Fine phonetic variation and sound change: A real-time study of Glaswegian

The grant: a brief history:

- *Duration:* October 2011-December 2014
- *Budget*: £235,682
- *Research team:*
  - Investigators: Stuart-Smith (English Language, Glasgow; 10hours/week); Timmins (Speech and Language Therapy, Strathclyde; 1hour/week); Torsney (2hours/week), then Evers and Neocleous (Mathematics and Statistics, Glasgow; 1hour/week).
  - RAs: RA1: Rathcke (years 1-2), then Macdonald; RA2: José.
  - LaBB-CAT consultant: Fromont (NZLBB, ChristChurch, NZ)
  - Collaborators: Sonderegger (McGill, Canada); Sóskuthy (York, UK); Rathcke (Kent, UK).
  - Advisers: colleagues at LANCHART (Copenhagen), IPS (Munich).
- *Research site:* Glasgow University.
- *Website:* http://soundsofthecity.arts.gla.ac.uk/

Objectives:

The main aim of our project was to investigate empirically the role of fine phonetic variation in sound change in a speech community over a period of time. Specifically, we used continuous acoustic phonetic analysis to observe Glaswegian vernacular over time to consider the following theoretical issues:

- ‘stable’ vs ‘changing’ sounds
- gradient vs abrupt change
- linguistic and social factors in sound change

This entailed methodological objectives, to:

- build a digital real-time corpus of spoken Glaswegian vernacular
- manage technical issues of recording quality and acoustic analysis
- overcome issues of comparability across recordings
- achieve optimal statistical modelling

Research activity:

1. The Sounds of the City corpus

We built a real-time digital spoken corpus of 142 speakers comprising over 700,000 words/60 hours of speech, with a real- and apparent-time span of around 100 years (Table 1). We also discovered an additional six recordings in the British Library’s *Berliner Lautarchiv* collection, of Glaswegian/Scottish soldiers recorded in German prisoner of war camps in WWI (Figure 1). This provided an unexpected opportunity to extend our real-time scope to the early 20th century.

Decade of Recording	Old 67-90 (Decade of Birth)	Middle-aged 40-55 (Decade of Birth)	Young 10-17 (Decade of Birth)
1970s	4 f, 6 m (1890s)	7 f, 7 m (1920s)	4 f, 8 m (1960s)
1980s	6 f, 6 m (1900s)	4 f, 12 m (1930s)	2 f, 5 m (1970s)
1990s	6 f, 6 m (1910s)	6 f, 6 m (1940s)	6 f, 6 m (1980s)
2000s	6 f, 6 m (1920s)	6 f, 5 m (1950s)	6 f, 6 m (1990s)

Table 1: Real- and apparent-time structure of the main Glasgow corpus; an additional 6 speakers born in the 1890s and recorded in 1916/17 are also available. Recordings are predominantly of unprepared spontaneous speech from sociolinguistic surveys, oral histories and broadcast media.

Figure 1: Wilhelm Doegen recording a speaker in a German Prisoner of War camp (*The Doegen Records Web Project,* http://doegen.ie/about; © Humboldt-Universität,Berlin; Stuart-Smith et al to appear).

2. Phonological variables

We analysed variables assumed to be potentially:

(1) stable:

- bimoraic vowels, /i e a o ɔ ʉ/
- initial /l/, e.g. *like, look*
- the Scottish Vowel Length Rule (SVLR)
- aspiration (VOT) in stops, /p t k b d g/

(2) changing:

- coda /r/, e.g. *car, beer*
- coda /l/, e.g. *kill, shelf*

3. Methodological challenges

(a) Phonemic segmentation

Manual segmentation of spontaneous speech is time-consuming. We carried out full automatic segmentation of the corpus using HTK forced alignment in LaBB-CAT. Inspection showed that the forced alignment was surprisingly successful for vernacular speech and was not dependent on recording type (Figure 2).

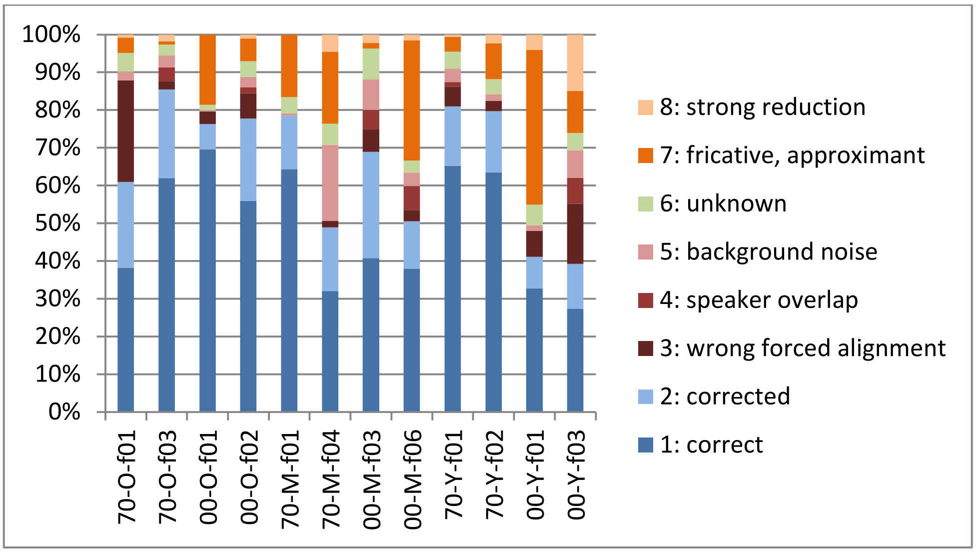


Figure 2: Proportions of tokens coded according to manual inspection of automatic Voice Onset Time predictions for 4491 voiced and voiceless stops (each bar represents a speaker, arranged according to decade of birth, from 70-O born in the 1890s to 00-Y born in the 1990s). 67% of the predictions were correct or easily corrected (Codes 1/2); Code 3 shows the proportion of errors arising from wrong forced alignment (mean: 7%). Stuart-Smith et al (2015).

(b) Recording quality

Our recordings vary in recording quality and background noise, both of which are challenging for obtaining reliable acoustic analyses. Digitization of tape recordings can also lead to spectral degradation (Figure 3).

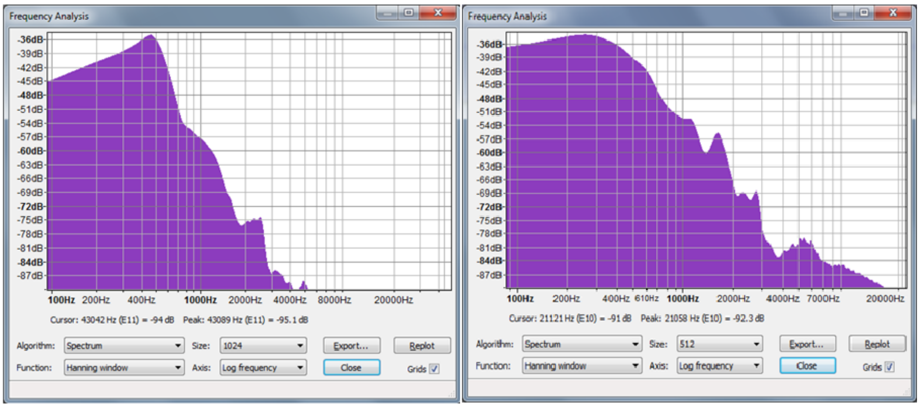


Figure 3: Examples of spectral energy distribution in 1970-G1-m01 recording before (left) and after (right) redigitization with corrected positioning of tape heads (with thanks to Gert Hansen, [LANCHART](), Copenhagen).

Analyses of noise in terms of Signal-to-Noise Ratio (SNR) and Spectral Balance (SB) revealed a real- and apparent-time structure to our recordings purely in terms of noise (Figure 4). Both kinds influence formant measures, for example, recordings with worse SNR also show statistically higher F1 values, thus confounding sound change with recording noise. The Lobanov method of vowel normalization emerged as the most effective way of eliminating the impact of noise on formant measures (Rathcke et al in preparation).

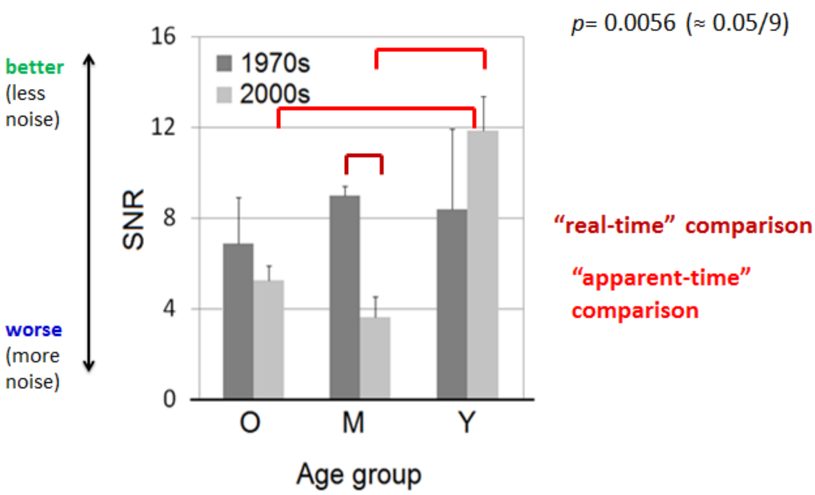


Figure 4: Real- and apparent-time differences in Signal-to-Noise Ratio in a subset of the Glasgow corpus; brackets indicate significant differences in Linear Mixed Effects modelling. (70 = 1970s recordings; 00 = 2000s; O = old; M = middle-aged; Y = young).

(c) Automated measures

Manual labelling/correction of acoustic characteristics in spontaneous speech is time consuming. We adopted (semi-)automated routines wherever possible to reduce time spent on manual annotation, and to increase token counts for robust statistical analysis. We also refined these procedures to optimize phonetic measures. Automatic formant measurement yielded 21,500 tokens of 6 vowels for 31 speakers; statistical procedures removed around a third of the data as outliers (Figure 5).

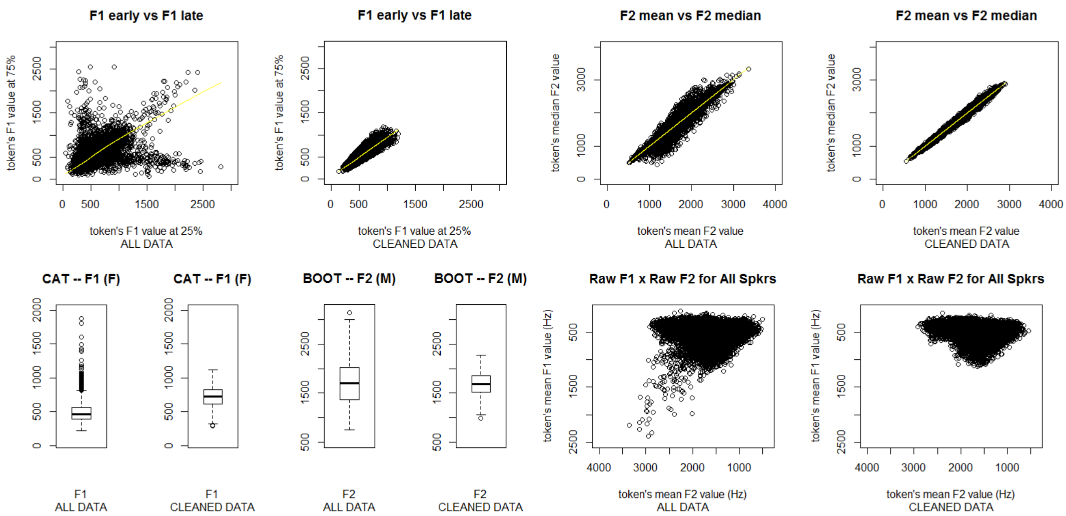


Figure 5: Vowel data before and after automatic data reduction (José et al in preparation)

With Sonderegger (McGill), we developed a swift semi-automatic procedure to extend his automatic classifier, AutoVOT, for analysing stops. We processed over 12,000 tokens from 23 speakers with substantially increased efficiency (Stuart-Smith et al 2015; Stuart-Smith et al to appear). We used Sóskuthy's new correction programme, Formant Edit, to obtain reliable dynamic acoustic views of coda /r/ and /l/ (Stuart-Smith et al to appear).

(d) Controlling for comparability

Our recordings range from oral history interviews to casual conversations. We adapted the LANCHART Discourse Context Analysis scheme to compare phonetic variation across the different genres and styles in our corpus, and tested it on a sample of 31 speakers. We found some effects of discourse context on vowel quality within particular age-time groups, but no significant impact on observed vowel changes over time, suggesting that comparisons across these diverse recordings are valid for fine phonetic variation (José and Stuart-Smith 2014).

4. Changes to the project plan

- The duration of corpus construction was longer than expected, as we needed to regularize and correct transcripts for automatic segmentation.
- Difficulties in locating recordings meant that we accepted shorter recordings for the 1980s.
- The corpus scope was extended by adding the *Berliner Lautarchiv* historical recordings.

Conclusions and achievements:

1. Corpus

- Our fully segmented digital corpus enabled us to carry out our intended analyses very quickly, and to develop successful collaborations for semi-automated phonetic analyses for stop aspiration (Sonderegger) and dynamic acoustic analyses of coda liquids (Sóskuthy).
- The corpus is also a flexible legacy resource, which is being extended with new recordings (2010s), linked with a new complementary corpus, and shared by other researchers.

2. Methodology

We have found solutions for methodological challenges, specifically, we have:

- dealt successfully with issues to do with digitization, recording quality/noise, and comparability
- used and developed (semi-)automated measures, including for dynamic acoustic analysis
- collaborated within and beyond the project team to find the most appropriate statistical analyses and ways of disseminating these statistical results (using Shiny)

3. Main finding: Most sounds appear to be changing in Glaswegian

We were surprised to discover evidence consistent with change not only for our 'change' variables (coda /r/: Figure 6; coda /l/: Figure 7), but also for most of our 'stable' variables (stop aspiration: Figure 8; initial /l/: Figure 9; vowel quality: Figure 10; vowel duration: Figure 11). We were therefore unable to compare variation in 'changing' variables with respect to 'stable' ones.

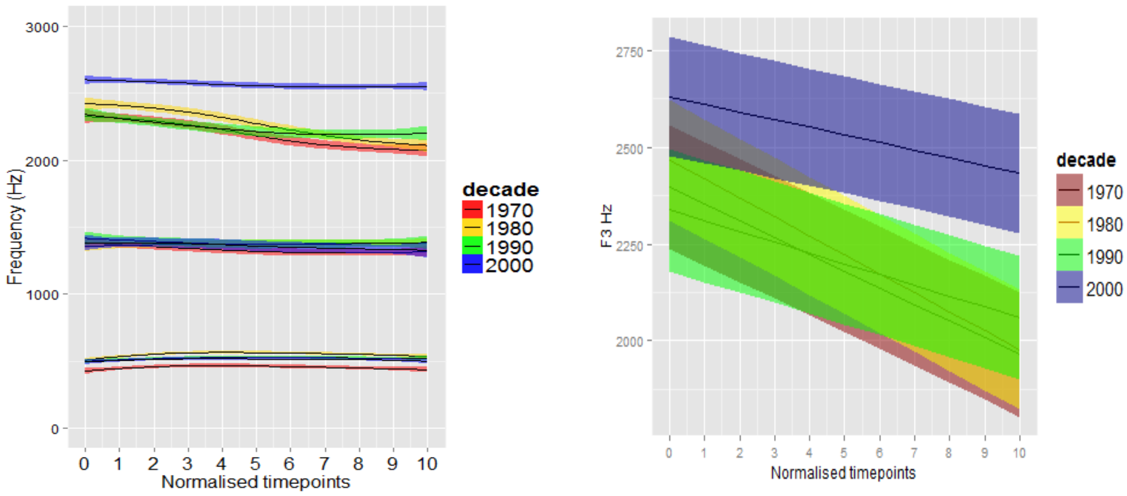
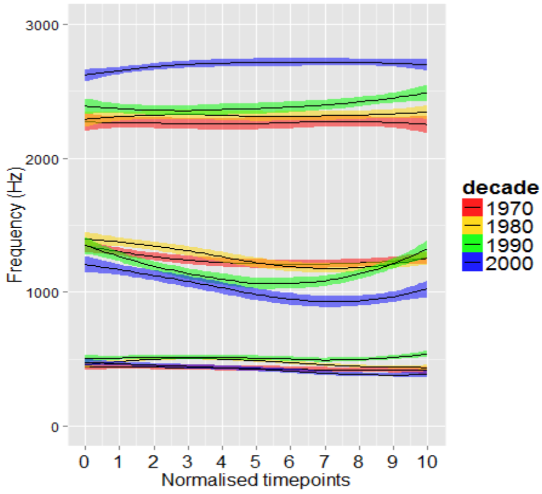


Figure 6: Dynamic view of change in coda /r/ in 8 Glaswegian men by decade of recording. Left: Time-normalized average formant tracks for the first three formants across the rhyme in word-final /r/ (0: vowel start; 10 rhotic end). Right: Predictions from Linear Mixed Effects (LME) modelling for the F3 tracks showing interaction of decade of recording/birth with normalised timepoint. The brake on the dropping of F3 in speakers born in the 1910s and recorded in the 1990s, and especially in the 1920s and recorded in the 2000s, confirms auditory weakening of rhoticity observed by previous auditory studies (e.g. Stuart-Smith et al to appear).



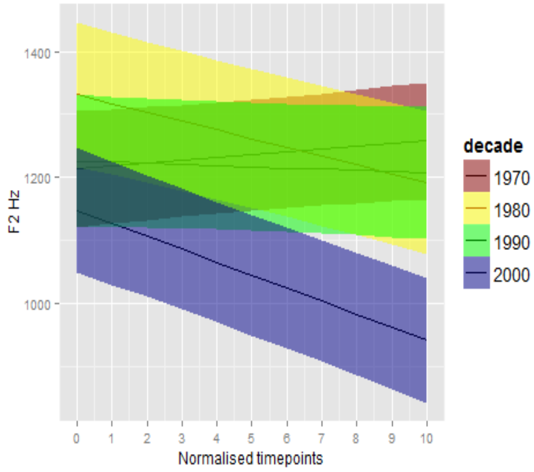
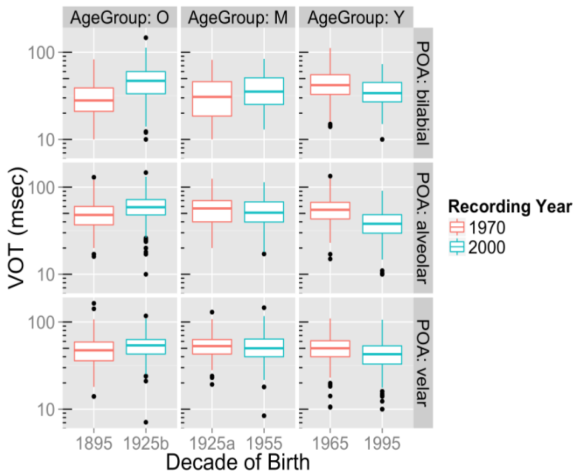


Figure 7: Dynamic view of change in coda /l/ in 8 Glaswegian men by decade of recording. Left: Time-normalized average formant tracks for the first three formants across the rhyme in word-final /l/ (0: vowel start; 10 lateral end). Right: Predictions from LME for the F2 tracks showing interaction of decade of recording/birth with normalised timepoint. F2 drops sharply for those speakers born in the 1920s and recorded in the 2000s, reflecting darkening of coda /l/ which appears to have been the precursor for new L-vocalization observed first by auditory studies in the 1980s (Stuart-Smith et al to appear).



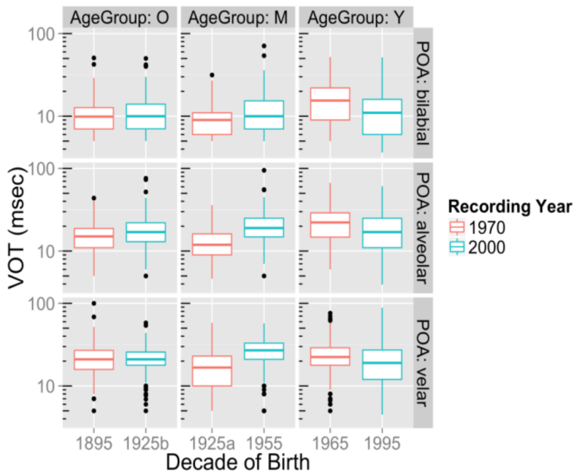


Figure 8: Change in a 'stable' variable: aspiration in initial stops in 23 female speakers (O = Old; M = Middle-aged; Y = Young). Boxplots of log/VOT by decade of birth and place of articulation for (left) voiced (*n* = 4088); (right) voiceless stops (*n* = 3247). VOT is lengthening over time for voiced and voiceless stops, except for adolescents born in the 1990s who revert to short lag stops (Stuart-Smith et al to appear).

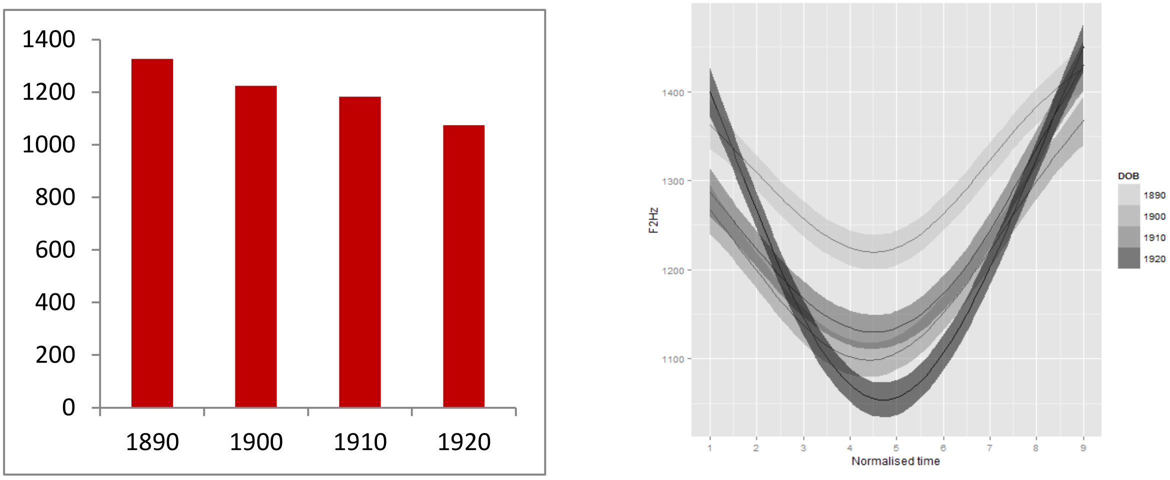


Figure 9: Static and dynamic views of change in a 'stable' variable: word-initial /l/ for 12 older speakers by decade of birth. Left: Estimates of F2 at the midpoint of the steady state; Right: formant tracks for /l/ in the phrase 'we lived' in female speakers. The value of F2 lowers significantly as the birth date becomes more recent, reflecting darkening of /l/ over time (Macdonald and Stuart-Smith, in preparation).

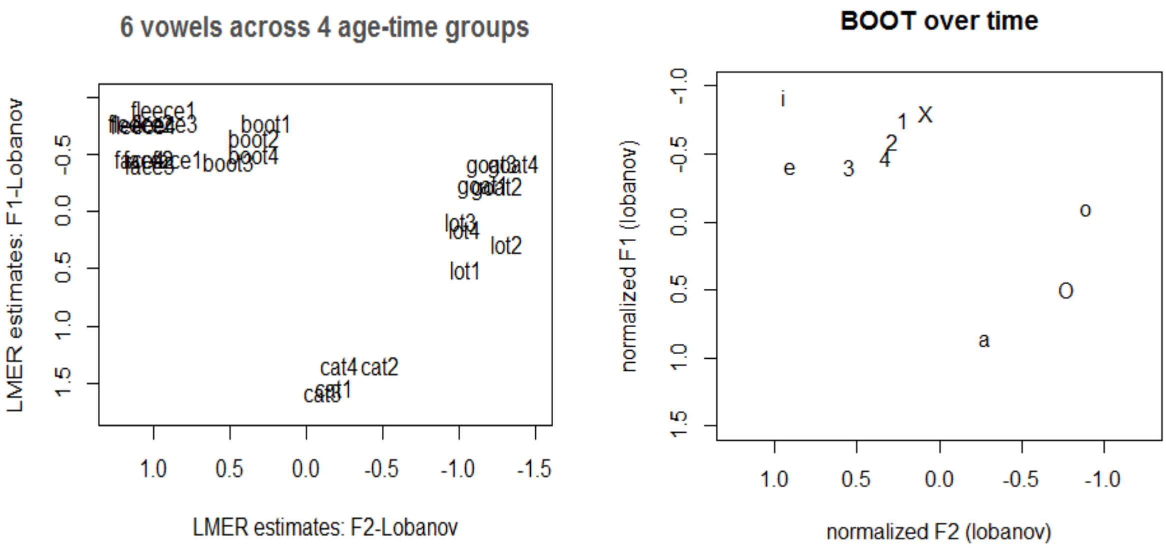


Figure 10: Change in 'stable' variables: acoustic vowel quality over time. Left: variation and change in 31 speakers (1 = born in 1890s, recorded in 1970s; 2 = born in 1920s, recorded in 2000s; 3 = born in 1960s, recorded in 1970s; 4 = born in 1990s, recorded in 2000s). LOT /ɔ/ and GOAT /o/ are raising, BOOT /ʉ/ is lowering. Right: variation and change in BOOT /ʉ/ in 19 male speakers (as before with X = born in 1890s, recorded in 1916/17). José et al in preparation; Stuart-Smith et al to appear.

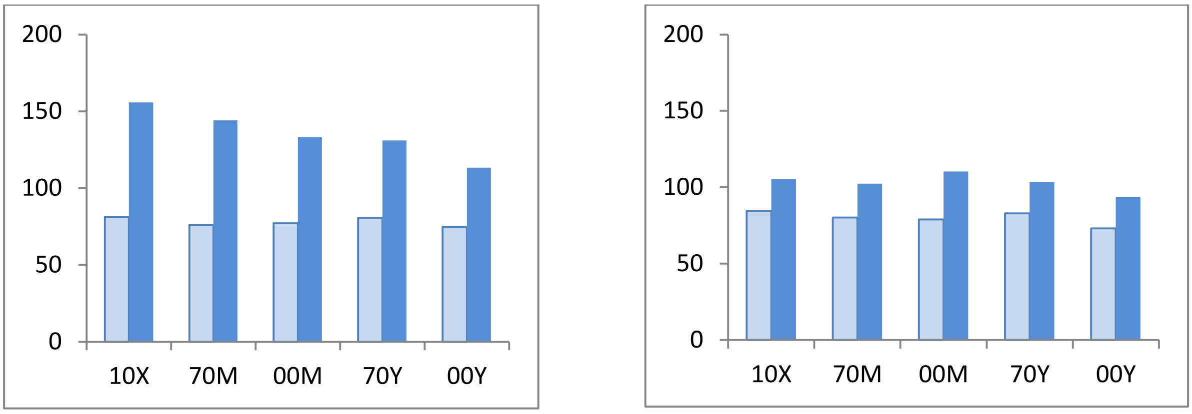


Figure 11: Change in a ‘stable’ variable: the Scottish Vowel Length Rule in 17 male speakers. LME estimates in milliseconds for /i ʉ/ in SVLR contexts in five speaker groups, in nuclear syllables (left) and non-nuclear syllables (right). Light bars show SVLR short contexts (e.g. *beat, bead*), dark bars long contexts (e.g. *bee, bees*; n = 1130). 10X = *BL* sample born in 1890s, recorded in 1910s; 70M = born in the 1920s, recorded in the 1970s; 00M = born in the 1950s, recorded in the 2000s; 70Y = born in the 1960s, recorded in the 1970s; 00Y = born in the 1990s, recorded in the 2000s. Long vowels become significantly shorter over time in the stronger prosodic context (Rathcke and Stuart-Smith to appear).

4. Gradience and change

Our evidence suggests that sound change is gradient, irrespective of whether measures are durational (VOT; SVLR) or spectral (liquids; vowel quality). While this may partly relate to the continuous nature of the measures, we also note:

- Coda /r/ shows a clear difference in F3 for speakers born in/after the 1910s, possibly reflecting an articulatory weakening and fronting of the tongue tip gesture
- Coda /l/ shows a sharp drop in F2, reflecting darker /l/ in speakers born in the 1920s, which may also reflect an earlier and more pronounced dorsal gesture.
- The darkening of coda /l/ in elderly speakers born in 1920s is a precursor for the apparently abrupt appearance of ‘new’ L-vocalization observed first in the 1980s (Figure 7).
- VOT shows gradual lengthening in speakers born in the 1890s to those born in the 1960s, but the most recent group revert to short-lag stops, more typical of vernacular Scots.

5. Linguistic factors and change

- The two variables considered using durational measures both confirm that prosodic factors play a role in sound change. SVLR weakening occurs in phrase-final position and nuclear position. VOT lengthening is greater in phrase-initial position.
- Real-time comparison of the dynamic patterns of F2 in coda liquids shows that the earlier additional distinction between clearer /r/ and darker /l/ becomes enhanced through further darkening of /l/ in speakers born in the 1920s and recorded in the 2000s (Figure 12).

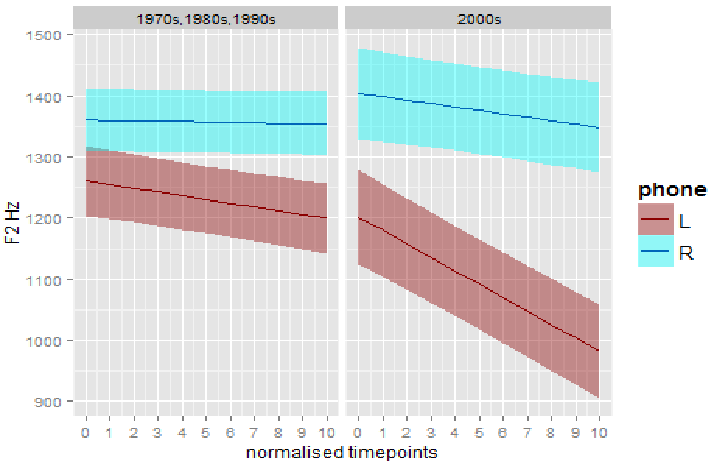


Figure 12: Predictions from LME for the F2 tracks for word-final liquids, showing interaction of normalised timepoint with phone and decade of recording (shown as two levels for ease of viewing). The drop in F2 in /l/ in speakers born in the 1920s and recorded in the 2000s enhances the contrast between /r/ and /l/ in word-final position (Stuart-Smith et al to appear).

- Lexical frequency is not a significant factor for any of the changes for which it was included (SVLR; initial /l/; coda liquids; VOT), though there is a tendency in the expected direction for VOT. This is intriguing given predictions for the role of lexical frequency on e.g. reduction processes, based on synchronic datasets; diachronic evidence seems to show a different picture.

6. *Social factors and change*

- The hypothesized impact of urban regeneration on phonetic variation is only discernible for one variable (SVLR weakening); the First World War appears to be generally more important (e.g. VOT lengthening; initial and coda /l/ darkening; /r/ weakening; /ʉ/ lowering).
- Despite extensive social and geographical mobility during the 20th century, all changes observed look dialect internal, promoted by linguistic and/or social factors operating within Glasgow and Scotland. This is interesting taken alongside recent evidence that the 'innovations' TH-fronting and L-vocalization also show signs of strong linguistic embedding, and only weak evidence for dialect contact (Stuart-Smith et al to appear).
- Little evidence has emerged for individual variation in patterns of change identified at the level of the group; but the groups are often fairly small, and further work is needed.

7. *Segment type and change*

There is no difference in the observable patterns of change for vowels vs consonants, either phonetically, or in terms of the influence of factors constraining/promoting change. It is not possible to say from these real-time variables that e.g. consonants are changing more than vowels, or that vowels are more local than consonants.

8. *Personal evaluation*

My own view is that the project has been very successful. Thanks to the hard work of the RAs and all who contributed, including our superb collaborators, we have carried out an enormous amount of work in a very short amount of time. Highlights have been: solving the methodological issues, including having the whole (phonetics) team learn to use R and implement mixed effects modelling; discovering the WWI recordings; developing stimulating new collaborations; and seeing two RAs take up permanent lectureships. The success of our work is substantiated by the enthusiastic reception of our work by the international academic community (see below), as well as my own nomination to lead a Niels Bohr project on sound change at Copenhagen (2016-2020).

Identifying less successful elements is hard: we were disappointed at the difficulty in finding middle-aged and younger speakers for the 1980s, which was not helped by the withdrawal of access to BBC Scotland after staff cuts. We have identified only a few instances of individual variation, but this may change as we continue analyses with larger numbers of speakers. Again, the finding that lexical frequency is not significant was intriguing, and is the subject of further collaborative work with Sóskuthy.

Publications and dissemination:[1]

(a) Publications: To date we have 6 peer-reviewed publications in print or in press (*LSA2014; ICLaVE7; ICPhS2015; Laboratory Phonology, Sense of Place; Listening to the Past*); one has been revised and is pending acceptance (*Language and Speech);* and 4 are in preparation with expected submission autumn 2015 (to *Language Variation and Change, Journal of Phonetics, JASA, Language in Society*).

(b) Presentations: Our research has been presented at 17 international meetings: BAAP 2012; 2nd Sound Change Workshop; SS19; NWAV41; ExAPP2013; ICLaVE7; UKLVC9; NWAV42; LSA2014; BAAP2014; PAC2014; 3rd Sound Change Workshop(plenary); LabPhon14; Methods15(invited workshop); SST2014 (plenary); PAC2015 (plenary); ICLaVE8.

(c) The Glasgow real-time corpus is currently a long-term resource for local, national and international postgraduate research beyond our project, as well as local undergraduate teaching.

(d) The project launch was widely covered in the broadcast and print media. We have delivered successful hands-on public workshops at the Glasgow Science Festival (2014, 2015).

(e) The project website, *Sounds of the City,* also provides an open-access resource, with stories, quizzes, activities and background information, to promote interest and awareness in language change, and to share our findings in an accessible format for the general public, which is also directly useful for teachers and pupils working with the Scottish Curriculum for Excellence.

Future research plans:

1. Continuing project research

- Kelvin-Smith-funded PhD between Mathematics/Statistics and English Language investigating the temporal dependence of variation in spoken corpora (Alexander: 2014-18).
- With Sonderegger, real-time change in the full stop voicing contrast with three acoustic cues, submitted for NWAV44 and intended written submission to *Language* this autumn
- With Jen Hay and colleagues (ChristChurch, NZ) on the SVLR in early Scottish settlers to New Zealand (UKLVC10).

2. Next steps

Our current grant (Carnegie: *How stable is the standard?* 2015) begins a real-time Scottish Standard English corpus, whose completion is the remit of a major pending ESRC bid. The linked corpora will provide a unique view of a community's spoken language across its entire recorded history.

Key words:

Language change; phonetics; sociolinguistics; analysis of digital corpora; Scottish English

[1] See http://soundsofthecity.arts.gla.ac.uk/Outputs.html