AN ACOUSTIC INVESTIGATION OF POSTVOCALIC /r/ VARIANTS IN TWO SOCIOLECTS OF GLASWEGIAN

Robert Lennon, Rachel Smith & Jane Stuart-Smith

Glasgow University Laboratory of Phonetics (GULP)
r.lennon.1@research.gla.ac.uk

ABSTRACT

This paper presents a small-scale acoustic investigation into postvocalic /r/ in both middle class and working class varieties of the Glaswegian accent. Tokens of /CVC/ and /CVrC/ minimal pairs (e.g. *hut/hurt*, *bead/beard*) were elicited from two middle-class and two working-class speakers, and the formant frequencies throughout the V(r) portion were analysed. The results show significant differences in the formant patterns across both varieties and across vowel environments, for minimal pairs such as *bead/beard* and *hut/hurt*. The middle class minimal pairs are acoustically distinct throughout the V(r) portion; the working class *hut* and *hurt* pairs differ only at the end, and only in F2, potentially causing misperception for listeners. These results support previous work on /r/ in working class speech in Glasgow. The results also support previous work on the characteristics of higher formants in bunched tongue configurations of /r/.

Keywords: sociophonetics; Scottish English; rhotic; derhoticisation; acoustic cues.

1. INTRODUCTION

In Glasgow, speakers are stereotypically rhotic [14]. However, recent sociophonetic research indicates a trend towards the loss of postvocalic /r/ in working class (WC) Glaswegian speech, leading to 'derhoticisation' [13], while a simultaneous increase in rhoticity in middle class (MC) Glaswegian speech is underway [6]. This means that the realisation of postvocalic /r/ is diverging over time between speech communities in Glasgow.

The change in postvocalic /r/ in Scottish English was first observed auditorily (e.g. [4] [12]), and has since been investigated extensively using Ultrasound Tongue Imaging (UTI). Lawson and colleagues (e.g. [6]), found that derhoticising working class speakers display a retracted tongue root configuration (causing a degree of pharyngealisation) in combination with a delayed, post-voicing tip-up gesture, leading to a vowel-like quality. They also found that hyper-rhotic middle class Scottish speakers use a bunched tongue configuration similar to the American English shape described by [1]. When comparing bunched and retroflex articulations in American English, [15] found that F4 and F5 are much closer in bunched /r/ than in retroflex /r/.

Acoustic work on Scottish /r/, however, has been very limited. When analysing working class /r/ in Glasgow, [13] found that in /Car/, /CarC/ and /CaC/ words (e.g. *car*, *heart*, *cat*), those with /r/ tended to have a longer rime and showed more retracted qualities in the vowel, than those without /r/. In these derhoticised /r/ variants, F2 was lower than in words without /r/, and F3 was higher, possibly reflecting uvularization.

There has been no detailed acoustic study of Scottish middle class rhoticity (though briefly discussed in [7]), however there has been acoustic analysis of hyper-rhoticity in other varieties. The proximity of F3 to F2 in approximant /r/ variants is noted by some authors to be important for a strong perception of rhoticity (e.g. [5] [9]). However, Heselwood, Plug and colleagues have written that the most important feature for rhoticity is a strong perceptual peak around the F2 region, whether that is achieved from a combination of F2+F3 or, as they found in experiments, absence of F3 entirely (as in [2] [3]).

The vowel-like nature of working class derhoticised /r/ variants can lead to misperception of certain words in some vowel environments, such as the minimal pair *hut/hurt*. In contrast, as middle class speakers are displaying an increase in rhoticity, this misperception would not be expected. [8] tested listeners' ability to discriminate between such pairs in middle class and working class Glaswegian speech, as a function of their experience with Glaswegian. Middle class minimal pairs were distinguished well by all listeners. However, only native Glaswegians could accurately identify derhoticised tokens of pairs like *hut* and *hurt*. Inexperienced listeners from Southern England identified these very poorly.

The present paper explores the acoustic underpinnings of these perceptual results, by providing a detailed analysis of the formant frequencies over time in the stimuli used in [8]. The aim of this research is therefore to examine the acoustic contrasts between V and Vr words (e.g. *hut/hurt*) for each of the two varieties in Glasgow, in order to better explain listeners' ability to distinguish between these minimal pairs.

2. METHODOLOGY

2.1. Recordings

The recordings analysed here were made for [8], and were segmented from connected speech, to be used as stimuli. Two pairs of native Glaswegian males (2xMC, 2xWC; 22-25 years) were recorded in a sound-attenuated booth, using lightweight Beyerdynamic headset microphones, at a sampling rate of 44.1kHz. Each pair of speakers was recorded separately. Each pair took part in a collaborative word-finding task: this meant that speech was as naturalistic as possible, while still ensuring that the full set of target words was produced by each speaker.

Table 1: Minimal pairs used in the experiment

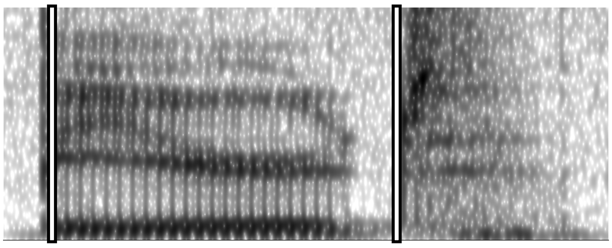
Close front	Unround. open-mid back
bead/beard	*bud/bird*
feed/feared	*hut/hurt*
weed/weird	*thud/third*

The words analysed here are 6 sets of minimal pairs, listed in Table 1. Each word was produced between 1 and 3 times by each of the four speakers (average 2.35 repetitions per speaker), totalling 113 tokens. The main hypothesis in the perceptual experiment concerned the potential for confusion between *hut* and *hurt* words in the working class variety, due to the proximity of pharyngealised /r/ to the location of the preceding unrounded open-mid back vowel. Therefore, this vowel environment was chosen for testing listener responses, along with the articulatorily distant unrounded close front vowel for comparison. The stimuli are representative of both middle class and working class Glaswegian speech.

2.2. Segmentation & formant analysis

In order to analyse the stimuli, each file had to be temporally segmented according to the vocalic portion of the word, including both vowel and /r/, if present (e.g. Figure 1). A similar segmentation protocol was used by Stuart-Smith [13] and Plug and Ogden [10], both of whom measured the rime in /r/-coda words. For each stimulus, the first boundary was placed at the onset of vocalic formant structure, and the second was placed at the start of the burst of the final alveolar plosive.

Figure 1: Segmented middle class *beard* stimulus.



Formants F1 to F5 are analysed here, as this allows for a very detailed investigation of formant patterns. To allow for more accurate plotting of formant tracks, the Python-based program Formant Editor [11] was used to correct any fluctuations in automatic formant tracking.

It has previously been noted that derhoticised /r/ is often accompanied by weakened formant amplitudes (e.g. [6]), and while the points at which the formants dropped in amplitude were tagged in the analysis, this paper focuses only on formant frequencies.

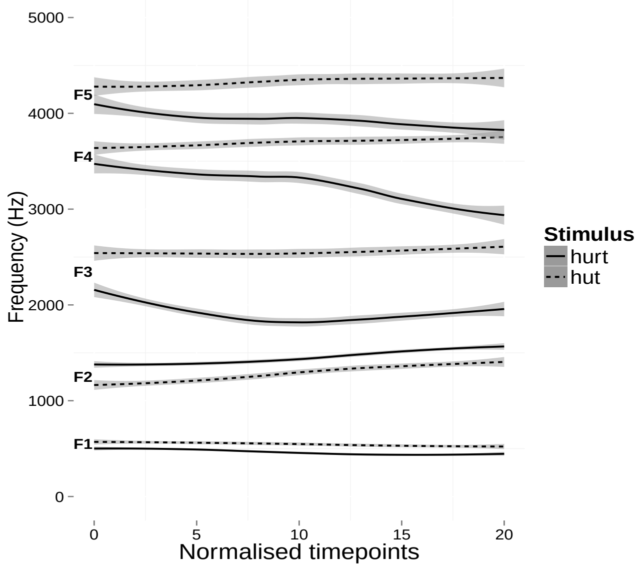
3. RESULTS

The results of the formant tracks for all five formants are shown in Figure 2. All five formants show clear differences relating to coda structure, vowel quality, and social class. These were confirmed by Linear Mixed Effects regression modelling using lme4 in R which took each formant as a dependent variable, and assessed variation across the time course of the track by modelling measurement_number as a fixed factor, with interactions for the factors of interest, Social class, Coda structure, Vowel quality, and Duration. After model comparison using the step() function, the optimal model for all five formants contained significant interactions ($p \leq 0.01$) for measurement_number*class*vowel, and for F2-F5 for measurement_number*class*coda ($p < 0.001$) (four way interactions could not be included because of sample size). We do not present the statistical results further here, but observations of differences in the descriptive tracks are also those which were found to be statistically significant ($p < 0.05$).

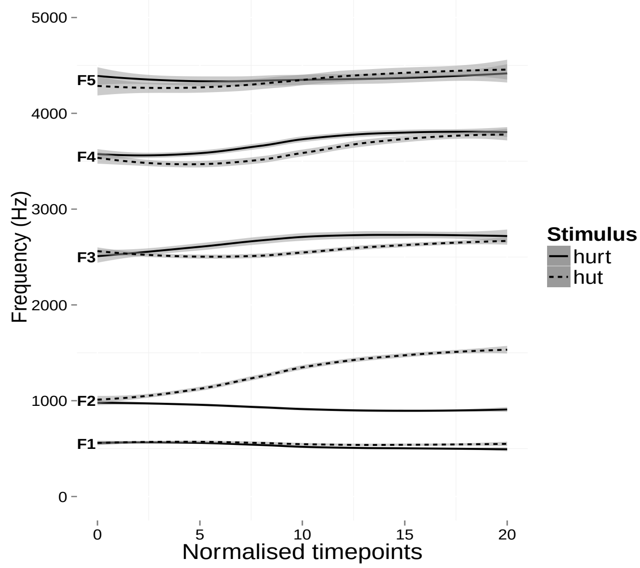
3.1. Acoustic variation and coda structure

Presence or absence of /r/ is signified by a difference in all formants between coda structure, for middle class speakers; but only in some formants for working class speakers. Most similar are the stimuli patterns in 2(b), where the only difference between

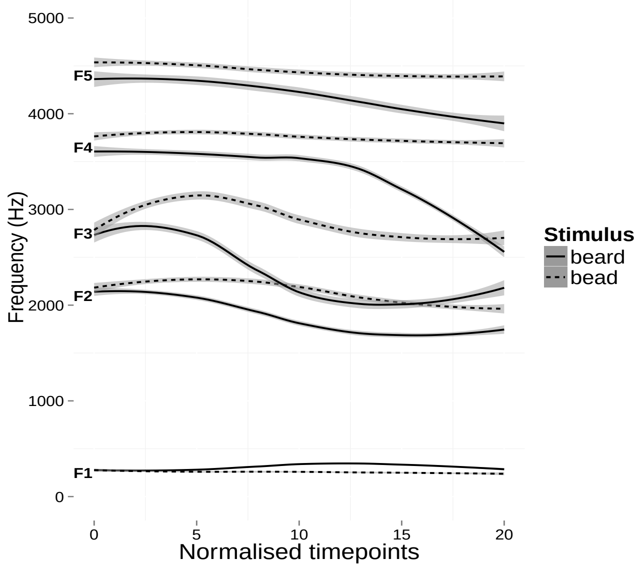
Figure 2: Formant tracks F1-F5 for all stimuli, by Class and Vowel. The /r/ stimuli are represented by solid lines, and /r/-less by dotted lines. E.g. *'hut'* represents all *bud,hut,thud* stimuli. Plotted in R using ggplot2's stat_smooth function to draw formant tracks. Shaded ribbons represent 95% C.I.



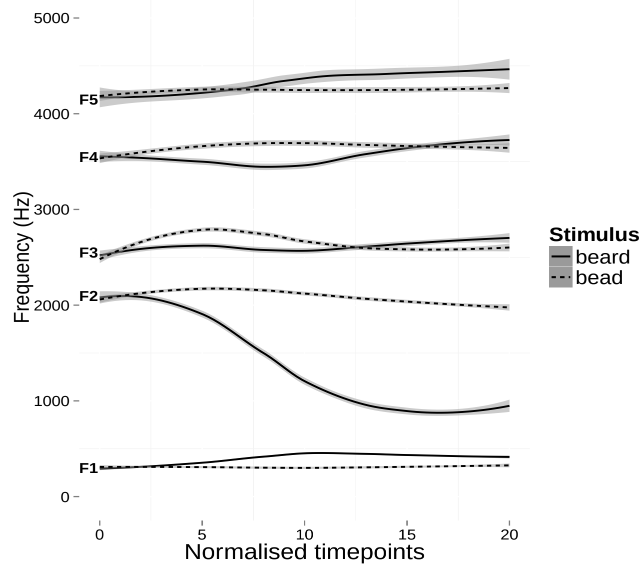
(a) Middle class *hut/hurt*, V(r) portion.



(b) Working class *hut/hurt*, V(r) portion.



(c) Middle class *bead/beard*, V(r) portion.



(d) Working class *bead/beard*, V(r) portion.

working class stimuli is rising F2 in *hut*.

3.2. Acoustic variation and vowel quality

The main difference between *hut/hurt* stimuli and *bead/beard* stimuli is the starting position of F2 in all graphs (compare 2(a&b) with 2(c&d)). Then, in *beard* stimuli F2 drops for the /r/, especially for working class speakers, whereas F3 rises slightly in working class speakers. In all front vowel stimuli, the contrast between /r/ stimuli and /r/-less stimuli is clear. The measurement_no*class*vowel interaction for F3 shows that F3 is lower for back vowels but higher for WC, and across the vocalic portion F3 rises. The measurement_no*class*vowel interaction for F4 shows that F4 is slightly lower for WC for back vowels across the vocalic portion. The same interaction for F5 shows that F5 is not as low for working class back vowel words, across the vocalic portion. Finally, the class*vowel interaction for F5 shows that F5 rises in working class speakers for back vowel tokens.

3.3. Acoustic variation and social class

In summary, F2 and F3 become very close in all middle class /r/ stimuli towards the end of the vocalic portion (solid lines in 2(a&c)), clearly showing their hyper-rhoticity. Conversely, the equivalent F2 and F3 tracks are much further apart for the working class speakers (solid lines in 2(b&d)), showing derhoticisation (similar to the large F2-F3 difference in [13]). Overall, the differences in these stimuli are further highlighted by the interactions described above.

3.4. Difference between higher formants

Table 2: Average higher formant values for all /r/ stimuli (unit: Hz), taken from normalised timepoints 10-15.

	MC /r/		WC /r/	
	beard	*hurt*	*beard*	*hurt*
F2	1728	1475	1009	901
F3	2044	1845	2604	2725
F2-F3 diff.	316	370	1595	1824
F4	3435	3235	3561	3773
F5	4139	3929	4480	4433
F4-F5 diff.	704	694	919	660

While no articulatory investigation has been conducted here, the differences between F4 and F5 in the middle class /r/ stimuli (Table 2: F4-F5 diff.) are comparable to the acoustic findings of Zhou et al. [15], who found that bunched /r/ in American English males showed a difference between F4 and F5 of just over 700Hz (compared with 1400Hz for retroflex). Although there is a similar F4 and F5 difference in working class /r/ stimuli (average 790Hz), there is no similarity between classes in the difference between F2 and F3 (MC average 340Hz, WC average 1709Hz).

3.5. Duration of vocalic portions

Table 3 shows average durations of the segmented vocalic portions of the stimuli. Like the rime durations described by [13], the vocalic durations in working class /r/ stimuli are longer than in their /r/-less counterparts, though more so for *bead & beard* than for *hut & hurt*. The middle class speakers also show longer durations in /r/ words than those without /r/, but this is perhaps less interesting, as there is much more difference, both acoustically and auditorily, between middle class minimal pairs than between working class minimal pairs.

Table 3: Average duration of vocalic portion V(r) for all tokens, by type (unit: ms).

	Middle class	Working class
bead	174	208
hut	175	238
beard	253	308
hurt	216	273

The statistical modelling for formants also found a significant interaction of measurement_no*class*duration for all formants bar F5, showing that middle class and working class speakers also show different formant trajectories according to the duration of the vocalic portion. This is presumably because they are likely using different articulatory gestures (e.g. [6] [7]) with temporal patterns.

4. DISCUSSION

This paper has examined in detail the acoustic characteristics of both working class and middle class postvocalic /r/ variants in Glasgow, providing a detailed description of the acoustic contrasts between V and Vr words. The main finding is that by far the most acoustically similar word types are minimal pairs *bud/bird*, *hut/hurt* and *thud/third*, produced by working class speakers, supporting previous work on derhoticisation in Glasgow, e.g. [13]. The potential for misperception in these pairs is likely very high. In contrast, middle class speakers are acoustically hyper-rhotic, primarily because of the proximity of F2 and F3 in /r/ words, and they also show higher formant characteristics similar to bunched /r/ (e.g. [15]), but without articulatory analysis this cannot be taken as evidence for tongue configuration. Their minimal pairs are acoustically much more distinct, meaning misperception is less likely.

In summary, this analysis is important support for the primary research on the perception of these stimuli, and provides valuable acoustic information about rhoticity in Glasgow today.

5. REFERENCES

[1] Delattre, P., Freeman, D. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics*.

[2] Heselwood, B. 2009. Rhoticity without F3: Low-pass filtering, F1-F2 relations and the perception of rhoticity in 'NORTH-FORCE', 'START' and 'NURSE' words. *Leeds Working Papers in Linguistics & Phonetics* 14, 49–64.

[3] Heselwood, B., Plug, L. 2011. The role of F2 and F3 in the perception of rhoticity: Evidence from listening experiments. *17th International Congress of Phonetic Sciences (ICPhS XVII)* (August), 867–870.

[4] Johnston, P. 1997. Regional variation. In: Jones, C. (ed). *The Edinburgh History of the Scots Language.* Edinburgh: EUP, 433–513.

[5] Ladefoged, P. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Wiley-Blackwell.

[6] Lawson, E., Scobbie, J. M., Stuart-Smith, J. 2011. The social stratification of tongue shape for postvocalic/r/in Scottish English. *Journal of Sociolinguistics* 256–268.

[7] Lawson, E., Scobbie, J. M., Stuart-Smith, J. 2014. A socio-articulatory study of Scottish rhoticity. In: Lawson, R. (ed). *Sociolinguistics in Scotland* Palgrave Macmillan, 53–78.

[8] Lennon, R. 2014. The effect of exposure in cross-dialect perception: Hearing ambiguous /r/ variants in Glaswegian. (poster presentation, AMLaP 20, University of Edinburgh, 3-6 September 2014).

[9] Lindau, M. 1985. The story of /r/. In: Victoria Fromkin (ed). *Phonetic linguistics: Essays in honor of Peter Ladefoged.* Academic Press, 157–168.

[10] Plug, L., Ogden, R. 2003. A Parametric Approach to the Phonetics of Postvocalic /r/ in Dutch. *Phonetica* 60(3), 159–186.

[11] Sóskuthy, M. 2014. Formant Editor: Software for editing dynamic formant measurements (Version 0.8.2) [Software]. Available from https://github.com/soskuthy/formant_edit.

[12] Stuart-Smith, J. 2003. The phonology of Modern Urban Scots. In: Corbett, J., McClure, D., Stuart-Smith, J. *The Edinburgh Companion to Scots.* Edinburgh: EUP, 110–37.

[13] Stuart-Smith, J. 2007. A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents. *Proceedings of the 16th International Congress of Phonetic Science* 1449–1452.

[14] Wells, J. C. 1982. *Accents of English* volume 1. Cambridge University Press.

[15] Zhou, X., Espy-Wilson, C. Y., Tiede, M., Boyce, S. 2007. An articulatory and acoustic study of "retroflex" and "bunched" American English rhotic sound based on MRI. *INTERSPEECH* 5–8.

6. ACKNOWLEDGEMENTS

This work was supported by the ESRC.