ACOUSTIC AND ARTICULATORY CHARACTERISTICS OF RHOTICITY IN THE NORTH-WEST OF ENGLAND

Claire Nance, Maya Dewhurst, Lois Fairclough, Pamela Forster, Sam Kirkham, Takayuki Nagamine, Danielle Turton, Di Wang

Lancaster University
c.nance@lancaster.ac.uk

ABSTRACT

We present a sociophonetic, acoustic, and articulatory analysis of coda rhoticity in East Lancashire, North-West England. We analysed data from 24 participants aged 8–73 recorded at a public engagement event in Blackburn Market (598 tokens). Auditory analysis shows coda rhoticity is declining across generations, with speakers born after 1990 being mostly non-rhotic. Audible rhoticity is realised by lowered F3 and raised F2. GAMMs fitted across the vowel(+rhoticity) interval show that audibly rhotic tokens have a significantly smaller distance between F3 and F2 than audibly non-rhotic tokens in all vowel contexts. Our ultrasound analysis compares minimal pairs e.g. 'core' and 'caw'. Principal Component Analysis of tongue splines shows that speakers use different tongue shapes in auditorily rhotic tokens.

Keywords: rhoticity, ultrasound, acoustics, sociophonetics, language variation and change

1. INTRODUCTION

In this paper, we consider the distribution and realisation of residual coda rhoticity in East Lancashire, North-West England, for example in words like 'farm' and 'core'. Coda rhoticity is one of the classic and most widely studied sociophonetic variables across the English-speaking world [1], see [2] for a recent overview. In England, rhoticity appears to have been in decline since early modern times [3]. Dialect survey work and analysis of oral history data from the 20th century indicates pockets of coda rhoticity remaining in Central/East Lancashire, (as well as rural West Yorkshire, and south-West England) [4, 5], though 21st century surveys and sociophonetic analysis suggests that East Lancashire rhoticity is now in decline [6, 7].

Acoustic approaches to English rhoticity have largely focussed on lowered F3 as a cue to rhoticity e.g. [5]. However, other work suggests that formants below F3 also contribute to perceptions of rhoticity [8, 9]. In East Lancashire in particular, previous work indicates that audible rhoticity is realised with lowered F3 and also raised F2 [10, 11]. We therefore chose F3−F2 as the acoustic measure best representing rhoticity. We present the raw values of F3−F2 as comparing the distance between formants should normalise to some extent for differing vocal tract lengths, and further normalisation may over-normalise the data [11].

Previous articulatory work on rhotics in English suggests that speakers broadly use a retroflex or bunched tongue gesture to realise rhoticity [12, 13]. Tongue shape is usually investigated in the midsagittal dimension. Due to the importance of the midsagittal perspective, ultrasound tongue imaging has been widely used to study rhotic articulation cross-linguistically e.g. [14, 15]. In English in the UK, ultrasound has especially been used to study variable rhoticity in Scottish English e.g. [16, 17]. Ultrasound analysis in Scotland has shown considerable sociolinguistic variation in rhoticity realisation [16]. Our analysis represents the first articulatory study of rhoticity in England.

In this paper we investigate the following research questions:

1. How is rhoticity distributed according to different demographic factors?
2. How is audible rhoticity realised in acoustics?
3. How is audible rhoticity realised in articulation?

2. METHODS

2.1. Participants and setup

Our data collection was carried out as part of a public engagement event in the Blackburn Market, East Lancashire, inspired by [18]. We recorded a total of 35 participants. In this paper, we present data from 24 participants (14f, 10m) who were long-term residents of East Lancashire, mostly Blackburn. They were aged 8–73 ($\bar{x} = 39$; $\sigma = 17.5$). Blackburn has a large British Asian population and we recorded one participant who identified as British Pakistani, and two who were born in India but now work in Blackburn. Our speakers represented a range of occupations, as well as children and non-working people. Due to this variety, we have not analysed social class further but note that this will be an important dimension in future work.

2.2. Materials and recording

Participants were asked to read a list of words which were presented in the Articulate Assistant Advanced (AAA) software [19]. Words were displayed orthographically and with pictures. Participants were asked to read the words twice, but we only collected one repetition from some speakers due to time constraints. Our word list included nine words potentially containing coda rhoticity, and four distractors/minimal pair counterparts (Table 1). Note that in Lancashire the NURSE and SQUARE lexical sets are often merged [20].

Word	Lexical set	Word	Lexical set
beard	NEAR	goat	GOAT
caw	THOUGHT	paw	THOUGHT
core	NORTH/FORCE	pour	NORTH/FORCE
cake	FACE	stair	SQUARE/NURSE
farm	START	stir	NURSE
fair	SQUARE/NURSE	worm	NURSE
fur	NURSE		

Table 1: Word list included in the study.

We used the Telemed MicrUs system to record our ultrasound data into AAA, with a 64-element probe, 20mm radius. The probe frequency was 2MHz, depth was 80mm, and field of view was 90–100% resulting in a frame rate of 80–90Hz. The probe was held in place using a plastic helmet [21]. The acoustic data were recorded with a headset microphone attached to the helmet. Acoustic data were recorded onto a laptop in AAA with a Sound Devices USBPre2 audio interface at 22,050Hz.

2.3. Analysis

Data were exported from AAA into Praat for auditory and acoustic analysis. Each token was first coded auditorily for the presence/absence of coda rhoticity. The interval containing the vowel plus any following rhoticity was then labelled, and extracted for analysis. The first three formants were estimated in Fasttrack [22]. The optimal analyses from Fasttrack were then analysed further in R. Acoustic data from one participant were excluded due to very poor audio quality. Ultrasound data were excluded for two participants due to poor image quality. TextGrids were then imported into AAA and the labels serve as landmarks for the articulatory analysis. Tongue splines were generated using DeepLabCut [23] in AAA at 11 key points on the tongue, and then rotated to the occlusal plane for each participant using bite plate traces [24] (except for one participant). Further analysis was done in R.

The analyses in this paper are structured around our research questions. Code and data for these analyses are available at https://osf.io/wb9m5/.

Analysis 1: Auditory coding results from words containing orthographic ‘r’ are presented according to speaker and vowel context. We then consider variation in the data according to age, gender and ethnicity. These results are analysed with a logistic mixed effects regression model: rhoticity $\sim$ year of birth + gender + (1|word) + (1|speaker). Ethnicity and vowel context are explored qualitatively due to small token counts. 598 tokens from 24 speakers.

Analysis 2: Acoustic data from words containing orthographic ‘r’ were divided by their auditory coding of ‘rhotic’ or ‘non-rhotic’. We then used GAMMs [25] to compare the F3$-$F2 values across 11 timepoints in the vowel(+rhoticity) interval in different vowel contexts. GAMMs were fitted following [26]. Predictor variables included a parametric term of auditory coding for rhoticity and smooth terms of normalised time and a normalised time-by-rhoticity interaction. We also fitted random smooths of time-by-speaker and time-by-word. For significance testing, we compared a full autoregressive model to nested models excluding all rhoticity predictors to investigate for significant differences in trajectory height, and then the full model to a model excluding rhoticity-by-time to investigate for significant differences in trajectory shape [27]. 394 tokens from 23 speakers.

Analysis 3: Ultrasound analysis focusses on two rhoticity minimal pairs: *core/caw* and *pour/paw* at 80% duration of the vowel(+rhoticity) interval. We compare the tongue shapes used by speakers visually, and then via Principal Component Analysis for statistical comparison of auditorily rhotic vs. non-rhotic tongue shapes [28, 15]. Values of PC1 accounted for 95% of the variation in the dataset. We therefore tested z-scored PC1 values comparing words containing ‘r’ (*core, pour*) against words not containing ‘r’ (*caw, paw*) for auditorily rhotic and non-rhotic speakers. This was carried out via linear mixed effects model of the formula PC1z $\sim$ r-ful word*auditory perception + (1|speaker) + (1|word). 120 tokens from 16 speakers.

3. RESULTS

3.1. Distribution of rhoticity according to demographic factors

Auditory perceptions of rhoticity are shown in Figure 1 for individual speakers and words.

Statistical testing considered age and gender of participants. Due to the correlation between vowel

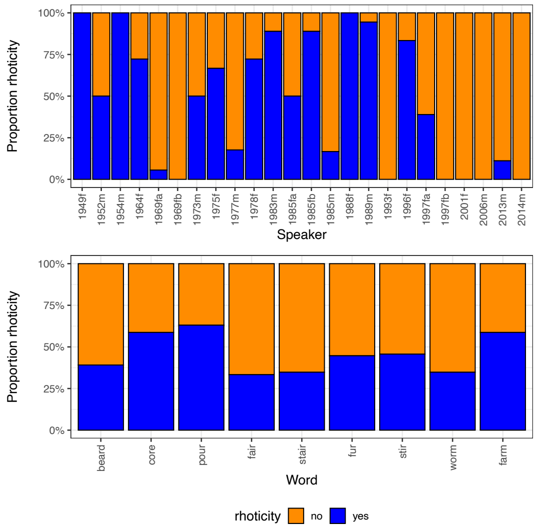

Figure 1: Top panel: Auditory coding of rhoticity for individual speakers. Speakers are ordered by year of birth oldest to youngest left to right. Speaker number shows year of birth and gender. Bottom panel: Auditory coding of rhoticity for individual words.

and word, we have modelled the vowel effect as a random intercept of word only. Visual inspection of Figure 1 indicates that rhoticity is most present in words from the NORTH/FORCE and START lexical sets, though there are not large differences between vowels. In terms of ethnicity, our British Pakistani speaker was completely non-rhotic. The two Indian speakers were almost 100% rhotic and used taps or trills, while the White British speakers used pharyngealisation or approximant rhotics. Model comparison shows a significant effect of age, where older speakers are more likely to be rhotic ($\chi^2(1) = 8.19$; $p(\chi^2) = .004$). Gender was not significant.

3.2. Acoustics of audible rhoticity

Here, we compare the F3−F2 values of auditorily rhotic and non-rhotic productions in the vowel(+rhoticity) interval via GAMMs. Results are shown in Figure 2, and the results of model comparisons testing for significant differences in trajectory height and shape are in Table 2. For significance testing via model comparison, we used an AR1 model with rho estimated as the autocorrelation at lag 1. See code for information (https://osf.io/wb9m5/). Results show significant differences for height in every vowel context, and significant differences for shape in NORTH/FORCE, SQUARE, and START words. In each case, F3−F2 is lower in auditorily rhotic tokens.

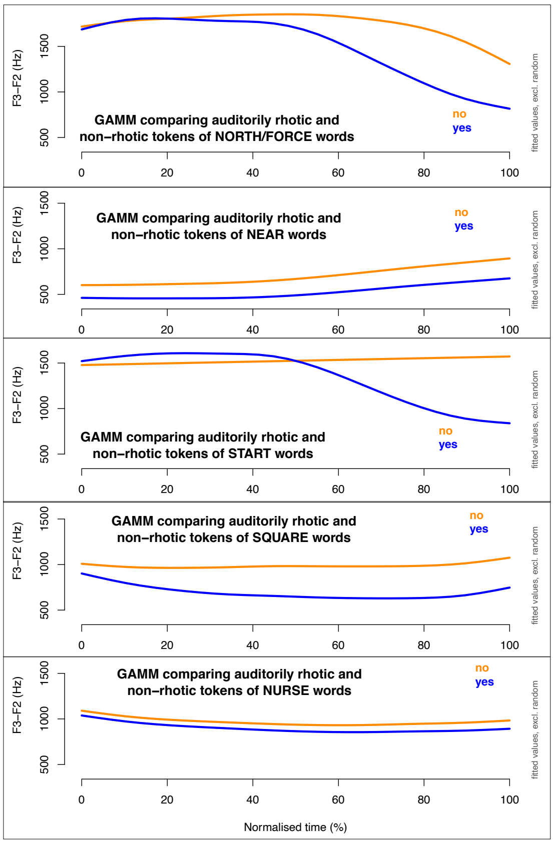

Figure 2: GAMMs comparing auditorily rhotic (blue) and non-rhotic (orange) tokens. Top panel NORTH/FORCE words; second panel NEAR words; third panel START words; fourth panel SQUARE words; fifth panel NURSE words.

Lexical set	Trajectory height	Trajectory shape
NORTH/FORCE	$\chi^2(3) = 26.7; p < .001$	$\chi^2(2) = 24.1; p < .001$
NEAR	$\chi^2(3) = 14.8; p < .001$	$\chi^2(2) = 0.32; p = .72$
START	$\chi^2(3) = 16.5; p < .001$	$\chi^2(2) = 16.5; p < .001$
SQUARE	$\chi^2(3) = 40.3; p < .001$	$\chi^2(2) = 8.50; p < .001$
NURSE	$\chi^2(3) = 7.52; p = .002$	$\chi^2(2) = 0.33; p = .71$

Table 2: Acoustic GAMM model comparison showing differences between auditorily rhotic and non-rhotic tokens.

3.3. Articulation of audible rhoticity

Here, we compare the tongue shapes used for *core/caw* and *pour/paw*. We have plotted the splines for the sixteen speakers analysed so far in Figure 3. The DeepLabCut output was smoothed using a 10 Hz first-order Butterworth low-pass filter. In the eight speakers who contrast for rhoticity, *core/pour* are produced with a raised and fronted tongue compared to the back vowel in *caw/paw*.

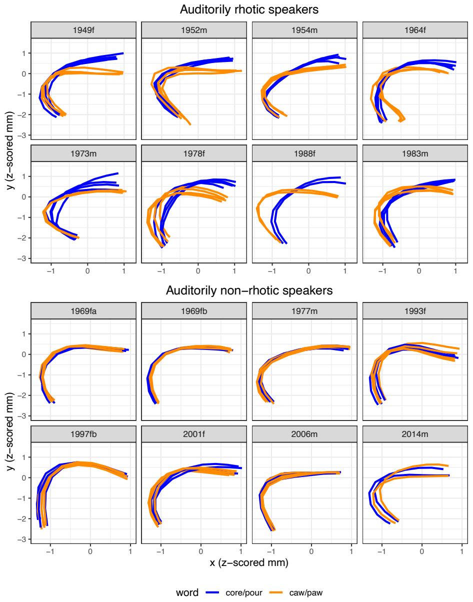

Figure 3: Tongues splines comparing rhoticity minimal pairs in auditorily rhotic and non-rhotic speakers. Code for this: https://osf.io/wb9m5/

The differences in tongue shape in Figure 3 were compared with a PCA run on the spline values. We report values of PC1 only, as this accounted for 95% of the variation in the dataset. (PC2 accounted for 2% and higher PCs even lower values). The variation in PC1 is plotted in Figure 4, along with a comparison of PC1 values in auditorily rhotic and non-rhotic speakers. PC1 appears to capture variation in tongue fronting and height.

The values of PC1 were modelled to test for an interaction between perceived rhoticity and minimal pair. Model comparison indicates that this interaction is significant ($\chi^2(2) = 23.6; p < .001$), meaning that for speakers who produce audible rhoticity, tongue shapes are significantly different for *core/pour* vs. *caw/paw*.

4. DISCUSSION AND CONCLUSIONS

Production of coda rhoticity appears to vary across our sample. However, we found a general pattern towards decline in rhoticity according to age, supporting previous work in [7, 6, 11]. In our data, young people born after 1990 are mostly non-rhotic. We were not able to fully investigate variation across vowel contexts due to the small number of words in our dataset, but it is interesting to note that NURSE words were not necessarily the most rhotic as has been suggested in other work on rhoticity in England [2]. The three British Asian speakers in our dataset show some interesting patterns based on the extent to which they have acquired a South Asian English variety (rhotic), or grew up in Lancashire (one young person, non-rhotic). This variation would be interesting to investigate in the future in more depth. Our acoustic results indicate that F3−F2 is a good measure of audible rhoticity in East Lancashire, especially for non-central vowels. It was less good for the NURSE vowel, though still showed significant differences in trajectory height. Our preliminary analysis of the ultrasound data shows differences in the tongue shapes used in minimal pairs where speakers produce audible rhoticity.

To conclude: rhoticity is still relatively widespread in East Lancashire, even though it is now very rare in the north of England generally [7]. This appears to be changing in apparent-time though future larger-scale work is now needed. We presented the first articulatory analysis of rhoticity in England and demonstrate that different tongue shapes are used for auditorily rhotic and non-rhotic tokens. Future work will look in more detail at the timing of rhotic gestures, as well as a larger set of lexical items. We are also interested in the possibility of differing voice qualities and/or articulatory settings across generations in East Lancashire (similar to the description of Glasgow in [29]) and would like to investigate this further with acoustic and articulatory methods.

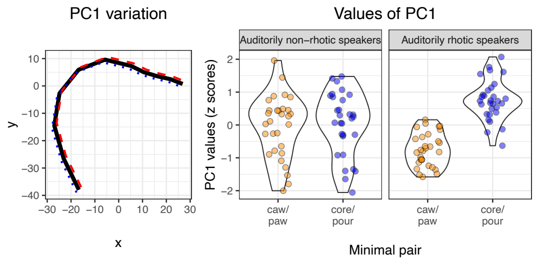

Figure 4: Left panel: variation explained by PC1. Black line shows tongue mean, red dashed line shows maximum values, blue dotted line shows minimum values.
Right panel: Values of PC1 according to perceived rhoticity and whether the word contains ‘r’.

5. ACKNOWLEDGEMENTS

Thank you to our participants, Blackburn Council, and Blackburn Market. Thank you to the ESRC Festival of Social Sciences for funding.

6. REFERENCES

[1] W. Labov, *The social stratification of English in New York City*. Washington D. C.: Center for applied linguistics, 1966.

[2] T. Blaxter, K. Beeching, R. Coates, J. Murphy, and E. Robinson, "Each person does it their way: Rhoticity variation and the community grammar," *Language Variation and Change*, vol. 31, no. 1, pp. 91–117, 2019.

[3] E. Gordon, L. Campbell, J. Hay, M. Maclagan, A. Sudbury, and P. Trudgill, *New Zealand English: its origins and evolution*. Cambridge: Cambridge University Press, 2004.

[4] H. Orton and E. Dieth, *Survey of English Dialects: The Six Northern Counties and the Isle of Man*, H. Orton and W. J. Halliday, Eds. Leeds: University of Leeds Press, 1962.

[5] S. Durkacz Ryan, H. Dann, and R. Drummond, "'Really this girl ought to be going to something better': Rhoticity and social meaning in oral history data," *Language in Society*, p. First view, 2022.

[6] W. Barras, "The sociophonology of rhoticity and r-sandhi in East Lancashire English," Ph.D. dissertation, University of Edinburgh, Edinburgh, 2010.

[7] A. Leemann, M.-J. Kolly, and D. Britain, "The English dialects app: The creation of a crowdsourced dialect corpus," *Ampersand*, vol. 5, pp. 1–17, 2018.

[8] B. Heselwood and L. Plug, "The role of f2 and f3 in the perception of rhoticity: Evidence from listening experiments," in *Proceedings of the 17th International Congress of the Phonetic Sciences*. Hong Kong: City University Hong Kong, 2011.

[9] D. Villarreal, L. Clark, J. Hay, and K. Watson, "From categories to gradience: Auto-coding sociophonetic variation with random forests," *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 11, no. 1, pp. 1–31, 2020.

[10] B. Heselwood, "Rhoticity without F3: Lowpass filtering, F1-F2 relations and the perception of rhoticity in NORTH-FORCE, START and NURSE words," *Leeds Working Papers in Linguistics and Phonetics*, vol. 14, pp. 49–64, 2009.

[11] D. Turton and R. Lennon, "An acoustic analysis of rhoticity in Lancashire, England," *Journal of Phonetics*, Under review.

[12] P. Delattre and D. Freeman, "A dialect study of American r's by x-ray motion picture," *Linguistics*, vol. 6, no. 29-68, 1968.

[13] H. King and E. Ferragne, "Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English," *Journal of Phonetics*, vol. 80, pp. 1–19, 2020.

[14] M. Tabain and R. Beare, "An ultrasound study of coronal places of articulation in Central Arrernte: Apicals, laminals and rhotics," *Journal of Phonetics*, vol. 66, pp. 63–81, 2018.

[15] C. Nance and S. Kirkham, "Phonetic typology and articulatory constraints: The realisation of secondary articulations in Scottish Gaelic rhotics," *Language*, vol. 98, no. 3, pp. 419–460, 2022.

[16] E. Lawson, J. M. Scobbie, and J. Stuart-Smith, "The social stratification of tongue shape in postvocalic /r/ in Scottish English," *Journal of Sociolinguistics*, vol. 15, no. 2, pp. 256–268, 2011.

[17] J. Stuart-Smith, E. Lawson, and J. M. Scobbie, "Derhoticisation in Scottish English: a sociophonetic journey," in *Advances in Sociophonetics*, C. Celata and S. Calamai, Eds. Amsterdam: John Benjamins, 2014, pp. 59–96.

[18] M. Heyne, X. Wang, D. Derrick, K. Dorreen, and K. Watson, "The articulation of /r/ in New Zealand English," *Journal of the International Phonetic Association*, vol. 50, no. 3, pp. 366–388, 2020.

[19] Articulate Instruments, *Articulate Assistant Advanced version 2.20.2*. Edinburgh: Articulate Instruments, 2022.

[20] K. Watson and L. Clark, "How salient is the Nurse-Square merger?" *English Language and Linguistics*, vol. 17, no. 2, pp. 297–323, 2013.

[21] L. Spreafico, M. Pucher, and A. Matosova, "UltraFit: A speaker-friendly headset for ultrasound recordings in speech science," in *Interspeech 2018*, 2018.

[22] S. Barreda, "Fast track: fast (nearly) automatic formant-tracking using praat," *Linguistics Vanguard*, vol. 7, no. 1, 2021.

[23] A. Wrench and J. Balch-Tomes, "Beyond the edge: Markerless pose estimation of speech articulators from ultrasound and camera images using DeepLabCut," *Sensors*, vol. 22, no. 3, p. 1133, 2022.

[24] J. M. Scobbie, E. Lawson, S. Cowen, J. Cleland, and A. Wrench, "A common co-ordinate system for mid-sagittal articulatory measurement," *QMU CASL Working Papers*, vol. 20, 2011.

[25] S. N. Wood, *Generalized Additive Models*. Chapman and Hall/CRC, 2017.

[26] M. Sóskuthy, *Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction*, 2017. [Online]. Available: https://arxiv.org/abs/1703.05339

[27] S. Kirkham, C. Nance, B. Littlewood, K. Lightfoot, and E. Groarke, "Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English," *Journal of the Acoustical Society of America*, vol. 145, pp. 784–794, 2019.

[28] D. Turton, "Categorical or gradient? an ultrasound investigation of /l/-darkening and vocalization in varieties of English," *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 8, no. 1, pp. 1–31, 2017.

[29] M. Sóskuthy and J. Stuart-Smith, "Voice quality and coda /r/ in Glasgow English in the early 20th century," *Language Variation and Change*, vol. 32, pp. 133–157, 2020.