoustic phonetic analysis across spoken corpora of diverse formats and sizes. The system aims to overcome the significant practical and methodological barriers to conducting essentially the same study across corpora, including necessary technical skills and non-comparability of results using non-standardized measures. Here, each audio corpus, which had already been force-aligned, was imported and then a custom sibilant measurement script was run, in ISCAN. The speed of this scaled-up automated analysis was already impressive, e.g. the 232-hour Raleigh corpus took only 4.2 hours to import and extract measures.

The initial count for all sibilants was 111,683 tokens. We anticipated that force-alignment followed by automated measurement would result in likely erroneous tokens and measures. We removed about 10% tokens with unexpected COG/peak less than 2400Hz, giving 100,246 tokens. Further data reduction took place during analysis. We used linear mixed effects modelling with *lme4* in R [28], and fit three models to gain three perspectives on sibilants and /s/-retraction.

(1) To consider the impact of phonetic context on sibilants in general, we fit a simple model with COG as the dependent variable, fixed factors of *Duration*, *Dialect*, *Gender*, and *Onset* (5 levels: s sC sCr str sh), and random intercept of *Word* and by Speaker slope for Onset. Further outliers were removed by running the model again on a dataset trimmed with respect to 2.5 standard deviations of the residuals from the first model [3].

(2) To examine the role of dialect and gender on prevocalic /s ʃ/, we fit a model with COG as dependent variable, with the same fixed and random factors, but now *Onset* had 2 levels (s sh), and we included the three-way interaction for *Onset*, *Dialect* and *Gender*.[2] This model was again run on a trimmed dataset to remove outliers.

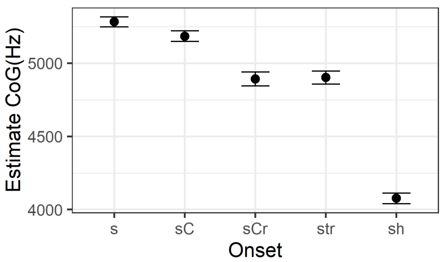
(3) The first two models used COG Hz measures without attempting to normalize the influence of speaker physiology on sibilant acoustics. They present indicative results of overall phonetic context, and allow a view of the potential 'limits' of sibilant space by dialect and gender. To investigate /s/-retraction controlling for speaker characteristics, we calculated retraction ratios [4], which express the position of /s/ in a cluster's COG in relation to the same speaker's mean COG of /s/ (closer to 1) and /ʃ/ (closer to 0); this orientation reflects the acoustic relationship of tokens to /s/ (higher) and ʃ/ (lower).

Calculating the retraction ratio on these corpus data resulted in numerous outliers, especially tokens with values more than 1 [7]. We removed all tokens from speakers with less than 8 tokens for either /s/ or /ʃ/, and then modelled only retraction ratios within a range of 2.5 to -1.5. The same fixed and random factors were used, but *Onset* had 3 levels (sC sCr str), and all two-way interactions between *Onset*, *Dialect* and *Gender* were included (counts did not permit three-way interaction). Results are considered significant when $p < 0.05$; statistics are given for terms of interest.

3. RESULTS

3.1. /s/ by phonetic context

Figure 1: Model estimates of COG (Hz) for English /s/ by phonetic context (n = 98,679)

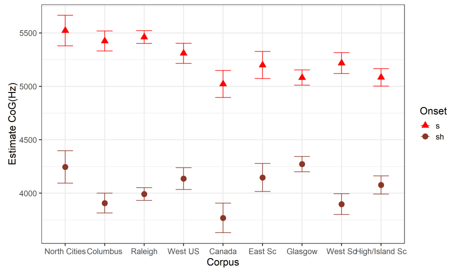


The first model returned significant effects for all fixed factors. As expected there were overall differences in sibilant COG by dialect (Canadian is lowest), gender (higher COG in female speakers) and duration (the longer the duration, the higher the COG). Of interest here is *Onset* ($\chi(18) = 19221, p < 0.0001$), see Fig. 1. Posthoc tests confirm that across all dialects and speakers, /s/ has the highest COG, /s/ in /sC/ is lower than /s/, but higher than /s/ in clusters with /r/; /ʃ/ is much lower. Unlike [32], c.f.[4], /s/ in /str/ is not lower than in /sCr/.

3.2. English prevocalic /s ʃ/ by dialect and gender

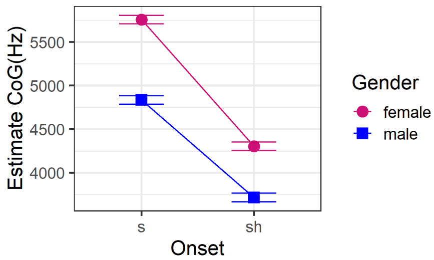
The second model showed a similar effect of duration on COG, and also showed significant interactions of *Onset* by *Gender* ($F(1, 380) = 38.34, p < 0.001$), and *Onset* by *Corpus* ($F(8, 415) = 11.37, p < 0.001$). There are clear differences in the production of both sibilants by dialect; Fig. 2. Pairwise comparisons show that U.S. dialects tend to show similar higher /s/ COGs to each other, while Canadian and Scottish English dialects pattern together in showing lower COGs. /ʃ/ varies substantially across dialects.

Figure 2: Model estimates of COG Hz for English /s ʃ/ by dialect (n = 75,465)



As expected, Fig. 3 shows that /s/ has higher COG than /ʃ/, and female speakers show higher COG than males, but there is also a larger gender difference for /s/ than /ʃ/.

Figure 3: Model estimates of COG Hz for English /s ʃ/ by gender (n = 75,465)



3.3. English /s/-retraction by dialect and gender

The significant interactions from the third model, for acoustic retraction in terms of retraction ratio, are shown in Fig. 4, for *Onset* by *Dialect* ($F(14, 648.5) = 3.86, p < 0.001$) and Fig. 5 for *Onset* by *Gender* ($F(2, 478.6) = 6.9, p < 0.001$). Duration was significant as before. /spr skr/ were too sparse in Northern Cities Shifted dialects, and this group was removed. Low numbers of tokens for /sCr/ also make the data points for West U.S. and Canada less reliable in Fig. 4.

Post-hoc tests confirm the main dialectal findings to be: /s/ in /str/ is only retracted in Columbus and Raleigh (cf [10, 34]), and Canadian English for which no studies appear to exist. There is no /str/ retraction in West U.S., or in Scottish English. Neither American dialect shows /str/ to be more retracted than the /spr skr/ clusters; /s/ COG does not vary according to cluster type in the Scottish dialects.

The difference between /sCr/ and /str/ clusters is found only when viewing the distribution by

Figure 4: Model estimates of retraction ratio for English /sC sCr str/ by dialect (n = 23,671)

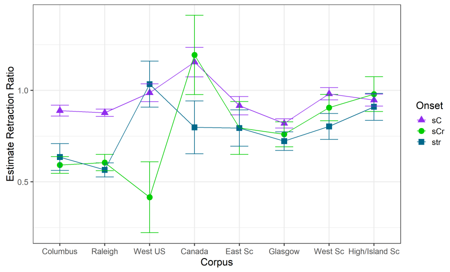
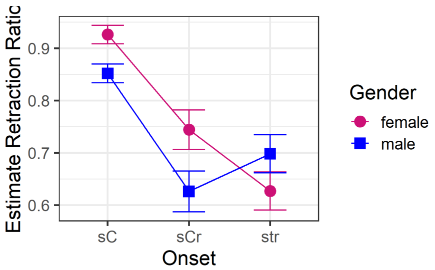


Figure 5: Model estimates of retraction ratio for English /sC sCr str/ by gender (n = 23,671)



Gender and *Onset*, see Fig. 5: female speakers show less retraction than males for /sC sCr/, but more retraction in /str/. Whilst we had suspected that a three-way interaction might uncover an additional relationship between Gender and Dialect with respect to /s/-retraction, this was not found. We fit a final model to a subset of the sample, removing four dialect groups showing low ns for one or more cluster; the same two interactions were confirmed as reported here.

4. DISCUSSION

This study has attempted to scale up our perspective on English /s/-retraction, by carrying out a large-scale, comparable, acoustic analysis across three national varieties and nine broad dialect groups. In terms of the influence of phonetic context on COG of /s/, our results largely align with those of previous studies on American and Australian English [4, 32, 7], which show relatively little retraction in /sp st sk/, and greater retraction in /spr skr str/ clusters. In this respect, we confirm Durian's [10] findings for Columbus, and Wilbanks' [34] for Raleigh. For the latter, we used slightly different measures for a good portion of the same dataset (though without respect to speaker age), and both the coincidental results, and the ability to position this dialect in context with others within and beyond North America is encouraging.

The Scottish varieties are different, since they pattern together showing the same moderate retraction for /s/ in all clusters, with and without /r/. Indeed it is difficult to anticipate the results of /s/-retraction for Scottish English, especially for cities like Glasgow, because prevocalic /s/ is known to be auditorily and acoustically retracted [21]. Only [11] has noted /s/-retraction as an audible process in Scottish English.

The coincidence of /s/ COG in Canadian and Scottish Englishes patterning together is intriguing, especially given the attested historical links in terms of migration. But we also note that the Canadian COGs are overall lower, despite the manual inspection of all recordings. Further Canadian data will resolve whether these lowered values can be interpreted.

We did not find evidence for greater acoustic lowering of /s/ in /str/ than in /spr skr/ clusters in the dialects considered here (see [4, 32, 34, 26]). But these clusters are distinguished in gendered productions, across all dialects, and within the five substantial dialect datasets which could be tested in a model with a three-way interaction. The fact that female speakers in general show lowered /s/ COG in /str/, contrary to any expectations from possibly smaller vocal cavities, suggests that for these dialects at least, female speakers may show a degree of socially-indexed control in producing /str/ (the same may apply in reverse to male speakers for /sCr/) [20]. Greater retraction could arise through coarticulation with specific articulatory variants of /r/, though for this cluster only; it could also arise from liprounding [10, 30].

We intend to continue the work presented here in a number of directions. Beyond expanding the regional dialect scope of our sample, we will also add time depth, both in terms of real- and apparent-time [16] where corpora permit. We intend to consider the impact on large-scale perspectives on sibilants and /s/-retraction from using alternative spectral measures [15], calculated from multitapered spectra [29]. Like Boylan [7], here the variable sibilant productions in spontaneous speech presented some issues with using the retraction ratio to characterise /s/-retraction; this needs closer attention. Finally, our ongoing work is now considering optimal ways of examining the considerable inter-speaker variation, which in initial plots, supports the assumption of the full range of /s/ realizations, also in 'non-retracting' dialects.

5. REFERENCES

[1] Altendorf, U. 2003. *Estuary English: Levelling at the interface of RP and South-Eastern British English*. Tuebingen: Gunter Narr Verlag.

[2] Anderson, W., Beavan, D., Kay, C. 2007. SCOTS: Scottish Corpus of Texts and Speech. In: *Creating and digitising language corpora*. Palgrave Macmillan UK 17–34.

[3] Baayen, R. H. 2008. *Analyzing linguistic data*. Cambridge: Cambridge University Press.

[4] Baker, A., Archangeli, D., Mielke, J. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23, 347–374.

[5] Bass, M. 2009. 'street or shtreet'? Investigating (str-) palatalisation in Colchester English. *Estro: Essex Student Research Online* 10–21.

[6] Boersma, P., Weenink, D. 2013. Praat: doing phonetics by computer.

[7] Boylan, S. P. 2018. */s/ retraction in the /str/ onset in Australian English: Is sound change in progress?* Mres Macquarie University.

[8] Dodsworth, R., Kohn, M. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change* 24, 221–245.

[9] Du Bois, J. W., Chafe, W. L., Meyer, S. A., Thompson, S. A., Martey, N. 2000. Santa Barbara Corpus of Spoken American English. Technical report Linguistic Data Consortium Philadelphia.

[10] Durian, D. 2007. Getting [S]tronger every day? More on urbanisation and the socio-geographic diffusion of (str) in Columbus, OH. *U. Penn Working Papers in Linguistics* 65–79.

[11] Glain, O. 2014. Introducing contemporary palatalisation. *Proceedings of the first Postgraduate and Academic Researchers in Linguistics at York conference (PARLAY 2013)* 16–29.

[12] Greenbaum, S., Nelson, G. 1996. The International Corpus of English (ICE) project. *World Englishes* 15(1), 3–15.

[13] Gylfadottir, D. 2015. Shtreets of Philadelphia: An acoustic study of /str/-retraction in a naturalistic speech corpus. *U. Penn Working Papers in Linguistics* 87–97.

[14] Harrington, J., Kleber, F., Reubold, U., Schiel, F., Stevens, M. 2018. Linking cognitive and social aspects of sound change using agent-based modelling. *Topics in Cognitive Science* 1–22.

[15] Koenig, L. L., Shadle, C. H., Preston, J. L., Mooshammer, C. R. 2013. Toward Improved Spectral Measures of /s/: Results From Adolescents. *Journal of Speech, Language, and Hearing Research* 56(4), 1175–1189.

[16] Labov, W. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.

[17] Labov, W. 1981. Field methods of the project on linguistic change and variation. Technical report Sociolinguistic Working Paper Austin, Texas.

[18] Labov, W. 2001. *Principles of Linguistic Change*. Oxford: Blackwell.

[19] Lawrence, W. 2000. /str/ -> /Str/: Assimilation at a distance? *American Speech* 70, 82–87.

[20] Levon, E., Maegaard, M., Pharao, N., (eds) 2017. *The Sociophonetics of /s/*. Linguistics.

[21] Macafee, C. 1983. *Varieties of English around the world: Glasgow*. Amsterdam: Benjamin.

[22] McAuliffe, M., Coles, A., Goodale, M., Mihuc, S., Wagner, M., Stuart-Smith, J., Sonderegger, M. 2019. ISCAN: A system for integrated phonetic analyses across speech corpora. *Proceedings of the 19th Congress of Phonetic Sciences* Melbourne.

[23] Meier, P. 1997. The International Dialects of English Archive (IDEA).

[24] Mielke, J., Baker, A., Archangeli, D. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory Phonology* 10, 699–730.

[25] Phillips, Jacob, B. 2016. Phonological and prosodic conditioning of /s/-retraction in American English. *Paper presented at Higher order structure in speech variability, LabPhon15 satellite workshop*.

[26] Phillips, Jacob, B., Resnick, 2018. Listeners' social attributes influence sensitivity to coarticulation in the perception of sibilants in nonce words. *Paper presented at NWAV47*.

[27] Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., Fosler-Lussier, E. 2007. Buckeye Corpus of Conversational Speech (2nd release). Technical report Department of Psychology, OSU Columbus, Ohio.

[28] R Core Team, 2013. R: A Language and Environment for Statistical Computing.

[29] Reidy, P. F. 2015. A comparison of spectral estimation methods for the analysis of sibilant fricatives. *The Journal of the Acoustical Society of America* 137(4), 248–54.

[30] Rutter, B. 2011. The acoustic analysis of a sound change in progress: The consonant cluster /str/ in English. *Journal of the International Phonetic Association* 41(1), 27–40.

[31] Shapiro, M. 1995. A case of distant assimiliation: /str/ -> /Str/. *American Speech* 70(1), 101–107.

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

[33] Stuart-Smith, J., José, B., Rathcke, T., Macdonald, R., Lawson, E. 2017. Changing Sounds in a Changing City. In: Montgomery, C., Moore, E., (eds), *Language and a Sense of Place*. Cambridge: Cambridge University Press 38–65.

[34] Wilbanks, E. 2017. Social and structural constraints on a phonetically-motivated change in process: (str) retraction in Raleigh, NC. *U. Penn Working Papers in Linguistics* 301–310.

---

[1] We are grateful to the Trans-Atlantic Platform initiative for funding this research; www.spade.arts.gla.ac.uk.

[2] The data were too sparse to fit a 3-way interaction for model 1.