Sex and gender differences in Glaswegian /s/

Jane Stuart-Smith*, Claire Timmins*, and Alan Wrench†
* Department of English Language, University of Glasgow, Scotland
†Department of Speech and Language Sciences, Queen Margaret University College, Scotland
E-mail: j.stuart-smith@englang.arts.gla.ac.uk, c.timmins@englang.arts.gla.ac.uk, awrench@qmuc.ac.uk

ABSTRACT

This paper presents the results of an acoustic analysis of /s/ as produced by a socially-stratified sample of 31 speakers of Glaswegian. Single spectral and whole spectrum measures together confirm a consistent effect of sex, whilst interactions of age and class in the same measures suggest that acoustic patterns indexical of gender are also present.

1. INTRODUCTION

It is well established that the acoustic characteristics of /s/ in English differ in male and female speakers. A number of studies have consistently demonstrated that female speakers show higher frequency energy than male speakers (see e.g. [7]). Reasons for observed differences in acoustics across females and males have tended to refer to biological or anatomical differences between males and females. Extending from the general expectation that females will have smaller vocal tracts than males [6], it is also expected that female speakers will have a shorter resonance cavity in front of the fricative constriction than men, and that this will cause the resulting turbulent air to show energy at higher frequencies. This rests on the accepted acoustic model of fricative production which links differences in the frequency and amplitude of fricative noise to different acoustic filters that the vocal tract forms when fricatives are articulated (e.g. [11]).

Flipsen et al [7] mention differences in cavity size as a possible factor to explain their own data, but also suggest that females in their sample may be using a more fronted articulation than male speakers, and even that the two are connected. Strand [12: 88] notes that 'general female-male vocal tract size differences ... exist mainly *behind* the area of the constriction ...', and goes on to describe an unpublished study by Naslund where males and females used different articulations to realise /s/. She uses this to argue for 'the development and salience of socially influenced fricative productions' for the samples in question. These observations move away from the assumption that males and females attempt the same articulation and that differences arise for reasons beyond their active control and from their biological sex, to the suggestion that males and females may use different strategies to articulate /s/, as a part of a more general construction of themselves as 'masculine' or 'feminine' within the culture and society in which they live. Such a distinction between biological 'sex', the biological determination as male and female, and 'gender', the socio-cultural construct of being male and female, is now widely made in the social sciences, and within sociolinguistics (e.g. [5]), and the term is now more common in phonetic research (e.g. [10], [12]).

The notion that the stream of speech, whilst carrying linguistic information, might also contain 'indexical' information about speakers' social grouping and identity, their individual status and their affective state, is not new to phonetics ([1]), and is an aspect which is of increasing interest in current research (e.g. [3], [4]). Abercrombie's discussion takes 'sex' as an idiosyncratic aspect of individuals, but notes at the same time the possibility of social indices, whereby 'in certain communities ... women have a different pronunciation from men' [1: 8], i.e. reflecting gender. It seems likely that we might find gender represented amongst the socio-indexical features of /s/ alongside those determined by sex.

A particular articulation of /s/ has been observed for working-class Glaswegian speech as distinct from middle-class speech. Macafee [9: 34] describes the articulation, which auditorily sounds rather 'retracted', as 'cacuminical' (apico-alveolar), i.e. produced with the tongue tip raised. Indeed, raising the tongue tip partially opens up the lower cavity which is formed during the articulation of [ʃ], '[h]ence any apical [s] may sound a trifle [ʃ]-like'; Catford [2: 157]. Preliminary auditory analysis of a recent corpus of Glaswegian speech identified a distinctive auditorily retracted pronunciation of /s/, but this was found as a property of male speech, and especially that of boys [13]. The subsequent detailed auditory analysis of the same data confirmed significant differences between male and female speakers, but again without the expected distribution according to class .

This paper considers the acoustics of /s/ in Glaswegian and attempts to answer the following questions: 1. What are the acoustic characteristics of /s/ in Glaswegian? 2. Do male and female speakers differ according to 'sex', and in the expected direction, i.e. females showing higher frequencies than males? 3. Is there evidence of 'gender' differences beyond those of 'sex'? 4. To what extent is class a factor?

 ISBN 1-876346-48-5

2. METHODOLOGY

The data discussed here were collected in 1997 from 32 speakers divided equally into the categories: 'older' (40-60 years) and 'younger' (13-14 years), working-class and middle-class, and male and female. Digital recordings were made using wide-frequency response clip-on microphones onto a Sony TC-D7 DAT recorder from speakers reading a wordlist and speaking in pairs. Only read materials were analysed here. All the wordlists were digitized into a Pentium 2 PC using *Xwaves* with a sampling rate of 16,000 Hz with 16 bits. This sampling rate is rather low for the analysis of a fricative like /s/ whose peaks may exceed 8,000 Hz, and the result can be aliasing, or the emergence of false peaks lower down the spectrum. Perhaps surprisingly aliasing was only a problem for /s/ in three words in one woman, though we excluded her data entirely from the analysis.

Quantitative analysis was preceded by a qualitative assessment of spectrograms for all 30 words containing /s/ for a sample of 10 speakers. This motivated a quantitative investigation of a subset of words which avoided consonantal clusters and contexts where /s/ followed a rounded consonant/vowel: *sieve, seven, somewhere, side, face, ice, icy, choice*.

Acoustic analysis was based on wide-band spectrograms (Hanning window, length 8 ms, 128 points) and DFT spectra (Hanning window, length 25 ms) obtained using *Xwaves*. A range of measures exist with which to capture aspects of the spectral energy of fricatives (e.g. [7]). We opted to use two different types of measure, partly to assess the extent to which results from the same data would coincide. We took single spectral measurements and whole spectrum measurements, all from a visually determined midpoint of the fricative. Single spectral measures characterize the spectral energy of a fricative as a single feature of the spectrum. We took three hand measurements, which we call here *minimum*, *cut-off* and *peak* frequency. The *minimum* frequency was determined as 'the frequency at which energy was first visible on the spectrogram', which usually occurred at the beginning of weak low-frequency bands which commonly occur well below the main band of energy associated with the place of articulation of the fricative. The *cut-off* frequency was defined as 'the frequency at which the main band of energy was first visible on the spectrogram', and reflects the increase in spectral energy on the approach to the peak. Both these measures relied on the settings of the spectrogram remaining constant throughout the analysis. The *peak* frequency was defined as 'the frequency at which the highest amplitude peak occurred in the spectrum'. Repeat hand measures were taken from a sample of 8 speakers by four other analysts. The original and repeat measurements showed strongly significant correlations for all three measures (p = .000).

We also took a set of four whole spectrum measurements using a moment analysis (also known as 'centroid' or 'centre of gravity', cf. Forrest et al [8]), in which '[t]he underlying assumption … is that the spectrum can be modelled as a single normal distribution which may reflect the dominant front cavity formant', Wrench [14: 460]. These measurements capture the characteristics of the overall distribution of spectral energy for a particular spectral slice. The four measures are called here *mean*, *spread*, *skewness,* and *kurtosis*. The *mean* gives the midpoint frequency at which the energy under the curve on either side of the point is equal. The *mean* is related to the *peak* but not in a simple way, since it is derived from all the energy under the curve, and not simply one single point. The other three measures describe the aspects of the spectral energy immediately flanking the mean. The *spread* refers to the 'bandwidth' of the energy on either side of the mean, the *skewness* to the asymmetry of the energy surrounding the mean, and the *kurtosis* to the peakiness of the distribution.

All the results were analysed statistically using 4-factor ANOVAs followed by post hoc tests to test for the extra-linguistic effects of SEX, CLASS, and AGE, and for the linguistic effect of POSITION IN WORD and subsequently WORD and their interactions. Only results which were significant with $p < .05$ are reported.

3. RESULTS

3.1. SINGLE SPECTRAL MEASURES

The results for *minimum*, *cutoff*, and *peak* frequency measurements are displayed in Figure 1. In this and the other figures, the numbers refer to social groups, with 1-4 representing females: 1 = middle-class women, 2 = middle-class girls, 3 = working-class women, 4 = working-class girls and 5-8 representing males: 5 = middle-class men, 6 = middle-class boys, 7 = working-class men, 8 = working-class boys

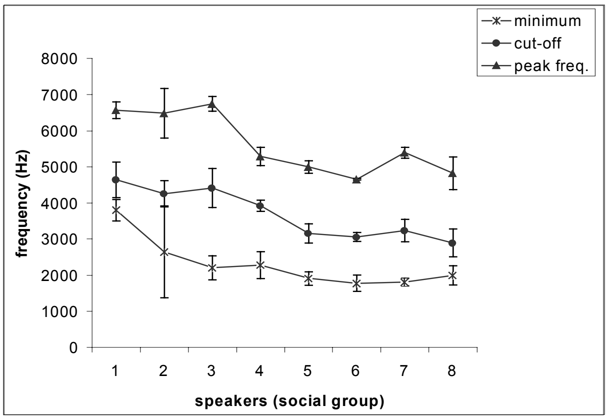

Figure 1: Means/sd in Hz for *minimum*, *cutoff*, and *peak* for the eight social groups (1-4: females; 5-8 males).

The only measures which showed significant effects of POSITION IN WORD were *cutoff* and *minimum* (initial vs final and initial vs intervocalic/final respectively) though only for *minimum* did this persist as an effect of WORD, translating into significant differences between *sieve* and *face/ice*. No interactions between WORD or POSITION IN WORD and the extralinguistic factors were found.

Figure 1 shows clearly that for the *minimum* frequency working-class females (3,4) pattern with male speakers (5-8) but that middle-class females (1,2) are higher, especially middle-class women (1). This is confirmed by significant factors of CLASS and SEX, AGE/CLASS, AGE/SEX, and CLASS/SEX interactions, and a post hoc difference of males and females within middle-class. The *cutoff* means for males and females occupy different frequency spaces, with a boundary just below 4,000 Hz. Within this working-class adolescents (4, 8) pattern differently, with each showing a lower mean than the adults. SEX is a significant factor, and this persists for all groups in post hoc tests. AGE too has an overall effect but does not emerge in post hoc testing, though this may reflect the observed differences in working-class adolescents. The clearest results in this group of measures are found in the *peak*. Figure 1 shows the grouping of the first three female groups, and the relative grouping of the four male groups, and, interestingly, how working-class girls (4) pattern with the male speakers. These impressions are confirmed through significant effects of SEX and AGE, and the interactions AGE/CLASS, CLASS/SEX, AGE/CLASS/SEX. Post hoc tests show sex differences for all groups bar the working-class (p = .059).

3.2. WHOLE SPECTRUM MEASURES

The results for *mean* and *spread* are shown in Figure 2, and for *skew* and *kurtosis* in Figure 3.

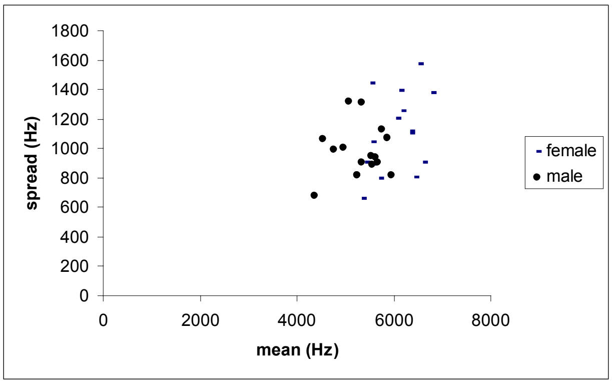

Figure 2: Scatter plot of *mean* against *spread* values (Hz) for all 31 speakers.

The *mean* was the only whole spectrum measure to show sensitivity to POSITION IN WORD (initial vs final), which emerged as WORD, and specifically to a difference between *seven* and *choice.* Again, as for the single spectral measures, no interactions with POSITION IN WORD or WORD were found with any extra-linguistic factors.

The patterning of working-class girls with male speakers found in the *peak* is repeated in the *mean* measure. (The two measures show a strong correlation with p = .000.) All effects are significant, as are the interactions of AGE/CLASS and AGE/CLASS/SEX. Post hoc tests reveal differences of sex for all groups bar the working-class (p = .06), and, interestingly, of age within male speakers, though the difference is not in a direction predicted by physiology, since working-class boys are lower than working-class men. Figure 2 shows how males (filled circles) are all below 6,000Hz, while females (dashes) tend to be above, but show outliers. These females are: the four working-class girls, one middle-class woman whose accent is most like English-English, and one middle-class girl, whose parents are from East Kilbride, a new town on the outskirts of Glasgow. This finding suggests that her parents may be working-class. The *spread* showed effects of SEX and AGE, interactions of AGE/CLASS and AGE/SEX and post hoc differences of older vs younger female speakers, and men vs women. These results boil down to a particularly wide spread in working-class women, especially with respect to middle-class girls and working-class men.

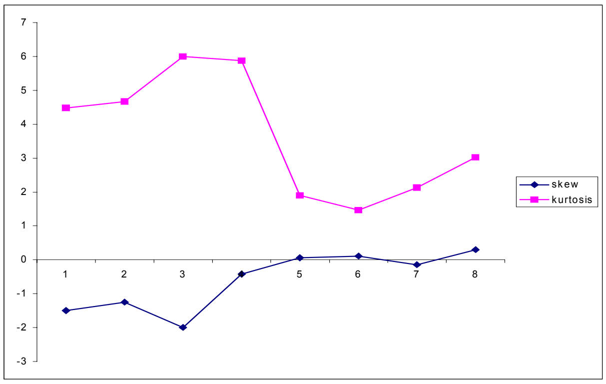

Figure 3: Means for *skew* and *kurtosis* for the eight social groups.

In Figure 3, means for *skew* show groupings of females (1-4) and males (5-8), with females showing negative values and males hovering about zero. Working-class females are distinctive: women show low values and girls much higher ones, verging on the same mean as working class men. SEX and AGE, and the interactions AGE/CLASS and AGE/CLASS/SEX are significant, with sex differences in all groups in post hoc tests. These results highlight working-class women as an independent group. *Kurtosis* again shows a clear division between females (1-4) and males (5-8), with females all above 4 and males all below 3. Having said that, within females the working-class females group together above the middle-class females, and within males, working-class boys are much higher than the other males. Despite these apparent trends, statistically this measure is the only measure to show SEX as the only factor, though in post hoc tests this does not hold within older speakers.

4. DISCUSSION

All of the seven measures show results determined by SEX, though of these only one shows SEX alone. The other six measures all show additional effects of AGE and CLASS, and all bar one, interactions of SEX with these additional effects. We interpret these findings as follows. The consistent finding of SEX suggests that aspects of the acoustic energy of /s/ in male and female speakers are determined by sex, i.e. by anatomical differences leading

ISBN 1-876346-48-5 © 2003 UAB

to resonance differences. We interpret the interactions of SEX with AGE and CLASS as indicative of additional gender differences in the data, assuming that these differences reflect indices of particular group identities. (We note that all AGE differences are manifested against expectations from physiology.) Thus, aspects of acoustic energy indexical of biological or anatomical sex provide a frame within which other aspects of acoustic energy, indexical of gender may be manipulated. The clearest group in terms of gendered identity of acoustic patterning emerges as the working-class girls, who group with male speakers in *peak* and *mean*. However, the relationships between measures and social groupings are not straightforward. These same speakers pattern with working class women in other measures (*minimum*, *kurtosis*). They are not the only group to emerge. Working-class women are distinct in their values for *skew*, middle-class women in theirs for *minimum*, and working class boys in their *mean*. Clearly being 'female' or 'male' is intricately bound up with being 'a working class girl' or a 'working class woman' at the very least. (This account simply operates with gross categories.)

Of the results that are comparable with those found elsewhere, we present similar findings. Like Flipsen et al [7] and the other studies discussed there, we find higher *peak* and *mean* frequencies in females, though with the reservation that this does not extend to working-class girls. Our results for *skew* are very similar to those of Flipsen et al [7], and are probably to be explained in the same way. They argue that negative skew values in females is an artifact of cutting off the spectrum at higher frequencies which has the effect of pushing the skew down the spectrum when the mean is high. That *skew* is tied to the mean explains the trend in working-class girls to show the highest skew, since they also show the lowest mean. We show different, sex-determined, results for *kurtosis*. Higher *kurtosis* values reflect a sharper peak whilst lower values a flattened distribution. We cannot explore the interaction of the measures here, but we note in passing the relation of *mean* and *kurtosis* in working-class females. Both demonstrate a sharp peak, a trend that distinguishes them from middle-class females, but we have to infer a high frequency sharp peak for women, but a lower frequency, but still sharp, peak for girls. We have also not explored the possible articulatory strategies behind these acoustic results.

5. CONCLUSIONS

We conclude by answering the research questions posed earlier. We have presented the acoustic characteristics of Glaswegian /s/ in some detail, but generalisations cannot be made without reference to speaker identity. Sex is a consistent factor in Glaswegian /s/, but so too are class and age, and with these, gender. Our analysis confirms the assumption of coexisting acoustic patterns indexical of sex and gender together.

ACKNOWLEDGEMENTS

Jane Stuart-Smith is very grateful to the Leverhulme Trust which supported the analyses presented here with a grant, and to the AHRB for a research leave grant which helped support further analysis and writing up the research. We are grateful to Alison Docherty, Valerie Sneddon, Karen Torrance, and Claire Young for cross-measuring the data.

REFERENCES

[1] Abercrombie, D. (1967), *Elements of General Phonetics*, Edinburgh: EUP

[2] Catford, J. (1977), *Fundamental Problems in Phonetics*, Edinburgh: EUP

[3] Docherty, G. (2003), 'Speaker, community, identity: Empirical and theoretical perspectives on sociophonetic variation', this volume

[4] Docherty, G. and Foulkes, F. (1999), 'Instrumental phonetics and phonological variation'. In P. Foulkes and G. Docherty, Eds, *Urban Voices*: *Accent Studies in the British Isles*, London: Arnold

[5] Eckert, P. (2000), *Linguistic Variation as Social Practice*, Oxford: Blackwell

[6] Fant, G. (1973*), Speech sounds and features. Cambridge,* Cambridge MS: MIT Press

[7] Flipsen, 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, 663-77

[8] Forrest, K., Weismer, G., Milenkovic, P. and Dougall, R.N. (1988), 'Statistical analysis of word-initial voiceless obstruents' *JASA*, 84, 115-123

[9] Macafee, C. (1983), *Glasgow*, Amsterdam: Benjamins

[10] Simpson, A. (2001), 'Dynamic consequences of differences in male and female vocal tract dimensions', *JASA*, 109, 2153-64

[11] Stevens, K. (1998), *Acoustic Phonetics*, Cambridge MS: MIT Press

[12] Strand, E. (1999), 'Uncovering the role of gender stereotypes in speech perception', *Journal of Language and Social Psychology*, 18, 86-99

[13] Stuart-Smith, J. (2003), 'The phonology of Modern Urban Scots', in *The Edinburgh Companion to Scots,* J. Corbett, D. McClure and J. Stuart-Smith, Eds. Edinburgh: Edinburgh University Press

[14] Wrench, A. (1995), 'Analysis of fricatives using multiple centres of gravity', *Proceedings International Congress of Phonetic Sciences*, 4, 460-463.