Desperately seeking 'English' sibilants: Discovering dialect norms and speaker variability for /s ʃ/ from large-scale multi-dialect analysis

Jane Stuart-Smith, Morgan Sonderegger, and The SPADE Consortium
Glasgow University Laboratory of Phonetics; McGill University; www.spade.glasgow.ac.uk

Introduction Understanding what constitutes the phonological margins for any sound, accent or speech community presumes that we also know the scale and nature of the requisite phonetic/phonological distributions themselves. But even for a language as well studied as English, until recently operational constraints on gathering and analysing phonetic data beyond a few dozen speakers per study means that our appreciation of constitutes 'normal' variation for the commonest sounds, is largely based on intuition. For example, the acoustic spectral targets for English prevocalic sibilants /s/ and /ʃ/ are generally assumed to be similar across British and North American dialects of English, with the odd exceptions, e.g. lower Scottish English /s/.[1,2] Sibilants are also thought to vary greatly by-speaker *within* a dialect ("within-dialect") not least for the construction of social identities – with asymmetry in the greater use of /s/ than /ʃ/,[3,4] which may arise from more variability *in general* for English /s/ (within dialect, across dialects) and/or from the sociophonetic focus on /s/ to the detriment of /ʃ/. A puzzle raised by these assumptions is – how can English sibilants be both highly similar (across dialects) and variable (across speakers)? This study uses large-scale acoustic analysis of spontaneous speech from British and North-American dialects to ask: (1) How variable is the production of /s/ and /ʃ/ across English dialects, and (2) within dialects (across speakers)? How do degrees of by-speaker and by-dialect variability compare? (3) Is English /s/ more variable than /ʃ/? **Methods** We report results for spectral Centre of Gravity (COG) from a 2400-8000Hz range, from the central 50% of all instances of word-initial stressed pre-vocalic /s/ and/ʃ/, computed using a custom Praat script and an automated data extraction pipeline.[5] To date we have analysed ~214,000 tokens from 1016 speakers from 12 spoken language corpora of UK and North American English (US and Canada), representing some 15 dialect areas[6] (median n/corpus = 3500, n/speaker = 117). We model COG using a Bayesian mixed-effects model (in brms/Stan),[7,8] with terms to capture by-dialect and by-speaker variability (nested in dialect) in COG of /s/ and /sh/, while controlling for speaker gender and control predictors (e.g. phone duration), with appropriate by-speaker and dialect random slopes. A Bayesian model allows us to explicitly obtain estimates and uncertainty for the degrees of speaker/dialect variability and compare degrees of variability, which is necessary to answer (1)-(3). **Results** The model predicts dialects and speakers (based on this sample) to vary following a nearly-normal distribution (Fig. 2); the widths $\sigma_d(s)$, $\sigma_s(ʃ)$ etc. are the parameters which capture "degree of variability" among dialects/speakers. Every dialect shows expected effects of gender and anteriority (female > male COG, /s/ > /ʃ/). Dialects do differ in /s/ and /ʃ/ COG (σ_d: /s/ = 316±124 Hz, /ʃ/ =201±110 Hz), but the degree of difference is small relative to the COG scale (Fig. 1), as well as relative to speaker variability within-dialect, which is substantial (σ_s: /s/ = 583±28 Hz, /ʃ/ = 545±26 Hz). There is strong evidence that speakers differ more (within-dialects) than dialects ($P(\sigma_s>\sigma_d)$ = 0.99 for /s/ and /ʃ/). Thus, (1) sibilant production is indeed broadly similar across English dialects, but this coexists with (2) high variability across speakers within dialect. And both sibilants show substantial opportunities for subdialectal variation amongst their speakers, which may well arise for sociophonetic reasons. There is also strong evidence that (3) /s/ is more variable than /ʃ/ both across dialects ($P(\sigma_d(s)>\sigma_d(ʃ))$=0.95) and across speakers within dialects ($P(\sigma_d(s)>\sigma_d(ʃ)$=0.99) (Fig. 3), suggesting that English /s/ may indeed be inherently "more variable" than /ʃ/, and hence more amenable for the expression of additional social-indexical meanings.

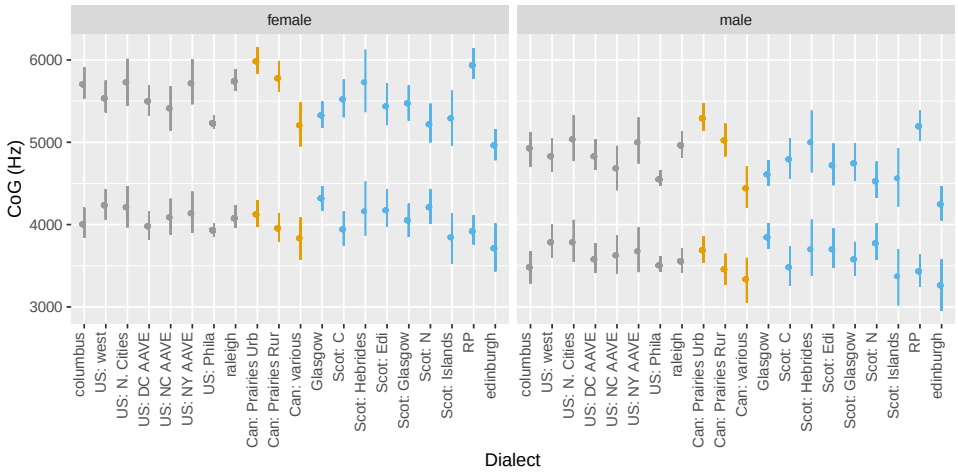

Figure 1: Estimated COG and 95% credible intervals for each dialect, for /s/ (upper) and /ʃ/ (lower), across US (grey), Canada (gold), and UK (blue) dialects.

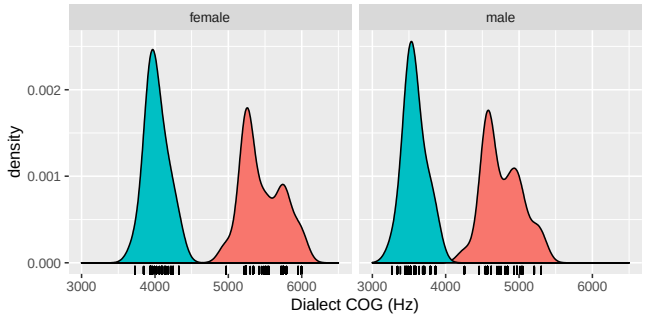

Figure 2: Estimated distribution of COGs across dialects, for /ʃ/ (blue) and /s/ (red), by speaker gender.

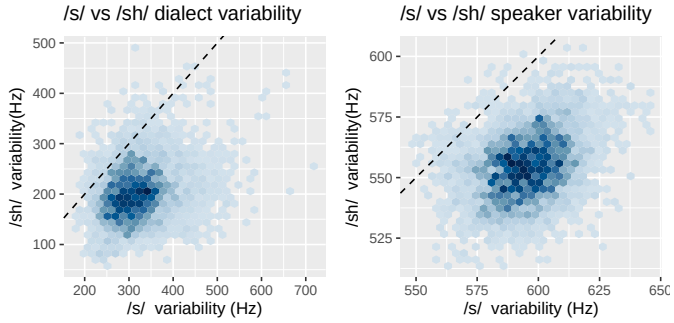

Figure 3: Model's posterior distribution of /s/ and /ʃ/ degree of by-speaker variability (left) and by-dialect variability (right). Darker shading = more likely values. Dotted line is y=x. Both plots show /s/ > /ʃ/ variability with high probability.

References 1. Wells, J. C. *Accents of English 2: The British Isles*. (Cambridge University Press, 1982). 2. Macafee, C. *Varieties of English around the world: Glasgow*. (Benjamin, 1983). 3. Eckert, P. & McConnell-Ginet, S. *Language and Gender*. (Cambridge University Press, 2003). 4. *The Sociophonetics of /s/*. (Linguistics, 55, 2017). 5. McAuliffe, M. *et al.* ISCAN: A system for integrated phonetic analyses across speech corpora. in *Proceedings of the 19th Congress of Phonetic Sciences* (2019). 6. https://spade.glasgow.ac.uk/ 7. Bürkner, P.-C. & others. brms: An R package for Bayesian multilevel models using Stan. (*J. Stat. Softw.* 80, 1–28, 2017). 8. Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F. & Kong, E. J. Bayesian data analysis in the phonetic sciences. (*J. Phon.* 71, 147-161, 2018)