DE GRUYTER

Jane Stuart-Smith*

Through the looking glass

Changing perspectives on /s/ and gender over time in Glasgow

Abstract: This paper considers the relative influence on sociophonetic interpretation of /s/ using ‘static’ and ‘dynamic’ acoustic analysis, where *dynamic* refers to the use of measures which capture the time-varying nature of segmental acoustics, and *static* to measures which are taken at a single point, or from an average across the sound (Docherty et al., 2015; Watson and Harrington, 1999). Static and dynamic Discrete Cosine Transformation (DCT) analyses of spectral Centre of Gravity (CoG) and spectral Slope measures/trajectories were carried out on gendered productions of /s/ and /ʃ/ for a real- and apparent-time 32 speaker sample in Glasgow dialect. Results of static CoG measures, reflecting place of articulation, indicate a reduction of gender differentation over time, such that girls born most recently revert to older vernacular (lower frequency) norms. Adding static spectral Slope, reflecting articulatory constriction, shows a change in gendered differentiation, whereby boys born most recently show a gestural shift. The DCT analysis both confirms the static results (through the first coefficient), and also reveals that dynamic characteristics of both sibilants carry key additional prosodic, linguistic and social information for this community. Together, our results reflect for the first time the usefulness of changing analytical perspectives on /s/, in terms of acousti analysis (from static to dynamic), and of linguistic (/s/ and /ʃ/) and social context (gender and time).

1 Introduction

A constant theme for sociophonetics (Foulkes, 2006; Foulkes et al., 2010; Hay and Drager, 2007), is how best to capture and characterise the relevant aspects of speech which relate to identified social constructs (Docherty et al., 2015). The corollary is how in turn these ‘pictures’ influence our conceptualisation of social-indexicality and speech—a core aspect of all scientific enquiry detailed in the role of ‘thing knowledge’ on scientific description and cognition (Baird, 2004; Dijkstra, 2012; Huggett, 2017).

Glasgow University Laboratory of Phonetics, University of Glasgow, email: Jane.Stuart-Smith@glasgow.ac.uk

This paper focusses on a specific aspect of 'viewing' speech, the relative influence on sociophonetic interpretation of /s/ using 'static' and 'dynamic' acoustic analysis, where *dynamic* refers to the use of measures which capture the time-varying nature of segmental acoustics, and *static* to measures which are taken at a single point, or from an average across the sound (Docherty et al., 2015; Watson and Harrington, 1999). Static acoustic analyses have repeatedly confirmed socially-conditioned variation in the production of /s/ across several varieties of English and for other languages (Levon et al., 2017; Stuart-Smith, 2007). Reidy (2016) shows that dynamic acoustic analysis reveals important cross-linguistic differences between English and Japanese /s/ not discernible from static measures. Here we ask how a dynamic acoustic representation improves our understanding of /s/ with respect to gender.

Specifically, /s/ and gender in Glaswegian are investigated using measures from static and dynamic (Discrete Cosine Transformation) analyses of spectral Centre of Gravity (CoG) and Slope, with respect to *time* and *phonological contrast*. First, are changes in social gender over time linked to changes in the realization of /s/? Synchronic work on acoustic shifts in /s/ production show how sensitive this sound can be in response to social identities (e.g. Podesva and Kajino 2014). Glasgow /s/ has always been thought to be a stable gender marker, but are shifts in the post-industrial cityscape since the Second World War linked to shifts in gendered /s/ production? Second, are changes in gender and /s/ over time related to /ʃ/? This perspective takes /s/ as one of the pair of sibilants, in e.g. *seat* beside *sheet*. There appears to have been little discussion of the social work done by /ʃ/, with the result that /ʃ/ seems to be rather dull compared to its hard-working counterpart /s/ (Eckert, 2003). Our results reflect for the first time the usefulness of changing analytical perspectives on /s/, in terms of acoustic analysis (from static to dynamic), and of linguistic (/s/ and /ʃ/) and social context (gender and time).

2 Background

2.1 Dynamic acoustic analysis in sociophonetic research

Watson and Harrington (1999)'s large-scale study of 19 monophthongs and diphthongs in controlled speech by 132 Australian English speakers, was the first demonstration of how acoustic vowel targets *and* time-varying acoustic information for the first two formants, together serve to differentiate vowel quality, including within monophthongs, especially the tense/lax pairs. Their decision to capture the dynamic behaviour of formant tracks in vowel production using Discrete Cosine Transformation (DCT), had the advantage that a single 'DCT analysis' of formant tracks gives coefficients which capture both target information usually obtained by static measures, through the first coefficient, and dynamic information, through the second and higher coefficients; for DCT, see Section 2.3. Williams and Escudero (2014) extended these conclusions for 16 diphthongs and monophthongs (especially /uː/) from two British English dialects, Southern Standard British English and Sheffield English. They found the first two DCT coefficients together provided optimal separation of vowel quality, noting (p.2754) that the first coefficient is 'indicative of the average resonances of the vocal tract over the course of a vowel token [and] is therefore quite likely to differ between different vocal tract sizes (e.g., gender) and shapes (e.g., tongue positions)', whilst the second 'corresponds to changes in those resonance frequencies over the course of a vowel token, broadly reflecting the extent and direction of change.'

Recent sociophonetic work on vowel monophthongs has also begun to show the value of using measures which capture the dynamics of formant trajectories, also known as Vowel Inherent Spectral Change (VISC; see e.g. Morrison and Assmann 2013). The work of Jacewicz and colleagues (e.g. Jacewicz et al. 2011) found that vowel change in three American dialects for nominal monophthongs of BIT, BED, BAD, results in shifts in the position in the vowel space, and variation in formant dynamics; that of Haddican et al. (2013) also used formant trajectories to unpick the role of social-indexicality in changes to GOAT and GOOSE in real- and apparent-time in York. Docherty et al. (2015)'s study concentrates of nine monophthongs and diphthongs in west Australian English, and focussed on method. Their comparison of two static acoustic measures from F1 and F2, temporal midpoint and target estimate from formant maximum, and one dynamic (SSANOVA on formant tracks) highlights the numerous *additional* differences across all nine vowels revealed by the dynamic method.

Dynamic acoustic analysis of consonant sounds for which acoustic transitions are integral, such as glides and /r l/, is increasingly common (e.g. Carter and Local 2007, Stuart-Smith et al. 2015a, on liquids). Kirkham et al. (2019) use Generalized Additive Mixed Models (GAMMs) to uncover dynamic differences in laterals in the neighbouring Manchester and Liverpool dialects. GAMMs take formant tracks as input to advanced mixed models which provide visual representations of modelled tracks, and significance testing for potential differences from specified fixed factors (currently up to two-way interactions; for GAMMs, see Sóskuthy 2017). Dynamic analysis of obstruents is much less usual. Here DCT analysis was used to look at /s/ and gender over time in Glasgow for the following reasons:

- sibilants are quite like vowel monophthongs in that they have both target-like frequency ranges of spectral energy, and they also show some variation in spectral energy over their timecourse. DCT analysis gives measures which capture both aspects of sibilants together (Harrington et al., 2018; Watson and Harrington, 1999).

- DCT analysis can be applied to tracks of variable length, without needing to make arbitrary decisions about how many points to take for a track, or time normalization (Watson and Harrington, 1999)
- DCT coefficients are continuous, mathematically-independent measures, amenable to linear mixed modelling to test for the influence of fixed and random factors, separately and in interaction.
- agent-based modelling of sound change which reduces acoustic trajectories, also for /s/-retraction, to a three-point multidimensional space using DCT analysis is proving effective (Harrington et al., 2018; Harrington and Schiel, 2017). This study provides the first corroboration of DCT analysis for identifying sociophonetic factors governing the realization of sibilants in naturally-occurring, casual, conversational vernacular speech.

2.2 Capturing the acoustics of /s/

The acoustic spectrum of /s/ reflects the resonances of the cavities behind and especially in front of the constriction made by the tongue tip/blade close to the alveolar ridge, and the shape and nature of the constriction itself, as the jet of air is forced through the constriction and then strikes the surfaces of the teeth as it leaves the vocal tract (Johnson, 2003). Those of /ʃ/ differ in the relative size of the front cavity, being made bigger/longer with the retraction of the constriction, the length and grooving of the constriction, and in English, lip-rounding. Acoustic variation relates to the changes to the size of the front cavity and the shape of the constriction, though these quasi-articulatory parameters are rather abstract, and actual mapping back onto articulation in the absence of articulatory data is tricky (Sundara, 2005).

The spectrum of /s/ shows peaks and troughs which can be captured in measures of spectral *Peak* and *(front)Slope*, the latter reflecting the low frequency ‘shoulder’ of energy visible on the spectrogram below the main bands of high frequency energy; Jesus and Shadle (2002); Reidy (2015); Figure 1. Moments analysis (Forrest et al., 1988) models the spectrum as a normal distribution, from which the main distributional properties, or ‘moments’, are calculated, and have been shown to distinguish fricative place of articulation (Jongman et al., 2000). The first moment—the *Centre of Gravity (CoG)*—captures the mean frequency of the spectrum, and is modelled here. Peak and CoG frequencies correlate with changes in front cavity, with smaller/shorter cavities showing higher Peak and CoG, and

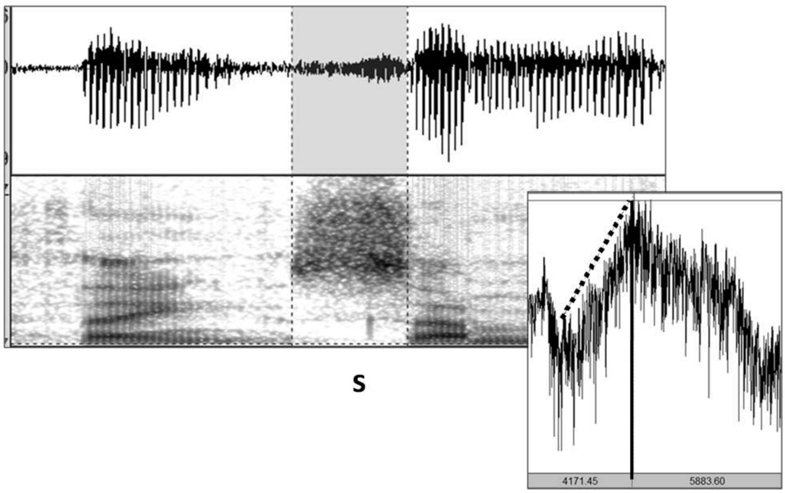

Figure 1: Long-term average spectrum (lower right) showing peak (solid line) and front slope (dashed line) for /s/ said by a working-class woman in the phrase, ‘I think some of the’. /s/ is shaded on the waveform/spectrogram (back).

vice versa; Slope relates to subtle changes in the shape and nature of the constriction, with more /ʃ/-like sounding sibilants showing higher Slope values (c.f. Sundara 2005).

2.3 Dynamic analysis of /s/ using DCT

Reidy (2016)’s cross-linguistic study of word-initial /s ʃ/ demonstrates how dynamic differences distinguish Japanese and English /s/. Reidy used polynomial growth-curve analysis on auditorily-transformed ERB spectral peaks from spectra taken at 17 points across the sibilants. English and Japanese sibilants differed from each other in overall level of spectral frequency and spectral shape. But likely the most interesting finding for sociophonetics is that Japanese and English /s/ differed not in Peak frequency, but in trajectory shapes. Subsequent work on Australian English (Stevens and Harrington, 2016) also showed qualitative differences in CoG trajectories taken across /s ʃ/ according to onset structure.

Figure 2 shows a DCT analysis for the /s/ token in Figure 1. The 9-point CoG track is compressed into three DCT coefficients (Watson and Harrington, 1999). The first coefficient, CoGk0, gives the *mean* value for the track, and so is similar

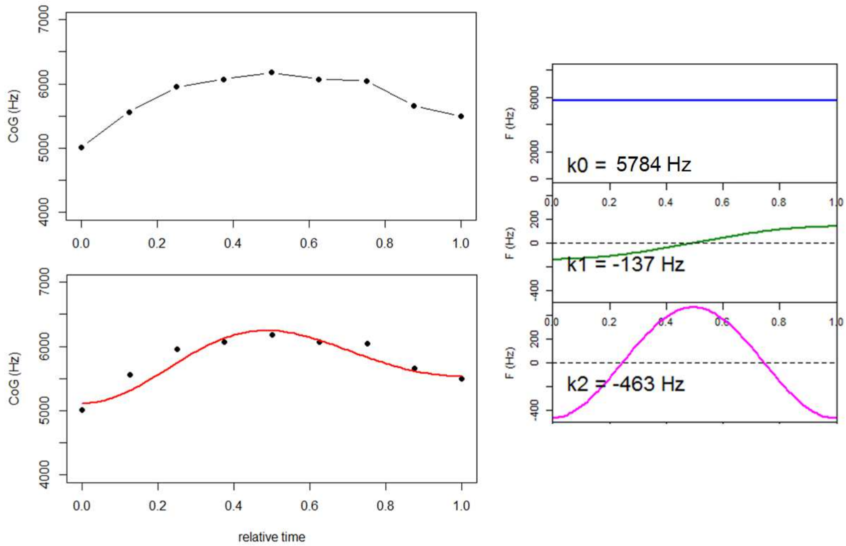

Figure 2: DCT analysis for /s/ in *some* shown in Figure 1. Upper left: time-normalized track of CoG measures. Right: Plots and values of the first three DCT coeffients, k0 (mean), k1 (slope) and k2 (curvature) of the CoG track. Lower left: DCT-smoothed CoG track.

to the static CoG measure taken from a long-window spectrum. The other two coefficients capture aspects of the time-varying shape of the trajectory. CoGk1 gives the degree and direction of the *slope* of the CoG trajectory, a negative value corresponds to an increase of CoG value across the trajectory (as a cosine it has an inverse relationship with the variable value). CoGk2 reflects *curvature* of the trajectory typical of sibilants, a negative value indicates more a humped-shape trajectory.

2.4 Gender in Glasgow /s/

Glasgow dialect has long been known for showing an auditorily-retracted /s/ (Macafee, 1983), likely from tip-raising, which increases the size of the front cavity (Johnson, 2003). Stuart-Smith (2007) analysed spectral Peak and Slope of /s/ in read wordlists from 32 speakers, stratified by age, gender and social class, recorded in the 1990s. Male and female speakers were generally different, but working-class girls additionally showed significantly lower frequency /s/ than all other female groups, such that they clustered with the men and the boys. Previous studies had consistently shown English female speakers with higher overall frequencies, assumed to arise from a smaller front cavity (Flipsen Jr et al., 1999). The Glasgow results evidenced the social construction of gender for /s/ production, since the working-class girls were clearly using articulatory strategies to overcome physiological constraints to produce acoustically lower frequency /s/. This was most likely to distinguish themselves from their middle-class counterparts, similar to Eckert (2000)'s polarization of 'burned out' Burnout from Jock girls. Similar results have been found for numerous different contexts and instantiations of social gender (Levon et al., 2017).

/s/ is thought to be a stable gender marker in Glasgow (Stuart-Smith et al., 2007). However, the social construction of gendered male and female roles has changed over the twentieth century, in conjunction with shifts in social and geographical mobility engendered during both World Wars and the latter half of the 20th century. Being children in the 1970s was rather different from in the 1990s: if social constructions of gender shifted over time in Glasgow, are these reflected in gendered changes in /s/? In particular, is lowered /s/ in working-class girls a recent development? And, if /s/ is changing, does this entail a shift in /ʃ/ too? These questions are summarized in the core research question here: *Does a dynamic acoustic representation improve our understanding of how social gender relates to sound change in /s/ and /ʃ/ in spontaneous Glaswegian vernacular over time?*

3 Method

All instances of word-initial, stressed /s ʃ/ were extracted from interviews and casual conversations from 32 working-class, Glaswegian speakers, from the electronic, automatically-segmented, LaBB-CAT *Sounds of the City* corpus (Fromont and Hay 2012; Stuart-Smith et al. 2017). The sample is structured by gender, presenting as ‘female’/‘male’, and real- and apparent-time, with recordings made in the 1970s and 1990s, from middle-aged (40-55 years) and younger (10-17 years) speakers, giving four *decades of birth*: 1920s, 1940s, 1960s, 1980s. ‘real-time’ comparisons are made on the basis of year of recording (so 1920s vs 1940s, 1960s vs 1980s), ‘apparent-time’ by comparing older and younger speakers within 1970s or 1990s (so 1920s vs 1960s, 1940s vs 1980s). The 1990s-recorded speakers are the same as those whose read speech was analysed in Stuart-Smith (2007).

The static measures were calculated in Praat (Boersma and Weenink, 2013) from spectra initially estimated using the *To spectrum* function (10ms Hamming window) across the central 70% duration of the sibilants. The sound files were downsampled to 22kHz, and high-/low-pass filtered at 1kHz and 11kHz, to remove extraneous low frequency energy and to prevent aliasing respectively, giving a frequency range of 1-11kHz across which CoG and Peak could be derived. An additional precaution was taken to eliminate erroneous tokens from wrong force-alignment and segmentation error, by removing all tokens showing perseverative voicing. CoG and Spread were calculated from the spectra, and then Peak and Slope measures were taken after conversion of the spectra using *To LTAS(1-to-1)*. Slope was measured using *LTAS:Get slope* over 1-4kHz frequency range (adapted from Jesus and Shadle 2002’s low frequency slope).

The dynamic measures were calculated using dct() in the emuR() package, from tracks of CoG (etc) measures calculated in Praat from a sequence of 10ms Hamming-window spectra, with no overlap, again across the central 70% sibilant portion, downsampled and filtered as before. An additional constraint was that only sibilants with durations $>$ 70ms were analysed. This, together with pruning of very low frequency tokens with CoG $<$ 2,4kHz, reduced the initial 7083 tokens to the subset dataset of 3392 tokens analysed here. While spectral Peak often gives a better static representation of social differences than CoG (e.g. Stuart-Smith 2007), CoG tracks showed less local variability across the sibilants than Peak tracks. Here, results are presented for CoG and Slope, which also showed better discrimination of linguistic and social factors than Spread.

All measures were modelled using lme4() and lmerTest() in R R Core Team (2013). Dependent variables were static CoG and Slope, and CoGk0-k2, Slopek0-k2 for the trajectories. Fixed factors were: sibilant *(log)Duration*, *fol-*

lowing phonological context, *Sound*, *Gender*, *Decade of birth*, and all interactions for *Sound*Gender*Decade of birth.* Models also included random intercepts for *Speaker, Word*, and random slopes for *Decade* by *Word*, and *Following context* by *Speaker.*

Before modelling, all variables were conceptually ‘centred’, either by scaling for (log)Duration, or sum contrasts coding, which reduces collinearity in the models and helps them to converge (Sonderegger et al., 2018). A grand mean is calculated across the levels of a factor, and then relates n-1 of the levels to that mean. *Gender* and *Sound* have one coded level: Gender1=female, Sound1=/ʃ/, male and /s/ are inferred. *Following phonological context–Post* has 2 coded levels: Post1= spread close front vowels vs grand mean of *Post*, Post2=central vowels; rounded back vowels/consonants is inferred. *Decade of birth* has 3 coded levels: Decade1=1920s(70-M), Decade2=1940s(90-M) and Decade3=1960s(70-Y), each vs grand mean of *Decade*; 1980s(90-Y) is inferred.

4 Results

The model summaries given in Tables **??** show that the linguistic factors of *Duration* and *following phonological context* significantly constrain the spectral characteristics, static and dynamic, of both sibilants as expected. For example, the spectral frequency of CoG and CoGk0, is higher for sibilants before spread close front, and central vowels, and lowered by the coarticulatory influence of following rounded back vowels and/or consonants (Baker et al. 2011). Here we focus on results for *Sound*, *Gender* and *Decade*, based on the relevant significant model estimates which are given in the Tables, plus targeted pairwise comparisons between levels inferred by the coding, and so which are not represented in the estimates given in the model summaries. The reporting of statistics for the relative contribution of significant terms and interactions to the models is limited to higher-order interactions.

4.1 First look: Static reflections

The significant three-way interaction of *Sound*Gender*Decade* (F(3,1977.8)=2.93, p=0.03) is shown in Figure 3. As expected, /s/ has higher CoG frequencies than /ʃ/, and female speakers show overall higher frequency CoGs than males (p=, but only for /s/ (Holliday et al. 2015). But we also see shifts by time. /ʃ/ shows a real-time increase for men (t=2.51, p=0.02), and apparent-time for 70s-recorded speakers

Factors	static CoG	mean(CoGk0)	slope(CoGk1)	curvature(CoGk2)
Intercept	4,691***	6,323***	31	−202***
Duration	−34**	−301***	−69***	71***
Sound1(/s/)	624***	767***	−50**	−59***
Post1(spread-front-vowels)	139***	174***	−100***	
Post2(central-vowels)	74*	84	−96***	
Gender1(female)	312***	366***		−13
Decade1(1920s)	−269	−395*	16	−35
Decade2(1940s)	166	259	7	28
Decade3(1960s)	161	193	7	−57*
Sound1:Gender1	223***	246***		−38***
Sound1:Decade1	18	−8	−2	
Sound1:Decade2	160***	193***	−30	
Sound1:Decade3	−90*	−155**	118**	
Gender1:Decade1	46	108		
Gender1:Decade2	−119	−100		
Gender1:Decade3	58	−45		
Sound1:Gender1:Decade1	−71*	−91*		
Sound1:Gender1:Decade2	53	89*		
Sound1:Gender1:Decade3	45	107*		
Note:			*p<0.05; **p<0.01; ***p<0.001	

Table 1: Estimates for modelling of static and dynamic measures for CoG.

Factors	static Slope	mean(Slopek0)	slope(Slopek1)	curvature(Slopek2)
Intercept	36.2***	41.7***	−0.3	−0.9***
Duration	−1.1***	−2.5***	−0.5***	0.3***
Sound1(/s/)	−4.6***	−6.1***	0.7***	0.5***
Post1(spread-front-vowels)	−0.7*	−0.5	−0.6***	0.3**
Post2(central-vowels)	0.2	−0.2	−0.2	−0.1
Gender1(female)	−3.1***	−4.3***	−0.2	0.1
Decade1	−0.6	−0.9	−0.3	0.01
Decade2	0.4	0.7	−0.1	0.5**
Decade3	−3.1**	−3.8*	0.1	0.1
Sound1:Gender1	−1.5***	−1.5***	0.01	0.2**
Sound1:Decade1	0.5	0.4	0.3	−0.2
Sound1:Decade2	−1.0***	−1.4***	0.1	−0.1
Sound1:Decade3	−0.5	−0.4	−0.4	−0.1
Gender1:Decade1	−1.0	−0.8	0.1	0.3*
Gender1:Decade2	−0.3	−0.6	0.8***	−0.02
Gender1:Decade3	2.2*	1.9	−0.2	−0.2
Sound1:Gender1:Decade1	1.1***	0.9**	0.03	−0.5***
Sound1:Gender1:Decade2	0.2	0.004	−0.6***	−0.01
Sound1:Gender1:Decade3	−0.02	−0.2	0.2	0.1
Note:			*p<0.05; **p<0.01; ***p<0.001	

Table 2: Estimates for modelling of static and dynamic measures for spectral Slope.

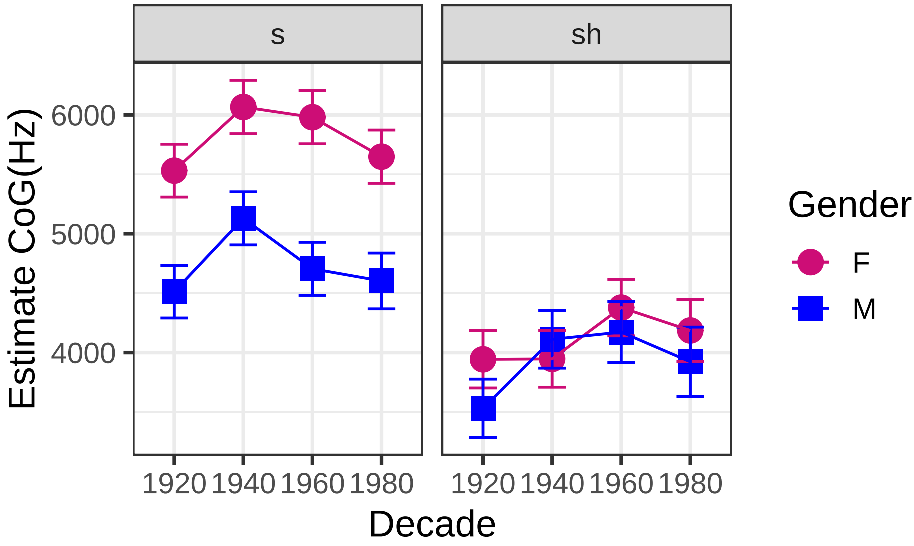

Figure 3: Model estimates of static CoG(Hz) plotted by sound, gender and decade of birth.

(females: t=2.87, p=0.008; males: t=2.63, p=0.02). For /s/, male speakers show generally low CoG values, with a rise in 1940s-born males (1940s vs 1920s: t=3.8, p=0.005; 1980s vs 1940s: t=-2.81, p=0.008). Females show a real- and apparent-time rise in CoG values, and then a reversal in the girls born in the 1980s, who pattern with the women born in the 1920s (1980s vs 1940s: t=-2.97, p=0.007; 1980s vs 1960s: t=-2.32, p=0.03).

Slope also shows a significant interaction of *Sound*Gender*Decade* (F(3,2000.9)= 5.7, p= 0.0007); Figure 4. But the pattern is rather different from that of CoG, showing socially-salient shifts not only to the ‘cavity size’ parameter of the sibilant, but also to its constriction shape. As expected, overall /ʃ/ shows higher Slopes than /s/, males have higher Slope values than females, and especially for /s/. But we also see changes to both sibilants. For /ʃ/, there is a real- and marginal apparent-time rise in Slope for female speakers (1920s vs 1940s: t=-3.13, p=0.004, 1920s vs 1960s: t=-1.8, p=0.08; 1960s vs 1980s: t=-1.93, p=0.06), and a real-time rise for boys (1960 vs 1980s: t=-2.4, p=0.02). For /s/, there is a real-time rise for girls (1960s vs 1980s: t=-3.24, p=0.003), but a substantial real-time increase in Slope for the boys (1980s vs 1960s: t=4.59, p<0.0001, 1980s vs 1940s: t=4.36, p=0.0001), such that the boys born in the 1980s no longer differ in Slope between /s/ and /ʃ/.

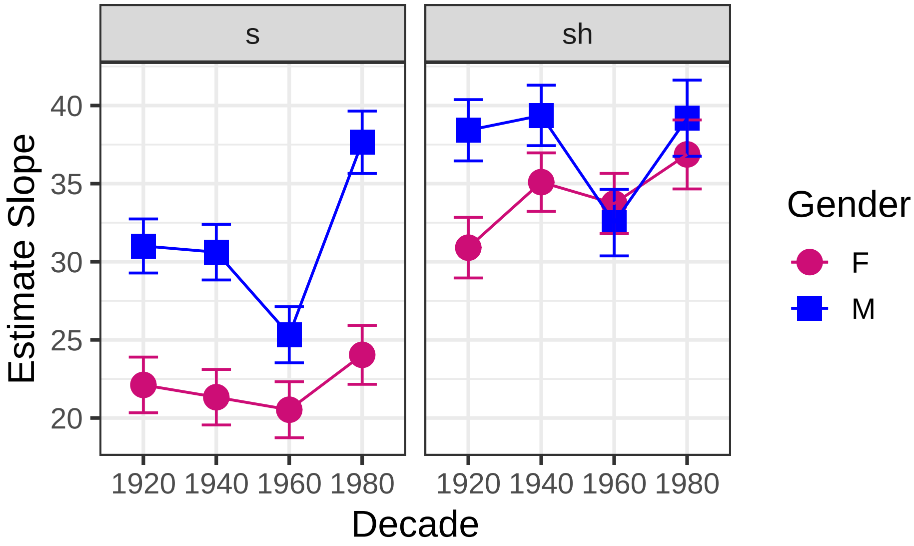

Figure 4: Model estimates of static spectral Slope plotted by sound, gender and decade of birth.

5 Second look: Dynamic reflections

The results for CoGk0, reflecting the mean of the CoG trajectory across the fricative, show similar, if more significant, effects for prosodic, linguistic, and social factors, than the static CoG measures, with the same patterning (*Sound*Gender*Decade*: F(3,2010.6)=4.44, p=0.004). CoGk0 is strongly correlated with CoG: r=0.91; t(3390)=125.18, p<0.0001.

The higher coefficients show that dynamic shifts to cavity size across the fricatives carry key additional prosodic, linguistic, and social information. CoGk1 reflects changes in magnitude and direction of the slope of the CoG trajectory over the fricative. /ʃ/ shows higher CoGk1 than /s/, corresponding to a greater lowering of CoG frequency across the trajectory for /ʃ/ than /s/, except for the 1960s-born adolescents (*Sound*Decade* F(3,81.1)=2.87, p=0.04).

Both sibilants show real- and apparent-time increases in CoGk2 (*Decade*: F(3,29.07)=3.01, p=0.046), reflecting decreasing humpiness in the curvature of the CoG trajectory over time. We also find a *Sound*Gender* interaction for CoGk2 (F(1,3055.19)=15.56, p<0.0001), such that male speakers show no difference in

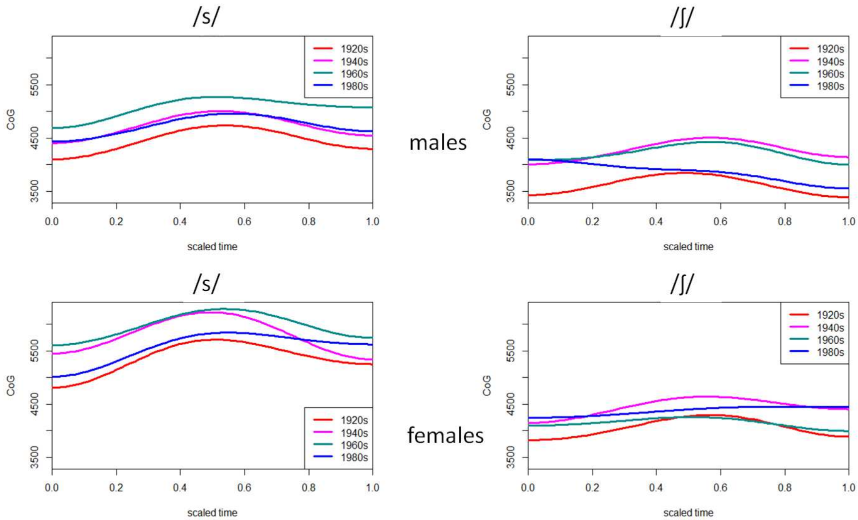

Figure 5: CoG tracks by sound, gender and decade of birth based on LME estimates of DCT CoGk0-k2.

curvature by sound, whilst females show much lower CoGk2 for /s/ than /ʃ/, so more humped /s/ and less humped /ʃ/.

The DCT results for the CoG tracks are summarized visually in the smoothed estimate tracks in Figure 5. Differences in spectral shape along with those of overall mean CoG frequency are clearly visible across the two sibilants, for male and female speakers, and by decade of birth: both 1980s-born girls and boys drop CoG frequency for /s/, but girls show a different CoG trajectory at the end of the fricative.

Slopek0 reflects the mean value of the spectral Slope trajectory, and so changes in constriction shape over the fricative. This measure is strongly correlated with static Slope ($r=0.85$; $t(3390)=93.09$, $p<0.0001$), and shows a similar patterning of significant prosodic, linguistic, and sociolinguistic effects as for static Slope; Table 2.

The higher Slope coefficients again reveal additional dynamic differences, reflecting that shifts in constriction shape carry prosodic, linguistic and social information. Slopek1 reflects the size and direction of shifts in the 'shoulder' of low-frequency energy during sibilant production. The significant interaction of *Sound*Gender*Decade* for Slopek1 ($F(3,1888.93)=4.67$, $p=0.003$) reflects variation in changes to the Slope trajectory such that it is steeper for /s/ than /ʃ/,

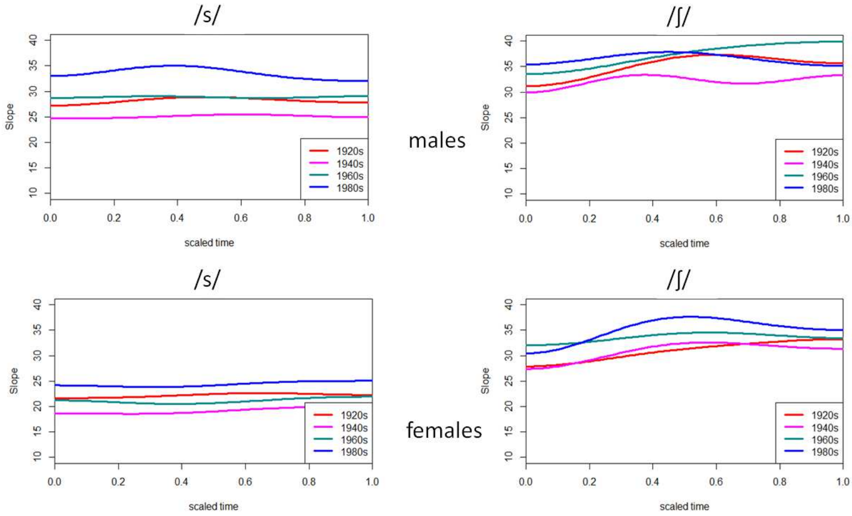

Figure 6: Slope tracks by sound, gender and decade of birth based on LME estimates from DCT CoGk0-k2.

which in turn over apparent-time increases for males (1940s vs 1980s: t=-2.26, p=0.01), and decreases for females (1940s vs 1980s: t=2.25, p=0.03).

The curved shape of the spectral Slope trajectory, how the 'shoulder' of low frequency rises and falls over the fricative, is captured by Slopek2. The significant interaction of *Sound*Gender*Decade* (F(3,1403.2)=4.88, p=0.002) reflects the less-humped Slope trajectory for /s/ than /ʃ/, with the reservation that over apparent-time, female speakers show more curved spectral Slope trajectories than males for /ʃ/ (1920s vs 1960s: t=2.39, p=0.02; 1940s vs 1980s: t=2.79, p=0.007).

Figure 6 gives a visual dynamic summary of differences in the spectral Slope trajectory over the course of the sibilants, by sound, gender and decade of birth. The nature of the sibilant constriction appears to have experienced changes which are structured by linguistic and social factors, and which impact on the sibilants' mean and spectral shape. As for the dynamic CoG results, we again see that both girls and boys born in the 1980s are doing something, but differently from each other, and this is even clearer from the dynamic perspective.

6 Discussion

6.1 The impact of changing spectral measures for 'viewing' /s/

Our study demonstrates how the selection of static spectral measures, CoG or Slope, fundamentally changes our interpretation of the sociophonetics of /s/. Viewed only from CoG, a choice taken by many to capture key social-indexical properties of /s/, the inference here would be a real-time reduction in the gender difference for Glasgow speakers, as girls born in the 1980s revert to earlier vernacular norms, and in so doing, approximate lower, male, spectral frequency. When spectral Slope is added, a different picture emerges. We see in terms of constriction shape, it is the boys born in the 1980s who change the most: as girls lower their spectral CoG frequency approaching that of the boys, so the boys shift another aspect of their fricative production, effectively shifting their sibilants away from those of the girls (and so much that the boys' /s/ Slope equals /ʃ/).

The dynamic DCT analysis using Discrete Cosine Transformation adds another crucial perspective. This allows us to capture with the first coefficient, the main 'target' acoustic characteristics, reflecting variation in sibilant cavity size and constriction shape. The higher coefficients reveal that dynamic aspects of fricative production carry key additional prosodic, linguistic, and social information. This in itself is not surprising: the articulation of speech sounds is known to be complex and dynamic. The assumption that all linguistic and social information is captured in steady-state aspects, when the process of producing speech in interaction is also dynamic, seems odd. Rather, we might expect that if we wish to model sociophonetic variation from highly-uncontrolled casual conversation, using dynamic analysis may be more appropriate. Harrington and Schiel (2017) justify the use of DCT analysis to model the role of interaction for sound change, partly because the higher DCT coefficients capture both the 'inherently dynamic synchronic processes such as coarticulation and undershoot' arising from phonetic constraints, and changes to trajectory shape arising 'as a consequence of interactions between the agents'. Our results confirm that both 'target' and dynamic aspects of sibilant variants carry key social-indexical information for this community.

6.2 Gendered changes to Glaswegian /s ʃ/

Static CoG measures show that Glasgow /s/ is always a certain kind of (social) /s/. They also suggest a reversal to 'old' variants in the adolescent girls born in the 1980s. Stuart-Smith (2007) interpreted the read speech results from these same girls as matching male norms; the perspective of time shows that the girls may be returning to vernacular female norms. Including static spectral Slope measures suggest that the girls are shifting to 'new-old' /s/-variants, perhaps with a different gesture. They also show that—perhaps unusually—as the girls shift their /s/ variants, so the boys shift theirs, and far more so than the girls. Clearly the Glaswegian sibilants are not as stable as previously assumed: as /s/ becomes more 'esh-y', so /ʃ/ in turn becomes 'esh-ier', especially in female speakers. ʃ/ also emerges as socially informative in this dialect, and is undergoing change too.

The real- and apparent-time differences in sibilants are most evident in the speech of the adolescents born in the 1980s. Their childhoods coincided with the reformation of close-knit networks in the inner-city and peripheral housing estates, following the period of substantial upheaval of urban regeneration and socio-spatial changes to the cityscape (Stuart-Smith et al., 2007). Stuart-Smith (2007) showed these same working-class girls frozen in a snapshot, polarised from the middle-class girls such that they patterned with men. Now we can infer that this was the result of a fairly recent change in /s/, likely enhanced by additional persona construction for reading the wordlist in the presence of the university fieldworker.

These changes to sibilants belong to a more general emergent local, socially-salient style used by adolescents born during this post-industrial period (Stuart-Smith et al., 2007). This style integrates phonological innovations (e.g. TH-fronting; Stuart-Smith et al. 2013) with a return to 'new-old' vernacular norms in fine-grained aspects of speech, such as /s/ here, and stop aspiration (Stuart-Smith et al., 2015b), and with increased use of Scots lexis, e.g. *hoose* for *house*; Stuart-Smith (2003). More generally, our initial hypothesis, that shifts in social gender over time might be reflected in subtle shifts in /s/ and /ʃ/ production is confirmed. /s/ in the girls born in the 1980s is less classically 'feminine', i.e. high-frequency. The boys' sibilants show clear evidence of gestural shift too. These results suggest close links between phonetic variation and social-indexical meaning for this vernacular community as its own social instantiations shifted over time.

7 Conclusions

This study returned to the Glaswegian speakers from the 1990s, whose read speech was analysed in Stuart-Smith (2007), and extended the analytical perspectives in terms of *time*, *phonological contrast*, and *spectral measures*. Focussing on spontaneous speech from working-class speakers, this study shows real- and apparent-time change for /s/ and /ʃ/ over an effective timespan of seventy years. Adding

/ʃ/ shows that both /s/ and /ʃ/ show gendered production. Our static analysis expanded from overall level of spectral frequency (CoG), reflecting shifts in front cavity size, to show that gendered changes are also taking place to subtle aspects of sibilant articulation, as reflected by spectral Slope. Our use of DCT analysis highlight that changes to gendered production of Glaswegian sibilants over time relate not only to ‘steady-state’ characteristics, but also to changes in spectral energy over the course of the fricatives themselves, and so to their dynamic characteristics.

It is clear that DCT analysis, just as demonstrated for dialectal variation in read vowels (Williams and Escudero, 2014), is also well-suited to sociophonetic investigation of sibilants in naturally-occurring speech. There is however further work to be done with respect to spectral analysis of sibilants more generally. Reidy (2015) reviewed different methods of calculating the spectrum itself, discrete Fourier transform (DFT) with and without pre-emphasis, and the multitaper spectrum. He found differences in the spectra, and hence the derived measures, but the same patterning for social and linguistic contrasts irrespective of spectral estimator. Again, the value of alternative spectral measures, such as those proposed by Koenig et al. (2013), needs to be considered for future sociophonetic studies. The impact of both spectral estimator and spectral measures on modelling dialectal and social factors for sibilants is the subject on our ongoing work using the substantial datasets being analysed in the SPADE project www.spade.arts.glasgow.ac.uk.

Finally, it is clear that changing the way we look at /s/, in terms of acoustic analysis, but also with respect to time and phonology, thoroughly changes our sociolinguistic interpretations. Just as for Alice, things do not appear as they did before, which in turn indicates the value of continually interrogating just how theoretical and analytical perspectives are influencing our conceptualisation of the very phenomena we are seeking to model (Baird, 2004).

Acknowledgment: This paper extends sketches published in early form in Stuart-Smith (2018, 2019). I am grateful: to the Leverhulme Trust for supporting the construction of the *Sounds of the City* corpus; to Jonathan Harrington for proper introduction to DCT; to Brian Jose, Rachel Macdonald, Tamara Rathcke, Claire Timmins, Wolf-Gerrit Fruh, for collaboration at different stages; to audiences at LabPhon15, P&P13, ICL20, NZILBB, Newcastle, QMUL, McGill, for feedback, and to the editors of this issue, Anne Fabricius and Nicolai Pharao.

References

Baird, D. (2004). *Thing knowledge: A philosophy of scientific instruments*. University of California Press, Berkeley.

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

Boersma, P. and Weenink, D. (2013). Praat: doing phonetics by computer.

Carter, P. and Local, J. (2007). F2 variation in Newcastle and Leeds English liquid systems. *Journal of the International Phonetic Association*, 37(2):183–199.

Dijkstra, E. (2012). *Selected writings on computing: A personal perspective*. Springer, Berlin.

Docherty, G. J., Gonzalez, S., and Mitchell, N. (2015). Static vs dynamic perspectives on the realization of vowel nucleii in West Australian English. In *International Congress of Phonetic Sciences (ICPhS)*.

Eckert, P. (2000). *Linguistic Variation as Social Practice*. Blackwell, Oxford.

Eckert, P. (2003). Elephants in the room. *Journal of Sociolinguistics*, 7(3):392–397.

Flipsen Jr, P., Shriberg, L., Weismer, G., Karlsson, H., and McSweeny, J. (1999). Acoustic characteristics of/s/in adolescents. *Journal of Speech, Language, and Hearing Research*, 42(3):663–677.

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: preliminary data. *The Journal of the Acoustical Society of America*, 84(1):115–123.

Foulkes, P. (2006). Sociophonetics. In Brown, K., editor, *Encyclopedia of language and linguistics*. Elsevier Press, Amsterdam.

Foulkes, P., Scobbie, J. M., and Watt, D. (2010). Sociophonetics. In Hardcastle, W. J., Laver, J., and Gibbon, F. E., editors, *The Handbook of Phonetic Sciences*, pages 703–754. John Wiley & Sons, 2nd edition.

Fromont, R. and Hay, J. (2012). LaBB-CAT: An annotation store. In *Proceedings of the Australasian Language Technology Association Workshop 2012*, pages 113–117.

Haddican, B., Foulkes, P., Hughes, V., and Richards, H. (2013). Interaction of social and linguistic constraints on two vowel changes in northern England. *Language Variation and Change*, 25(03):371–403.

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

Harrington, J. and Schiel, F. (2017). /u/-fronting and agent-based modeling: The relationship between the origin and spread of sound change. *Language*, 93(2):414–445.

Hay, J. and Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, 36:89–103.

Holliday, J. J., Reidy, P. F., Beckman, M. E., and Edwards, J. (2015). Quantifying the robustness of the English sibilant fricative contrast in children. *Journal of Speech, Language, and Hearing Research*, 58(3):622–637.

Huggett, J. (2017). The Apparatus of Digital Archaeology. *Internet Archaeology*, (44).

Jacewicz, E., Fox, R. A., and Salmons, J. (2011). Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of Phonetics*, 39(4):683–693.

Jesus, L. M. T. and Shadle, C. H. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, 30:437–64.

Johnson, K. (2003). *Acoustic and Auditory Phonetics*. Blackwell, Oxford, 2nd edition.

Jongman, A., Wayland, R., and Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 108(3):1252–1263.

Kirkham, S., Nance, C., Littlewood, B., Lightfoot, K., and Groarke, E. (2019). Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English. *The Journal of the Acoustical Society of America*, 145(2):784–794.

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

Levon, E., Maegaard, M., and Pharao, N., editors (2017). *The Sociophonetics of /s/*. Linguistics.

Macafee, C. (1983). *Varieties of English around the world: Glasgow*. Benjamin, Amsterdam.

Morrison, G. S. and Assmann, P., editors (2013). *Vowel Inherent Spectral Change*. Springer, Berlin.

Podesva, R. J. and Kajino, S. (2014). Sociophonetics, gender, and sexuality. In Ehrlich, S., Meyerhoff, M., and Holmes, J., editors, *The Handbook of Language, Gender, and Sexuality*, pages 103–122. Wiley Online Library.

R Core Team (2013). R: A Language and Environment for Statistical Computing.

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.

Reidy, P. F. (2016). Spectral dynamics of sibilant fricatives are contrastive and language specific. *The Journal of the Acoustical Society of America*, 140(4):2518–2529.

Sonderegger, M., Wagner, M., and Francisco, T. (2018). *Quantitative Methods for Linguistic Data*. Montreal, v. 1.0 edition.

Sóskuthy, M. (2017). Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. *arXiv preprint arXiv:1703.05339*.

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

Stuart-Smith, J. (2003). The Phonology of Modern Urban Scots. In Corbett, J., McClure, D. J., and Stuart-Smith, J., editors, *The Edinburgh Companion to Scots*, pages 110–137. Edinburgh University Press.

Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. In Cole, J. and Hualde, J., editors, *Change in Phonology: Papers in Laboratory Phonology 9*, pages 65–86. Mouton de Gruyter.

Stuart-Smith, J. (2018). Sound perspectives? Speech and speaker dynamics over a century of Scottish English. In Mesthrie, R. and Bradley, D., editors, *The Dynamics of Language: Plenary and focus lectures from the 20th International Congress of Linguists*, pages 74–92. University of Cape Town Press, Cape Town.

Stuart-Smith, J. (2019). A tale of one city: Phonological variation and change over 100+ years of Glasgow English. In Glain, O., editor, *English-speaking towns and cities: Memories and narratives. Proceedings of the Workshop, St Etienne, 20-21 October 2016*. St Etienne.

Stuart-Smith, J., José, B., Rathcke, T., Macdonald, R., and Lawson, E. (2017). Changing Sounds in a Changing City: An Acoustic Phonetic Investigation of Real-Time Change over a Century of Glaswegian. In Montgomery, C. and Moore, E., editors, *Language and a Sense of Place: Studies in Language and Region*, pages 38–65. Cambridge University Press, Cambridge.

Stuart-Smith, J., Lennon, R., Macdonald, R., Robertson, D., Soskuthy, M., José, B., and Evers, L. (2015a). A Dynamic Acoustic View of Real-Time Change in Word-Final Liquids in Spontaneous Glaswegian. In *XVII International Congress of Phonetic Sciences (ICPhS 2015)*.

Stuart-Smith, J., Pryce, G., Timmins, C., and Gunter, B. (2013). Television can also be a factor in language change: Evidence from an urban dialect. *Language*, 89(3):501–536.

Stuart-Smith, J., Sonderegger, M., Macdonald, R., and Rathcke, T. (2015b). The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*, 33(4):505–459.

Stuart-Smith, J., Timmins, C., and Tweedie, F. (2007). Talkin' 'Jockney'? Variation and change in Glaswegian accent. *Journal of Sociolinguistics*, 11(2):221–260.

Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *The Journal of the Acoustical Society of America*, 118(2):1026–1037.

Watson, C. I. and Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *The Journal of the Acoustical Society of America*, 106(1):458–468.

Williams, D. and Escudero, P. (2014). A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *The Journal of the Acoustical Society of America*, 136(5):2751–2761.