VARIATION IN VOICE ONSET TIME ALONG THE SCOTTISH-ENGLISH BORDER

G.J. Docherty[a], *D. Watt*[b], *C. Llamas*[b], *D. Hall*[b] *& J. Nycz*[b]

[a]Newcastle University & [b]University of York, UK
g.j.docherty@ncl.ac.uk; dw539@york.ac.uk; cl558@york.ac.uk; jn527@york.ac.uk; d514@york.ac.uk

ABSTRACT

This paper presents findings on VOT variability in voiced and voiceless plosives produced by 159 speakers from four locations straddling the English-Scottish border. The analysis reveals key effects of both social and phonetic factors, highlighting the importance of both in accounting for variability.

Keywords: VOT, sociophonetic variability

1. INTRODUCTION

Increasing convergence across the phonetics/ sociolinguistics divide has not only led to the development of new integrated theoretical positions [7, 8, 9, 13], but has also opened up a wider range of explanatory accounts for observed patterns of variation, with the emphasis on ensuring that variation which is attributed to social factors might not be a secondary consequence of some other phonetic factor (and vice versa). Docherty [6] illustrated a case in point; across a range of moderate speech rates, the incidence of glottal realisations of /t/ among speakers of Tyneside English was insensitive to cross-speaker differences in speech rate, but the incidence of glottal realisations *did* correlate with more extreme rates of speech, with a high incidence being found at fast rates (> 6.5 sylls/sec) and very low incidence at low rates. However, rate alone was not the determining factor in respect of glottaling of /t/ since not all those speakers who produced relatively low speech rates (as it happens, all elderly) produced low rates of glottaling, and low rates of glottaling were found in some speakers whose speech rate was entirely in the mid-range. This suggests that the explanation needs a dimension relating to the phonetic consequences of differences in speech rate while at the same time acknowledging that there is a learned component to this type of realisational variation in that variety of English.

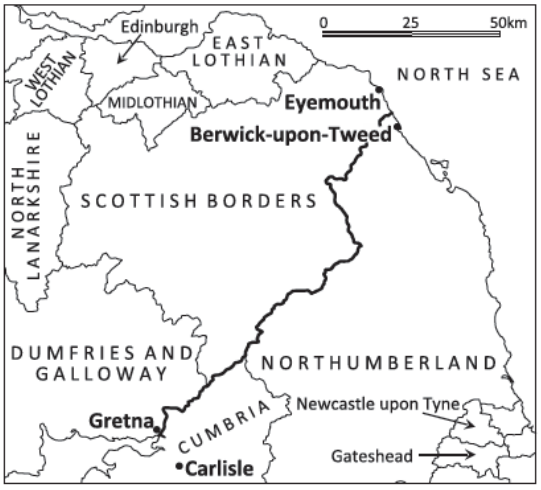
In this paper we consider another case where similar considerations must apply: variation in voice onset time (VOT [4, 5, 10]) for voiced and voiceless stops across young and old speakers recorded in four locations straddling the English-Scottish border. This project is part of a larger study (*Accent and Identity on the Scottish-English Border*-AISEB) tracking phonological variation and change across a political border which has been described as aligning with "the most numerous bundle of dialect isoglosses in the English-speaking world" [1] p 895. Previous comparative accounts of VOT in English vs. Scottish varieties suggest that for many speakers north of the border, /p t k/ are unaspirated [3, 16]. Recent studies of phonological variation in the north-east of England have not focused on VOT primarily because it was not perceived to be realised differently to what was reported for English varieties more generally, although there are reports (e.g. [16]) of some northern varieties being marked by shorter VOTs in voiceless stops, and this is occasionally a feature that surfaces in imitations of northern English accents. Within the AISEB project, analysis of a range of other variables has brought to light a number of parameters where there are significant age-related differences within a particular location, and has also pointed to different levels of cross-border permeability attributable at least in part to prevailing language ideologies. Given this context, the present study investigates the patterning of VOT across the four communities concerned with the aim of examining the complex interaction between phonetic and social factors in the realisation of a phonological variable.

2. METHOD

The AISEB project recorded speakers from four localities along the English-Scottish border (Berwick-upon-Tweed [English] and Eyemouth [Scottish] to the east, and Carlisle [English] and Gretna [Scottish] to the west – see Figure 1). Forty

informants per location were recorded, stratified by age (Old: 57+; Young: 16-25), sex, and social class (working class vs. middle class). Two speech styles were elicited: read speech arising from a word list and from a narrative text, and unscripted conversational speech arising from responses to a structured questionnaire administered by a fieldworker. Recordings were made in quiet surroundings (most often in speakers' homes) using a Zoom H4n recorder with built-in variable XY stereo microphones recording direct to .wav format (44.1kHz / 16-bit), and were subsequently analysed using *Praat*.

Figure 1: Location of the four sites investigated in the AISEB project (the political border is shown as a bold line) [12].



4,662 tokens of stressed syllable-initial plosives were extracted from the corpus for analysis in respect of VOT (3,319/1,343 tokens of voiceless and voiced stops respectively). All tokens consisted of plosives produced in the word-list task in word-initial position (with the exception of the stressed /t/ of *potato*) with a following stressed syllable nucleus. VOT was measured as the interval between the onset of the burst of energy in a spectrogram associated with the release of the stop closure, and the first vertical striation indicating the onset of voicing (a negative VOT was registered if the voicing began in advance of the stop release). Measurements of the voiceless stops and the voiced stops were each subsequently analysed using multiple regression in order to test for the effects of the independent variables built in to the design of the corpus (Country [England/Scotland], Coast [East/West], Speaker Age/Sex/Class) on VOT. The present paper highlights in particular the effect of Speaker Age.

3. RESULTS

Figures 2-5 show the distributional density of VOT for voiced and voiceless stops for each of the four locations pooled across all speakers, but differentiated by Speaker Age. It can be seen that in all sites, voiceless stops are realised predominantly with relatively long VOTs in the range that is associated with audible aspiration. For voiced stops, the distribution of VOT is more variable (both within and between speakers), with VOTs occurring in both the pre-voiced (negative VOT) range and in the short-lag (positive VOT) range typically associated with the production of voiceless unaspirated tokens. A bi-modal distribution of this sort is not an unusual finding in respect of English voiced stops [5, 10].

Figure 2: Distributions of VOT for Berwick speakers as a function of Speaker Age.

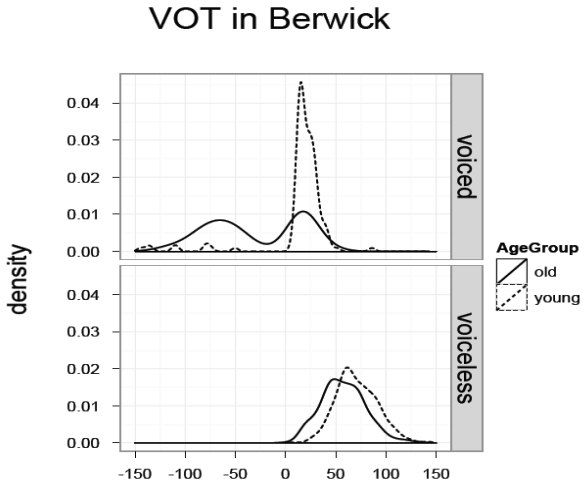


Figure 3: Distributions of VOT for Eyemouth speakers as a function of Speaker Age.

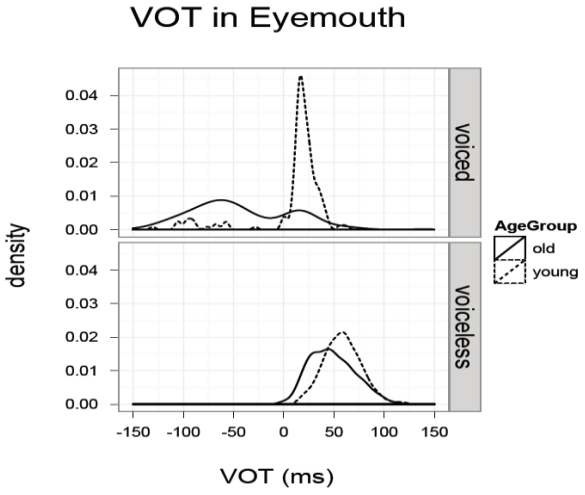


Figure 4: Distributions of VOT for Carlisle speakers as a function of Speaker Age.

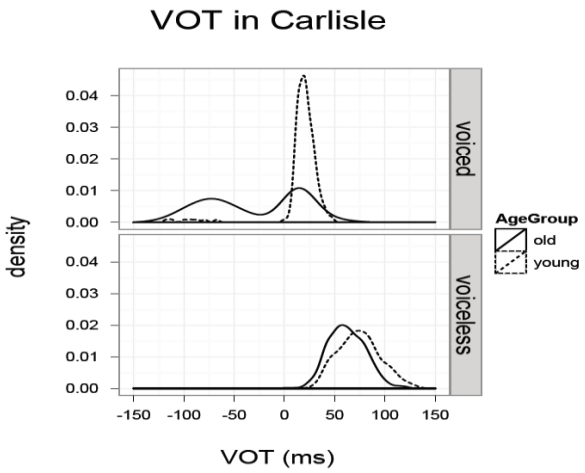
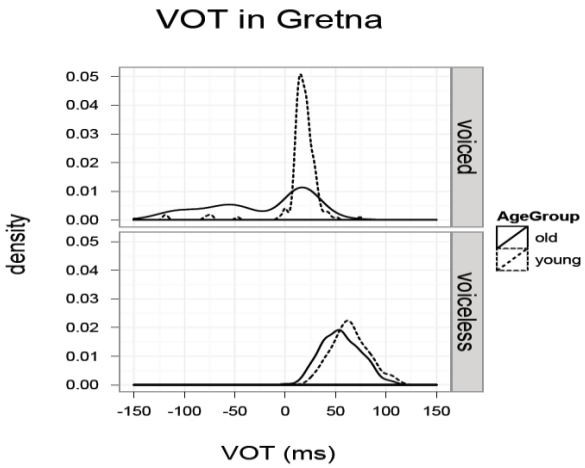


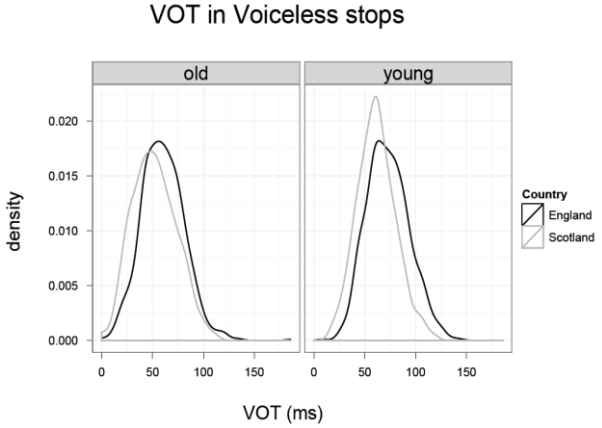
Figure 5: Distributions of VOT for Gretna speakers as a function of Speaker Age.



Visual inspection of the figures, confirmed by the results of the multiple regression analysis, reveals a striking effect of Speaker Age. Younger speakers are much less inclined to use prevoicing for their voiced stops, so that with other variables held constant, their VOTs in /b d g/ are significantly longer overall than those produced by old speakers ($t_{1,1337}$ = 20.02, p<0.001). For young speakers across all four localities, voiceless stop VOTs are significantly longer than those for older speakers, notwithstanding a large overlap in density distributions ($t_{1,3313}$ = 16.15, p< 0.001).

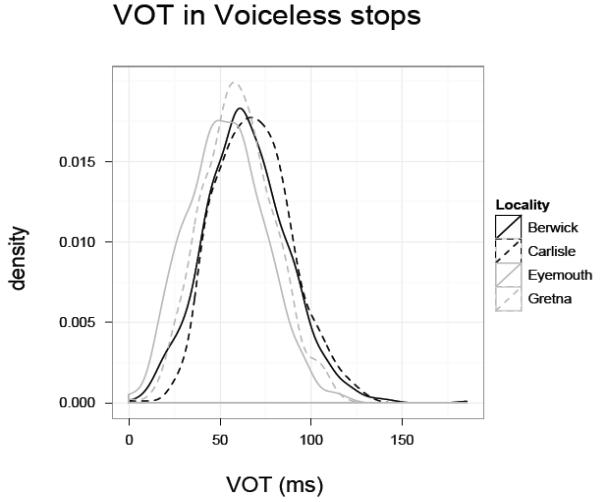
There are also effects of the Country in which the recordings were made. Figure 6 shows density distributions for /p t k/ VOT sub-categorised by Age and Country. It can be seen that for both young and old speakers there is a small but significant effect of Country ($t_{1,3313}$ = 14.19, p< 0.001), with Scottish locations associated with a 10ms shorter VOT than those locations south of the border, a finding which is consistent with the view that Scottish varieties of English may be characterised by shorter VOTs for /p t k/ (although as just noted, this is clearly an effect which interacts with speaker age).

Figure 6: Voiceless stop VOT by Speaker Age and Country.



There is also a small but significant effect of Coast, with the eastern localities having slightly lower voiceless VOTs than their western counterparts in the same country ($t_{1,3313}$ = 7.40, p< 0.001). Figure 7 shows the VOT distribution densities for /p t k/ for all speakers sub-categorised by recording site. While there is a good deal of overlap, it can be seen that distribution for the speakers from Eyemouth is positioned somewhat lower in the VOT range than that for speakers from Gretna, while Berwick occupies a slightly lower range than Carlisle.

Figure 7: Voiceless stop VOT by Country.



4. DISCUSSION

From a sociolinguistic perspective one interpretation of the consistent effect of Age on VOT for voiceless stops would be that it reflects a change in 'apparent time', with the younger generation moving to enhance the audible aspiration of /p t k/. However, these findings must also be evaluated in light of previous studies of the effects of age on the acoustic properties of speech [2, 11, 14, 15] which provide evidence that, across languages, younger speakers tend to have longer VOTs overall for voiceless stops than older speakers (although the [presumably] physiological basis for this has not been established). It is noteworthy that previous work has not focused on the prevalence of pre-voicing for /b, d, g/ in older vs younger speakers; the present results suggest that negative VOTs for /b, d, g/ are relatively infrequent in younger speakers, and one might speculate whether this is connected to the fact that those speakers have a tendency to use relatively long VOTs for their voiceless plosive realisations.

Notwithstanding the fact that some of the results reported here are consistent with patterns seen more generally across languages/age groups (something which renders the change in apparent time account somewhat less plausible), the differences as a function of Country/Coast point to a somewhat more complex picture. While the young Scottish speakers (from Eyemouth and Gretna) do indeed have VOTs for /p t k/ which are longer than older speakers from the same location (and have largely abandoned the use of prevoiced realisations), the distribution of VOT which is deployed for /p t k/ is significantly lower than that found for the English locations, and in a direction which is consistent with the claims that speakers from Scotland will be characterised by having shorter periods of aspiration than speakers from England. The finding that Eyemouth speakers have shorter VOT distributions than Gretna speakers is consistent with other results from the AISEB project [12] which suggest that speakers in Eyemouth have a greater sense of 'Scottishness' than Scottish speakers from Gretna at the western end of the border, and that this correlates with their use of particular phonological variables (e.g. their patterns of rhoticity, or use of the 'Scottish Vowel Length Rule').

In sum, as well as providing the first detailed analysis of the realisation of the voicing contrast by speakers located either side of this very significant linguistic border, these results provide further illustration of the value inherent in considering both phonetic and social factors when building accounts of phonetic variation, and the need for these to be integrated within models of speech production.

5. REFERENCES

[1] Aitken, A. 1992. Scots. In: McArthur, T. (ed.), *The Oxford Companion to the English Language*. Oxford: Oxford University Press, 893-899.

[2] Benjamin, B.J. 1982. Phonological performance in gerontological speech. *Journal of Psycholinguistic Research* 11, 159-167.

[3] Catford, J.C. 1988. *A Practical Introduction to Phonetics*. Oxford: Oxford University Press.

[4] Cho, T., Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, 207-229.

[5] Docherty, G.J. 1992. *The Timing of Voicing in British English Obstruents.* Berlin: Foris Publications.

[6] Docherty, G.J. 2007. Prosodic factors and sociophonetic variation: Speech rate and glottal variants in Tyneside English. *Proc, 15th ICPhS* Barcelona, 1517-1520.

[7] Docherty, G.J., Foulkes, P. 2006 The social life of phonetics and phonology. *Journal of Phonetics* 34, 409-438.

[8] Docherty, G.J., Mendoza-Denton, N. In press. Speaker-related variation – sociophonetic factors. In Cohn, A., Fougeron, C., Huffman, M. (eds.), *The Oxford Handbook of Laboratory Phonology*. Oxford: OUP.

[9] Hruschka, D.J., Christiansen, M.H., Blythe, R.A., Croft, W., Heggarty, P., Mufwene, S.S., Pierrehumbert, J.B., Poplack, S. 2009. Building social cognitive models of language change. *Trends in Cognitive Sciences* 13(11), 464-469.

[10] Lisker, L., Abramson, A.S. 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10, 1-28.

[11] Liss, J.M., Weismer, G., Rosenbek, J.C. 1990. Selected acoustic characteristics of speech production in very old males. *Journal of Gerontology* 45, 35-45.

[12] Llamas, C. 2010. Convergence and divergence across a national border. In Llamas, C., Watt, D. (eds.), *Language and Identities.* Edinburgh: Edinburgh University Press.

[13] Pierrehumbert, J., Clopper, C. 2010. What is LabPhon? And where is it going? *Laboratory Phonology* 10, 113-132.

[14] Ryalls, J., Zipprer, A., Baldauff, P. 1997. A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research* 40, 642-645.

[15] Torre, P., Barlow J.A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42(5), 324-333.

[16] Wells, J.C. 1982. *Accents of English* (3 vols.). Cambridge: Cambridge University Press.