

Methodological Issues in a Real-Time Study of Glaswegian Vowels: Automation and Comparability

Brian José & Jane Stuart-Smith*

LSA Annual Meeting, Minneapolis, January 2-5, 2014

1. Introduction. Our poster, still viewable at <http://soundsofthecity.arts.gla.ac.uk/Outputs.html> at the time we submitted this text to the LSA, presents just one aspect of the research being conducted for the *Sounds of the City* project (or, more formally, the project on *Fine Phonetic Variation and Sound Change: A Real-Time Study of Glaswegian*) at the University of Glasgow. Our particular focus, here, is on a subset of the Glaswegian vowel space, specifically those vowels that Scobbie, Turk, & Hewlett (1999: 1617) refer to as the six bimoraic monophthongs of Scottish English: FLEECE /i/, FACE /e/, BOOT /ʊ/, CAT /a/, LOT /ɔ/, and GOAT /o/. We analyze these in a real-time cross-sectional (or trend) sample of 31 speakers stratified for age, sex, and time of recording as summarized in Table 1, with the following goals: (i) to highlight the type of automation procedures being used for the Glasgow real-time project and, thereby, to contribute to current discussions about the range of tools available for such purposes (cf. Labov, Rosenfelder, & Fruehwald 2013) and (ii) to participate in ongoing evaluations of data comparability in sociolinguistics (e.g., Bailey & Tillery 2004; Gregersen, Beck Nielsen, & Thøgersen 2009; Tillery & Bailey 2003, among others). We also seek (iii) to add to the ever-growing body of real-time sociolinguistic literature (e.g., Gregersen 2009; José 2010; Rickford & Price 2013; Sankoff & Blondeau 2007; Van Hofwegen & Wolfram 2010; among others) as well as (iv) to add to the expanding body of research on language variation in Glasgow, Scotland (e.g., Macafee 1983, 1988; Macaulay 1977, 2001; Stuart-Smith, Pryce, Timmins, & Gunter 2013; among others).

2. Methods. The 31 recordings in this sub-sample consist of sociolinguistic (70-O, 70-Y) and oral history (70-O, 00-O) interviews and free, unmonitored conversations between self-selected pairs of friends (00-Y). Additional types of recordings exist in the corpus more generally.

2.1. AUTOMATION. The audio recordings and time-aligned orthographic transcriptions of them are stored in a searchable “LaBB-CAT” database (Fromont & Hay 2012) which allows the user to apply a hidden Markov model (HTK) algorithm to align the speech signal with phonemic transcriptions, subsequently enabling the automatic extraction and measurement of vowel tokens, among other actions. We performed searches for the lexically-stressed instances of each of the six vowels in the database from our designated 31 speakers. Using LaBB-CAT’s built-in integration with Praat (Boersma & Weenink 2012, etc), LPC-generated formant measurements were obtained for the first three formants at each of three separate measurement points: 25%,

* We gratefully acknowledge financial support from the Leverhulme Trust in the form of a research grant, number RPG-142, awarded to the second author and from the School of Critical Studies in the form of a travel grant in partial support of the first author’s attendance at the 2014 LSA meeting. We would also like to express our appreciation, in approximate chronological order, to Robert Fromont for expert technical support with LaBB-CAT; to Astrid Ag, Frans Gregersen, & Randi Skovbjerg Sørensen for DCA training and guidance; to Hannah Birtwistle-Craine, James Butler, Peter Campbell, Shouchou (‘Cheery’) Chien, Fraser Dallachy, Kerry McPhillips, Lewis Thomson, & Louise Wilson for DCA coding; to Ellen Gallagher for invaluable assistance with data preparation; to Ludger Evers & Tereza Neocleous for valuable statistical advice; and to Rachel Macdonald for helpful discussions and for proofreading a draft of this submission. Naturally, none of these individuals are to be held accountable for any errors or shortcomings in the analysis and/or interpretation of the data; full responsibility for these rests solely with the authors. Authors: Brian José, University of Glasgow (bjose.ling.phd@gmail.com) & Jane Stuart-Smith, University of Glasgow (jane.stuart-smith@glasgow.ac.uk).

		Age and Sex of Speakers			
		“Old” = 67-90 yrs		“Young” = 10-17 yrs	
		M	F	M	F
Decade of Recording	1970s	4	3	4	4
	2000s	4	4	4	4

Table 1: Thirty-one speaker sub-sample

50%, and 75% of the distance between the HTK-calculated beginning and end points of the vowel. The raw formant values as well as the means of the three 25-50-75% measures were then normalized following the Lobanov (1971) method implemented using the online NORM suite (Thomas & Kendall 2007). Approximately 14,000 combined tokens were ultimately retained for analysis. Statistical analyses look for (i) differences among the 6 vowels at any one point in “time” (70-O, 00-O, 70-Y, 00-Y) and for (ii) movement across those four age-time groups for each of the six vowels individually. In each case ((i) and (ii)), the normalized 3-point mean F1 and F2 measures are modeled separately using linear mixed effects models with fixed factors for either the vowel (i) or the age-time group (ii), the speaker’s sex, and the preceding & the following places of articulation and with random factors for the individual speaker and word.

2.2. COMPARABILITY CODING. These 31 transcripts (plus an additional 19 not analyzed here (=50)) were tagged for a series of discourse context categories adapted from the Discourse Context Analysis (DCA) coding protocols developed at the University of Copenhagen’s LANCHART Centre (cf. Gregersen et al 2009). For this pilot work adapting and adopting their system for our research, we retain 4 of their original 6 categories, as summarized in Table 2. It’s noteworthy that the 4 of their 6 categories that we have retained survive in modified form. Thus, their 5 types of *speech events* have become 8 types of *participant structures* for us; similarly, we’ve replaced one of their *speech genres* (jokes) with a different one of our own (arguments), and one of our *activity types* (monologue) is a type of *interaction structure* in their system. We disregard the DCA-P category for now because there’s a complementary distribution effect in our data: the 70-O, 00-O, and 70-Y groups are only and always interviews and the 00-Y group is only and always peer-pair conversations. The DCA-C category serves, so far, only as a diagnostic measure for pre-emptively removing all quoted and/or *others’* speech from the analysis. In the future, we may investigate whether our speakers’ renditions of others’ speech is

<i>LANCHART’s 6 DCA categories</i>		<i>Our 4 DCA-inspired categories</i>
• S = type of speech event (5 possible types)	=>	• P = participant structure (8 possible types)
• A = activity type (6 possible types)	=>	• A = activity type (6 possible types (≠))
• I = type of interaction structure (9-4=5 possible types)		
• M = type of macro speech act (5 possible types)		
• G = type of speech genre (8 possible types)	=>	• G = type of speech genre (8 possible types (≠))
• U = ‘enunciations’ (Danish <i>udsigelse</i>) (1 possible type (<i>yes or no</i>))	=>	• C = citations (1 possible type (<i>yes or no</i>))

Table 2: Adaptation of Discourse Context Analysis coding categories for the Glasgow project

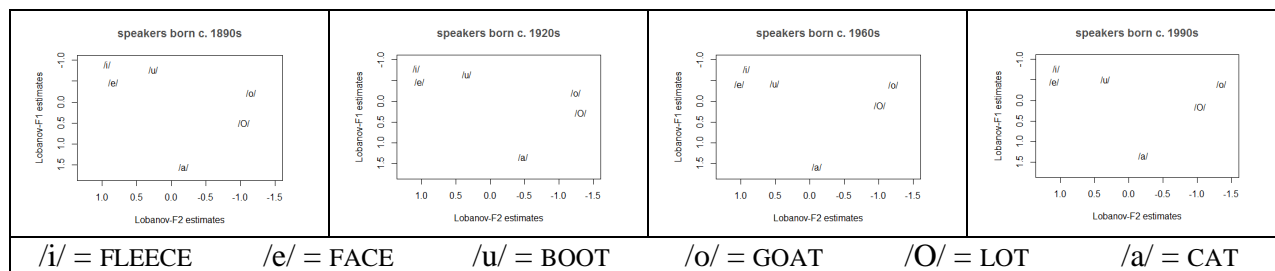


Figure 1: Normalized F1 x F2 positions of the 6 bimoraic monophthongs at 4 points in time

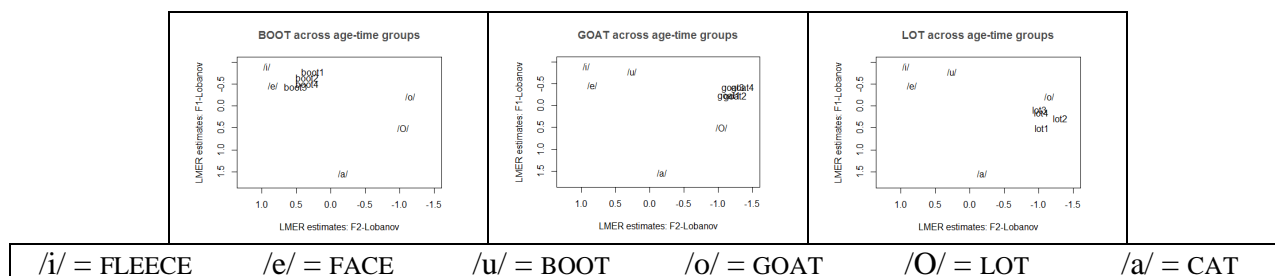


Figure 2: Movement of BOOT, GOAT, LOT vowels over time

measurably different from their own, but for now we simply disregard the approximately 5-10% of the data coded as *C-Quot*.

3. Results. Here, we provide a necessarily abbreviated summary of our main findings.

3.1. REAL-TIME DEVELOPMENTS IN THE GLASWEGIAN VOWEL SPACE. Figure 1 plots the normalized F1 x F2 vowel spaces of the 6 bimoraic monophthongs in each of our 4 age-time groups. Changes over time can be ascertained by comparing the distribution of the vowels within that space at one point in time to a subsequent point or points in time: Fig 1A shows the 70-O group (born c. 1890s), Fig 1B shows the 00-O group (born c. 1920s), Fig 1C shows the 70-Y group (born c. 1960s), and Fig 1D shows the 00-Y group (born c. 1990s). Alternatively, Figure 2 plots the changes of primary interest: lowering of the BOOT vowel (Fig 2A), raising of the GOAT vowel (Fig 2B), and raising (as well as a back-then-forward oscillation) of the LOT vowel, each of which is identified as significant or marginally significant in the LMER analyses. The other three vowels (FLEECE, FACE, and CAT) also exhibit some significant development(s) over time, but only in interactions with other factor groups; thus, while we can say that the influence of those other factor groups ---e.g., the preceding and/or following POAs--- has changed over time, we can't identify any overall, general changes in the quality of those vowels. Therefore, we limit our observations, here, to just the 3-4 robust developments illustrated in Figure 2; nor will we mention, here, any of the significant interactions involving age-time for the vowels in Figure 2.

In each panel of Figure 2, *vowel1* represents that vowel's position in the 70-O age-time group, *vowel2* represents its position in the 00-O age-time group, and so forth; the other 5 vowels are held stationary in their 70-O starting positions. The BOOT vowel (comprising both the FOOT and GOOSE lexical sets) lowers successively from 70-O to 00-O to 70-Y and then rebounds slightly from 70-Y to 00-Y. The GOAT vowel shifts upwards between 00-O and 70-Y. The LOT vowel mirrors the developments of the BOOT vowel: it rises successively from 70-O to 00-O to 70-Y and then rebounds slightly from 70-Y to 00-Y; on the F2 dimension, it is also farther back in the 00-Y group than in the other three age-time periods (which, likely, isn't a *change* but merely an idiosyncrasy in the 00-O group).

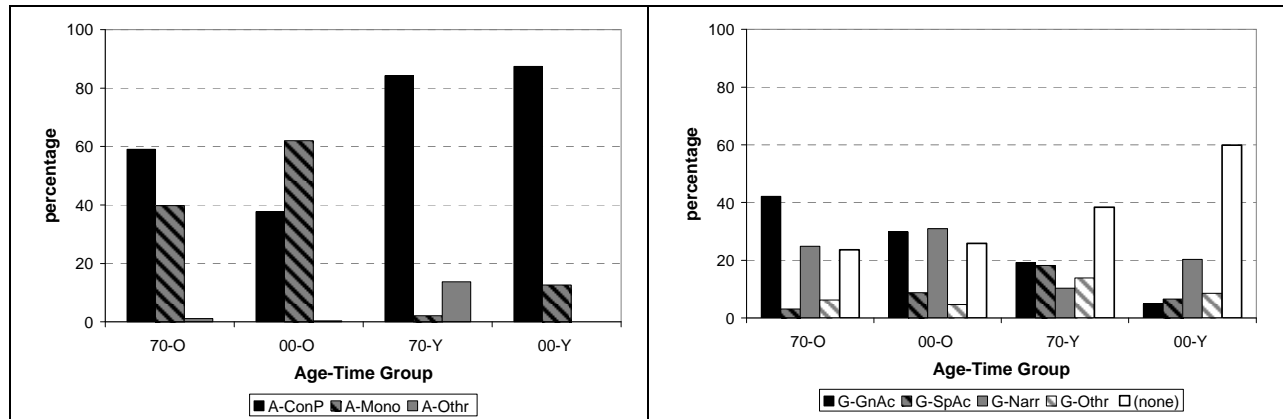


Figure 3: Distribution of DCA-A (L) and DCA-G (R) re-codes in the current data set

3.2. DCA COMPARABILITY. Figure 3 illustrates how our data is distributed within our modified DCA-A and DCA-G categories. Within DCA-A (Fig 3A), the vast majority of the data is divided between conversation with another ratified participant in the session (A-ConP) and monologues (A-Mono); however, because these aren't clearly separated types in our system, many of the monologues may simply be long(er) uninterrupted passages that occur within conversations. Nonetheless, for our initial statistical analysis, we recode the data into a two-way opposition between A-ConP on the one hand and everything else (including A-Mono) on the other hand. Within DCA-G, we find that general accounts (G-GnAc), specific accounts (G-SpAc), and narratives (G-Narr) are each fairly well represented and that other types are also relatively robust when pooled together; it's also possible for speech not to be classifiable into any of our designated eight speech genres (thus, *none*). For our initial statistical analysis, we recode this category into a four-way opposition: G-GnAc vs G-SpAc vs G-Narr vs all others / 'NA'.

The statistical models reveal that speech that occurs during conversation with a participant (A-ConP) and speech that occurs elsewhere (again: including, but not limited to, within monologues) are effectively equivalent, at least with respect to their influence on the changes depicted in Figure 2: i.e., the formant estimates of those vowels are not significantly different according to DCA-A types. Likewise, the respective influence of the four recoded types of speech genres do not contribute significantly differently to the formant estimates of those vowels. There is some evidence that (i) the DCA-G category and (ii) an interaction between the age-time group & the DCA-A category have a weak effect on the normalized F2 values of the LOT vowel, but neither of them meets our critical alpha level (adjusted, due to multiple comparisons, from $p \leq 0.05$ to $p \leq 0.001$). In conclusion, then, the preliminary evidence suggests that our two activity types are generally comparable with each other, as are our four speech genres, and, thus, that DCAs can be exploited beneficially for research on language variation and/or change.

(Selected) References

- Fromont, Robert & Jen Hay. 2012. LaBB-CAT: An annotation store. *Proceedings of the Australasian Language Technology Workshop*. 113-117.
- Gregersen, Frans, Søren Beck Nielsen & Jacob Thøgersen. 2009. Stepping into the same river twice: On the discourse context analysis in the LANCHART project. *Acta Linguistica Hafniensia* 41. 30-63.
- Labov, William, Ingrid Rosenfelder & Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89. 30-65.
- Scobbie, James M., Alice E. Turk & Nigel Hewlett. 1999. Morphemes, phonetics, and lexical items: The case of the Scottish Vowel Length Rule. *Proceedings of the XIVth International Congress of Phonetic Sciences*. 1617-20.