

QML 2025 Final Reflection

Your Name

Group Project title	Exploring the FOOT-STRUT split across different accents
Group members	Eric Ferrari, Cai Fletcher, Calvin Jubb, Eleanor Butterworth

Reflection

Introduction

The distinction between the FOOT and STRUT vowels - also known as the FOOT-STRUT split - is a widely studied phenomenon in English dialectology (c.f. Beal, 2012). Many varieties of English, such as Southern British or American English, realise these vowels as acoustically and phonologically distinct (split). Other varieties, including Northern English and Irish accents, do not show this contrast and instead employ a singular vowel for both lexical sets (merged) (Schreier et al., 2013).

Our study examines the FOOT-STRUT realisations in three accent communities - Scottish, Northern English, and American - highlighting the distinctions between them on the grounds of vowel differences in a specific FOOT-STRUT test. With two communities having expected splits (Scottish and American) and one having an anticipated merge (Northern English). By analysing participants' acoustic data, we assessed the extent to which the Foot and Strut vowels were distinct or merged, and highlighted the specific vowel realisations employed to further illustrate the quantitative differences between the data sets. Numerical frequency measurements of vowel quality—specifically F1 and F2 values—provided a basis for comparing vowels, and their phonetic differences across accents, while also revealing the internal variations within each accent group.

Ultimately, our findings suggest that each community maintained its expected pronunciation of the target words, with some minor inter-community variation among the Northern English participants.

Background

Historically, the FOOT-STRUT split refers to the divergence of the Middle English vowel /u/ into two distinct vowels. In most American accents, the split is complete: foot tends to have /ʊ/ while strut has /ʌ/. (Honeybone et al., 2008) Scottish English, however, developed differently, and despite not undergoing the same historical split as Southern British English, it maintains a clear contrast between foot realised with the high central vowel /ə/ and strut realised with the vowel /ʌ/. (Smith et al, 2024, p. 157) Comparatively, many Northern English accents lack the split and instead employ /ʊ/ (or similar variants) for both sets of words (Turton et al., 2021 p. 163-201).

A number of studies have examined these patterns across dialects. J. C. Wells's Accents of English (1982) provides a foundational description of the split and its uneven geographic distribution globally. (Wells, 1982). Further sociolinguistic research, such as Hughes' investigations on the regional varieties of English in England, analysed the absence of the split in Northern communities (Hughes et al., 2005). Furthermore, the work of Aitken explored Scotland's unique history and experience with this split, highlighting its distinct developmental path from the Southern English and American split (Aitken, 1984).

Methodology

To elicit examples of the FOOT and STRUT vowel, the study takes the form of a speech production experiment. Twelve sentences were used in the study as stimuli. Six of the sentences contained the target vowels: three instances FOOT vowels and three instances STRUT vowels for an even split. The target vowels were placed in identical phonemic environments so that they were between the phonemes /l/ and /k/ to create the lexemes 'look' and 'luck.' The choice to stick with one phonemic environment was made to control the vowel quality. The six remaining sentences were filler sentences. All sentences were taken from the Merriam Webster Dictionary (n.d.).

Ethics approval was obtained by the LEL Research Ethics panel by the course organizer. Participants were recruited from social networks and from within the research group. Six speakers, two each from Glasgow, Lancashire, and Massachusetts, were recruited. They were presented with the nature of the study and were prompted to sign ethics forms if they consented to participate. They were directed to a soundproof room with one of the experimenters. They were seated roughly 30 centimeters in front of an Andrea SG-110M Shotgun Microphone to record the sentences displayed on a screen. These recordings were captured in Praat with a sampling rate of 44100 Hz. The sentences were randomized to determine their order of presentation, but this randomized order was fixed for every participant. They repeated each sentence three times, leading to 18 tokens per participant by saying each of the three sentences with the FOOT vowel three times, and each of the three STRUT sentences three times. This produced 108 instances of the target vowel across all studied speakers.

The recordings and ethics forms with the target vowels were then loaded into separate, secure folders within the University of Edinburgh OneDrive. The target recordings were loaded into Praat. Each vowel was manually segmented, and the formants for each vowel were manually extracted at their midpoint and loaded into a spreadsheet for the analysis. The collected data was normalised using Lobanov normalisation. Within speaker z-scores were calculated for each formant; these z-scores were then multiplied by the across speakers' standard deviation and added to the across speaker mean. This normalises the data removing the effect of different length vocal tract on the formant frequencies, allowing for better comparison between speakers. It was additionally checked that no z-scores were more than 3 S.D. away from the mean. Later, normalisation was carried out at the across speaker level to preserve data for multilevel regression models using by-speaker intercept varying terms which are lost in Lobanov normalisation.

Initial Exploratory Analysis

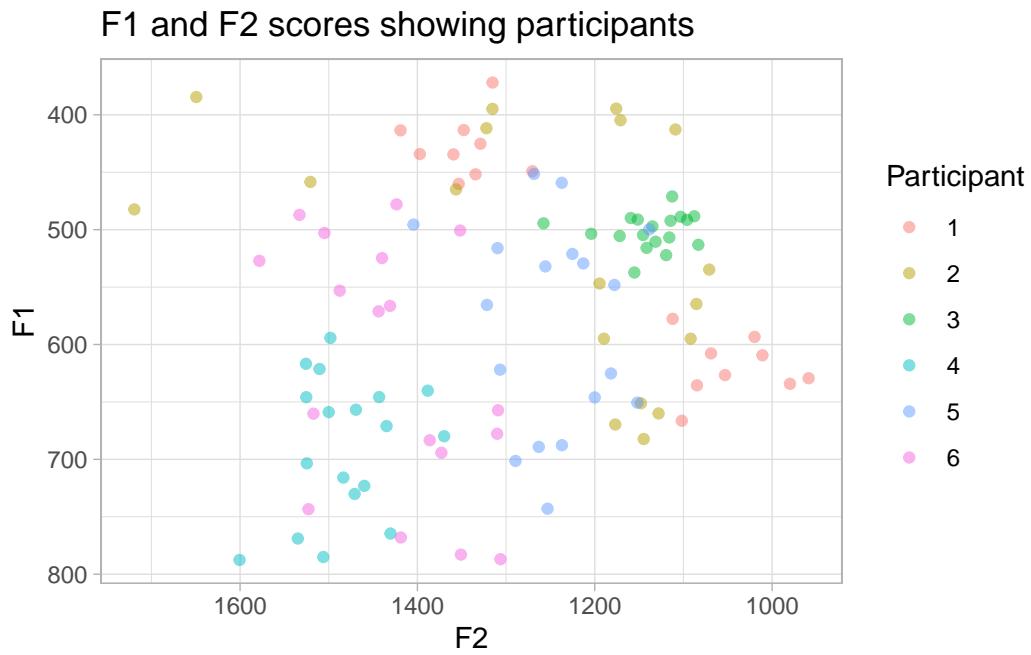


Figure 1: Formant and Participant data scatterplot

The graph above Figure 1 shows all the initial F1 and F2 data in a scatterplot prior to normalization. In this figure, there are already defined splits and merges among the accent groups. Participants 1 and 2 are Scottish, 3 and 4 are Northern English, and 5 and 6 are American. We can see that the Northern English participants, green and teal, are lumped into two tightly defined clusters, one for each Northern English participant. These suggest the initial represen-

tations of a merger in the data. Interestingly, this also shows differing vowels for the different participants – both still adhere to the merged expectation but utilise different vowels to do so. This is an irregularity that will be resolved upon normalisation of the data; however, it could be explained by the demographic data of this participant (their gender, age, where they live, native languages, etc.). The data for Scottish participants, in red and yellow, is spread out across the graph. There is, however, a clear intermixing of red and yellow in similar spots, such as the top middle and left middle. This suggests a split in the FOOT and STRUT vowels for the Scottish accent group. Similarly, the American data, in blue and purple, is rather spread out over the graph. This also suggests a split, albeit broader, in which individuals use a wider range of FOOT and STRUT vowels in their vocal production. However, given that this is non-normalised data, these scatterings and differing clumps could result from vocal features such as participant gender, pitch, or intonation. See Figure 2 for normalised data and a more comprehensive view of the splits and mergers.

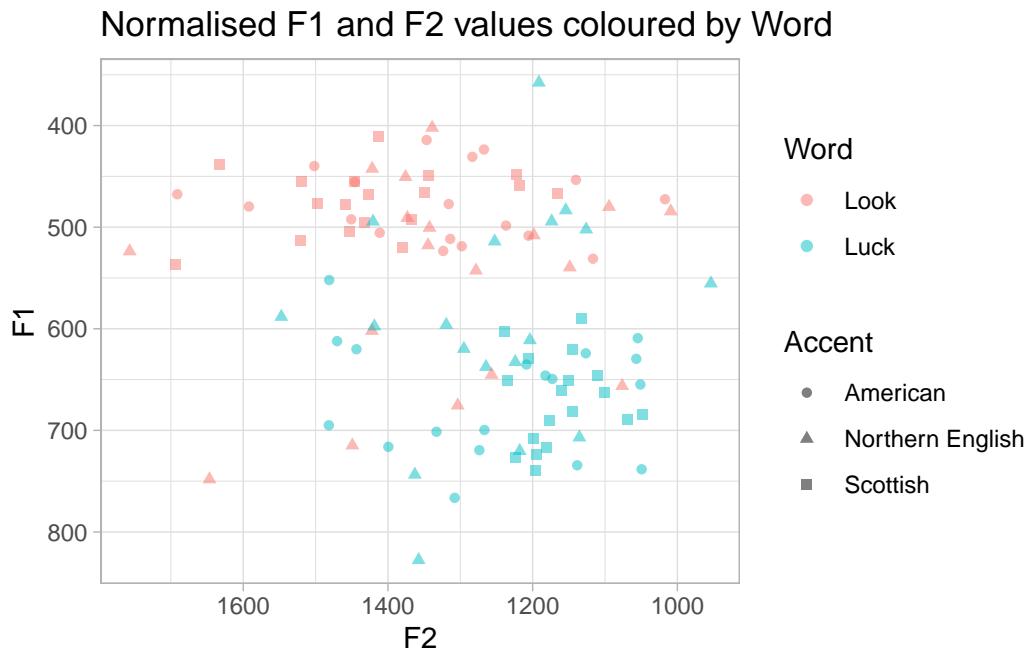


Figure 2: Scatter plot showing the F1 and F2 values of the speakers after Lobanov normalisation

Figure 2 illustrates the normalised data in scatterplot format. This data was normalised using the Lobanov processes and provided with a z-score. In doing so, the data is measured in standard deviation (SD) and any distinct pitches or vocal qualities were removed within the formant data. Following this, the z-scores were converted back to Hz using the SD across all speakers, thereby creating a more accurate model with standardised vocal ranges for each participant. This time Figure 2 focuses specifically on colour-coding for each word instead of the specific accent community (which can be seen in Figure 3). Ultimately, there is still a

visible split within both the Scottish and American accents, but the Northern English Accent datapoints show a clearer merger in the normalised values, as seen by its extensive distribution and intermixing across the entirety of the graph.

F1 and F2 distributions

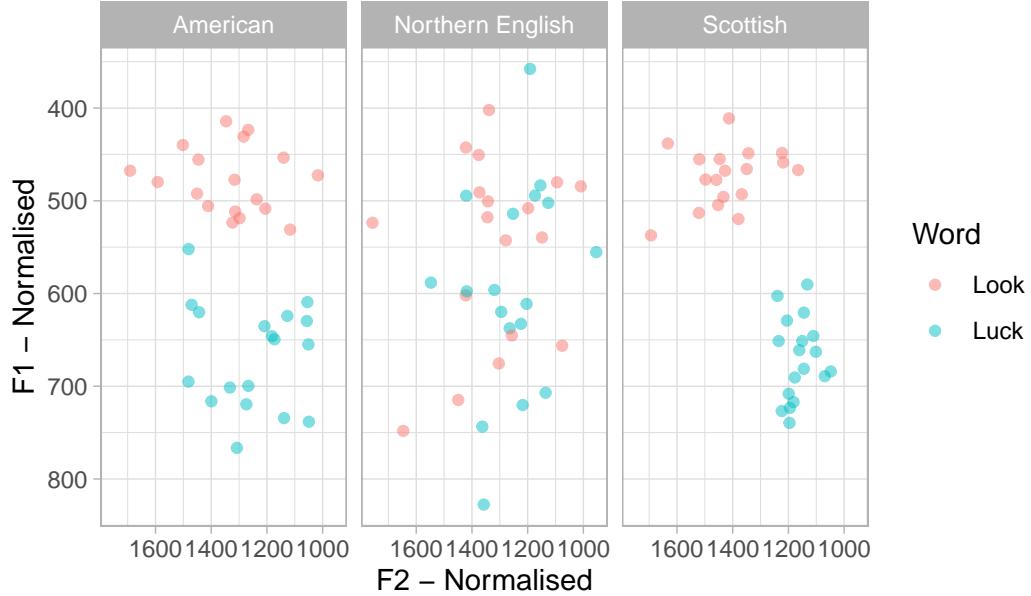


Figure 3: Scatter plot showing the F1 and F2 values of the different accents

Figure 3 takes the previous graph Figure 2 and normalised data and breaks into the specific accent communities. Within these accent groups, a clearer representation of the vowel productions is visible. In the American dataset there are two clear clusters, one for the vowel in the FOOT set and another for the vowel in the STRUT set. These differing sections illustrate a sharp contrast between the two vowels, with the FOOT vowel sitting at a lower F1 and the STRUT vowel at a higher F1. Ultimately, this means that FOOT occupies a higher vowel space, as a lower F1 equates to a higher vowel (Nasution et al, 2023) and STRUT occupies a lower vowel space, as a higher F1 signals a lower/back vowel (Nasution et al, 2023). Importantly, both these vowels occupy a similar F2 space (1600 – 1000Hz) which is representative of a back vowel (Nasution et al, 2023). A similar pattern can be observed for the Scottish accent group, with two clear groupings representative of a split. However, the STRUT vowel lands in an even lower F2 space (1200-1000Hz) meaning that it tends to be a further back vowel than the American realisation. Comparatively, the Northern English accent group is intermixed for both their vowels, signifying a merge. The vowels are employed broadly regardless of word distinction. However, they still occupy a high back vowel space, as is expected for the FOOT and STRUT vowel sets (o and u sounds are represented by high back/mid vowels).

Analysis of Regression Draws

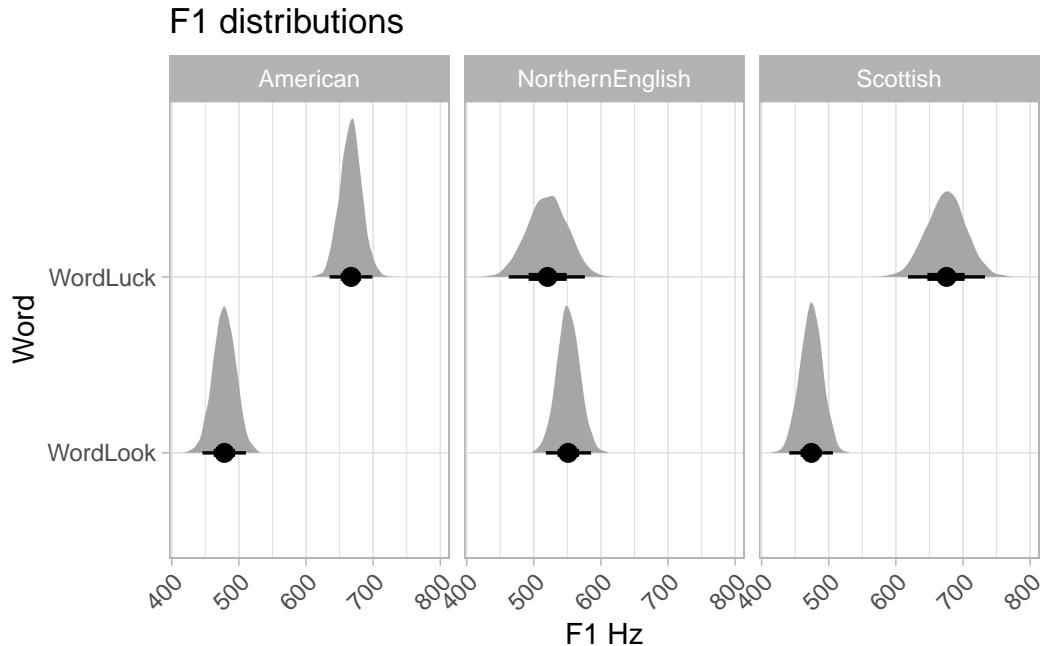


Figure 4: Gaussian Distributions of the F1 values within different accents/words

Figure 4 features a Bayesian Gaussian regression model to assess the F1 in Hz of vowels by the speakers' accent and the word they were saying. The model was fitted with the brms package (Bürkner, 2017) in R (R Core Team, 2025), with the default priors, four chains with 2000 iterations each, of which 1000 for warm-up. We used the normalised F1 as the outcome variable. We included accent (American, Northern English, and Scottish) and the word spoken ('Look' and 'Luck') as the regression predictors, which was coded with the default treatment contrasts (American word Look as the reference level). The numbers were then extracted using summary before being graphed.

According to the model, there is a 95% probability that the F1 for an American speaker saying 'Look' is between 445 and 510 Hz (mean = 478 Hz, sd = 17). For the word 'Luck' there is a 95% probability that the F1 of their vowel is between 143 to 235 Hz higher (mean = 189, sd = 23). Using this we can estimate the mean F1 value for an American speaker to be 667Hz.

There is a 95% probability that the F1 of Northern English speakers saying 'Look' is 28 to 119 Hz higher than for American speakers (mean = 73 Hz, sd = 24), giving a mean of 551 Hz. The F1 for Northern Speakers' 'Luck', is between 82 and 212 Hz lower than the American 'Luck' (mean = -146, sd = 34), which gives a mean of 521 Hz. The difference between the mean F1 of 'Look' and 'Luck' for Northern English accents is 30 Hz. This small difference demonstrates the predicted merger in Northern English accents.

For Scottish speakers saying ‘Look’ there is a 95% probability that the difference from the American baseline is between -50 and 42 Hz (mean = -4 Hz, $sd = 23$), giving a mean of 441 Hz. The F1 of their ‘Luck’ is between –58 lower to 73 Hz higher than the American ‘Luck’ (mean = 8 Hz, $sd = 34$), which gives a mean of 642 Hz. The difference between the mean F1 of ‘Look’ and ‘Luck’ for Scottish accents is 201 Hz. Therefore these differences of 189Hz and 201Hz respectively demonstrate the FOOT-STRUT split for American and Scottish accents.

The residual standard deviation is between 62 and 81 Hz (mean = 71 Hz, $sd = 5$).

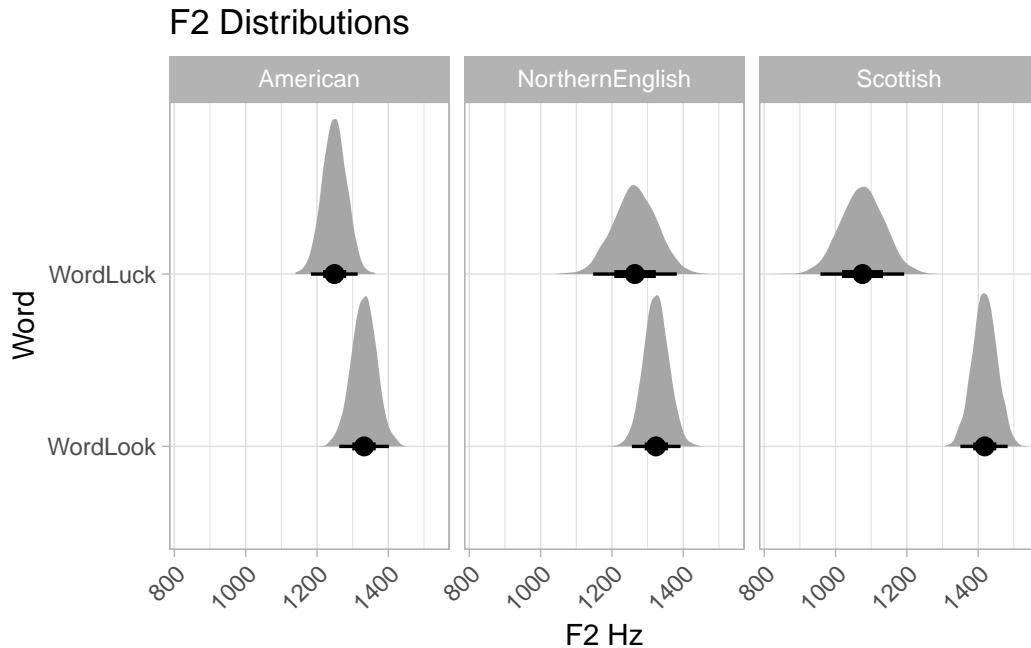


Figure 5: Gaussian Distributions of the F2 values within different accents/words

Figure 5 features a Bayesian Gaussian regression model for F2, following the same procedure described for F1 in Figure 1. The model included accent (American, Northern English, Scottish) and word (Look, Luck) as predictors, used default priors, and was run with four chains of 2000 iterations each (1000 warm-up) in brms (Bürkner, 2017) in R (R Core Team, 2025). The outcome variable was the normalized F2, with default treatment contrasts applied (American Look as the reference level).

According to the model, there is a 95% probability that the mean F2 for American speakers saying ‘Look’ is between 1226 and 1402 Hz (mean = 1332, $sd = 35$). For the word ‘Luck’ there is a 95% probability that the F2 of their vowel is between 176 Hz lower to 15 Hz higher (mean difference = -83, $sd = 49$) than them saying ‘Look.’

There is a 95% probability that the F2 of Northern English speakers saying ‘Look’ is 103 Hz lower to 89 Hz higher than for American speakers (mean difference = 8, $sd = 49$), giving a

mean of 1340 Hz. The F2 for Northern Speakers' 'Luck', is between 116 Hz lower and 149 Hz higher than the American 'Luck' (mean = 16, $sd = 70$), which gives a mean of 1265 Hz. The difference between the mean F2 of 'Look' and 'Luck' for Northern English accents is 75 Hz. This small difference demonstrates the predicted merger in Northern English accents.

For Scottish speakers saying 'Look' there is a 95% probability that the difference from the American baseline is between -103 and 89 Hz (mean difference = 87, $sd = 49$), giving a mean of 1419 Hz. The F2 of their 'Luck' is between -38 to 310 Hz lower than the American 'Luck' (mean difference = -174, $sd = 69$), which gives a mean of 1075 Hz. The difference between the mean F2 of 'Look' and 'Luck' for Scottish accents is 344 Hz. Therefore the graph clearly demonstrates the FOOT-STRUT split for American and Scottish accents.

The residual standard deviation is between 62 and 81 Hz (mean = 71, $sd = 5$).

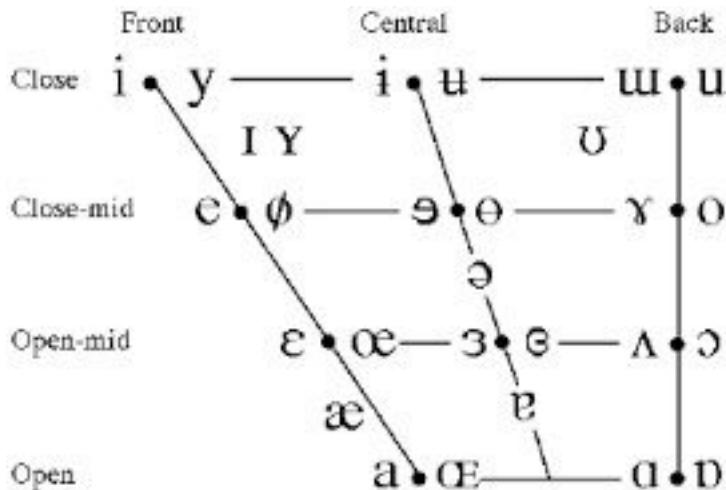


Figure 6: IPA vowel chart

Collectively Figure 4 and Figure 5 paint a much clearer picture of the vowels employed in our test. The differing F1 values for the American and Scottish productions of Look and Luck, respectively, illustrate different vowel heights for the LOOK and LUCK vowels in each community; whereas the Northern English production has similar F1 values, making the vowel sounds much closer in height than their accent counterparts. When analysing the F2 distributional graphs, the Scottish participants showed a much wider distribution of F2 values, with greater F2 values for the LOOK – suggesting a more central vowel. Whereas the Northern English and American participants sit in similar ranges, suggesting an all-around back vowel. These findings are consistent with the expected vowel productions outlined in the background. (Honeybone et al., 2008) (Smith et al., 2024, p. 157) (Turton et al., 2021, p. 163-201).

See Figure 6, which labels the vowel found in the International Phonetic Alphabet of English (Wieling et al., 2011). The American vowels /ʊ/ (FOOT) and /ʌ/ (STRUT) appear in places that are consistent with their identified F1 and F2 values - /ʊ/ is high and back, while /ʌ/ is mid

and back. The Northern English vowel /ʊ/ (FOOT and STRUT) is a high-back vowel, hence the consistent and similar placements for LOOK and LUCK in both F1 and F2 values. The Scottish vowels /u:/ (FOOT) and /ʌ/ (STRUT) also follow F1 and F2 expectations, with /u:/ being in a F1 range, similar to that of /ʊ/ (as seen in the American data) and F2 being considerably higher (more frontal) than the American /ʌ/ - making it closest to the vowel /ʊ/.

Gaussian Mixture Models for different accents

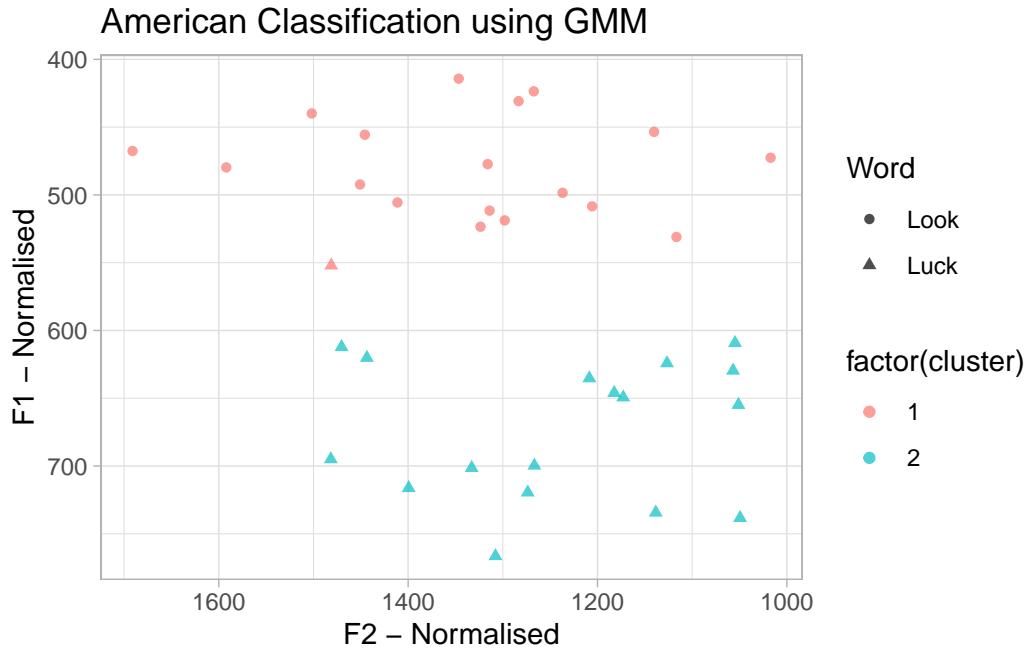


Figure 7: Gaussian Mixture Model Clustering on the American participants data.

The graphs in Figure 7, Figure 8 and Figure 9 are all classification models, using the Formant values of F1 and F2 to classify into a number of groups chosen by the BIC (Bayesian Information Criterion) to maximise between the models fit and simplicity. The American and Scottish graphs show that the GMM chooses two groups, and mostly correctly groups based on the word, demonstrating that there is a clear split in pronunciation. Contrastingly, the model for the Northern English formants classifies it all as one group, showing that there is a merger in the pronunciation for these words in the Northern English accent. Ultimately what this graph shows is a computational representation of the split rather than our interpretation. Thus a mathematical split was found.

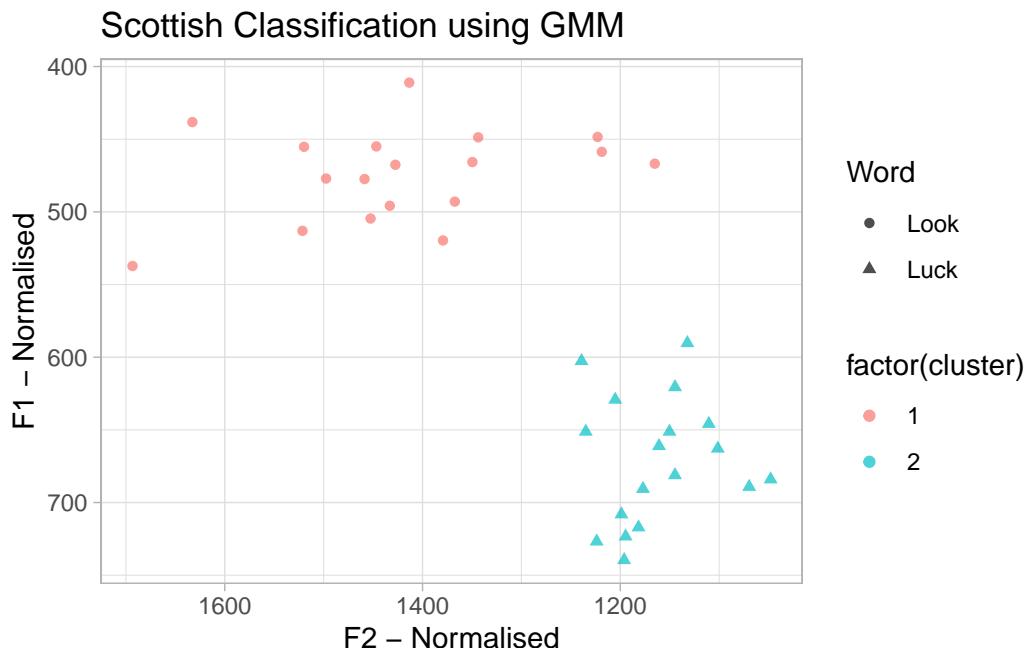


Figure 8: Gaussian Mixture Model Clustering on the Scottish participants data.

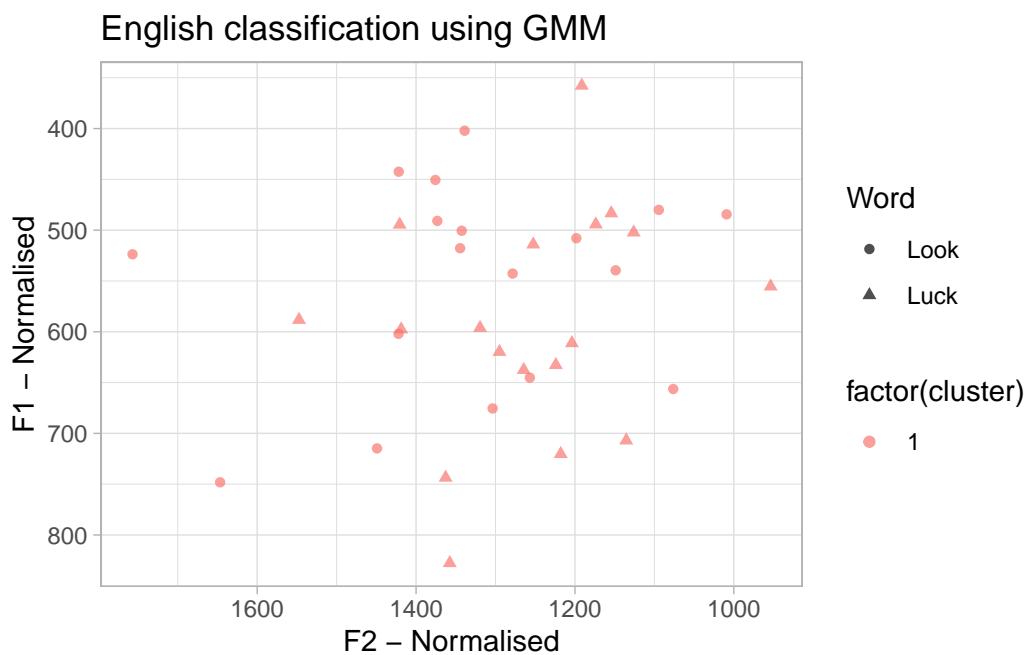


Figure 9: Gaussian Mixture Model Clustering on the English participants data.

Alternate Routes

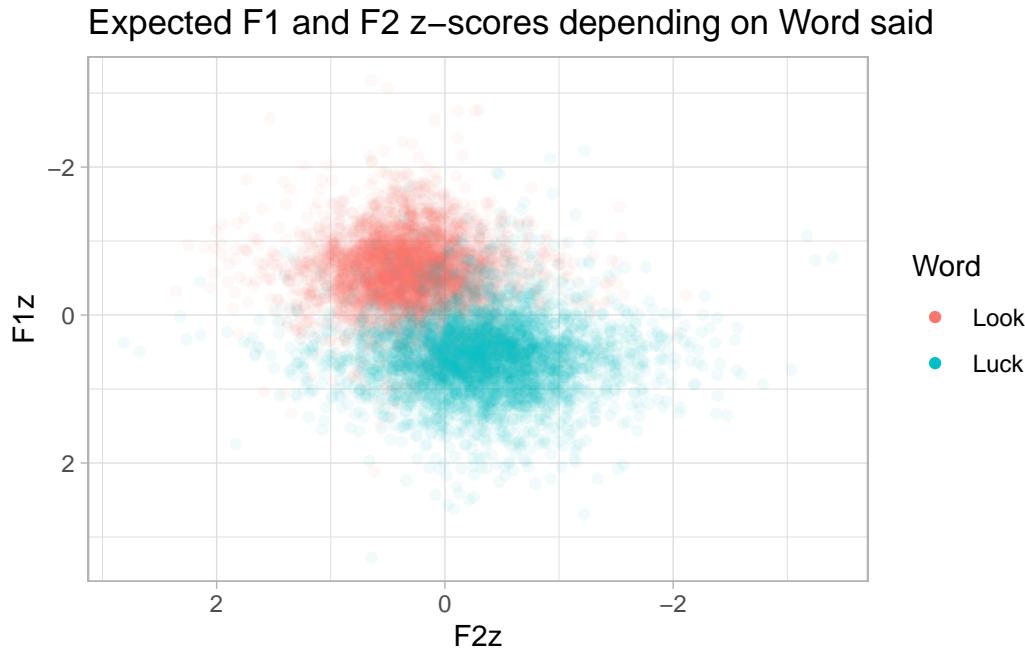


Figure 10: Multivariate multilevel regression model diagram

This graph in Figure 10 instead of using the formant features to predict class, instead uses the Word and the Accent to predict posterior distribution of the F1 and F2 values.

We fit a multivariate multilevel regression model using F1 and F2 z-scores as outcome variables. To keep things simple we just add which word is being said as a predictor and we include speaker-specific varying terms for intercept and word. Finally we then use `epred.draws` to give a series of predictive F1 and F2 z-score values for each word. Plotting those with a different colour for each word, we can see that overall there is a split between the two words pronunciation, however there is some overlap, this is likely because of the Northern English data.

Conclusion

This study investigated the FOOT-STRUT split over three different accents (General American, Northern English, and Scottish) with the tokens look and luck. The results suggest that the data fits in line with previous work that details a split between the vowel in American and Scottish accents, and a merger for Northern English. It was found that American and Scottish F1 were split while the F1 for Northern English participants remained overlapped. The F2 values were closer for American participants, while they remained distinctly separate for Scottish

participants. The study was limited by a general lack of participant numbers, but the findings uphold previously tested hypotheses about the split; that American and Scottish speakers of English have the split. It is of interest to note again that the two Northern English speakers, while a merge is present, use different vowels for FOOT-STRUT each, so that each speaker is not merging on the same vowel. Although this may be to the wall dataset, it may be warranted to look further into the quality of vowels own Northern English speakers to investigate whether different speakers are using different vowels.

Transparency

All our data and code is available on github via <https://github.com/jubbhubb/QML>. All of our raw code is in the file Coding.qmd, our data is available in csv format in file, data.csv And this final report is available in .qmd and .pdf format as Final-Report. All materials are under CC BY-NC © 2019. This work is openly licensed via CC BY-NC 4.0.

References

- Aitken, A. J. (1984). Scots and English in Scotland. Originally published in Peter Trudgill (Ed.), *Language in the British Isles* (Cambridge University Press), 517–532.
- Beal, J. C. (2012). ‘By Those Provincials Mispronounced’: The strut Vowel in Eighteenth-Century Pronouncing Dictionaries. *Language & History*, 55(1), 5–17. <https://doi.org/10.1179/1759753612Z.0000000001>
- Bürkner, Paul-Christian. 2017. “Brms: An r Package for Bayesian Multilevel Models Using Stan.” *Journal of Statistical Software* 80 (1): 128. <https://doi.org/10.18637/jss.v080.i01>
- Honeybone, P., & Salmons, J. (2008). *Handbook of English Phonology*. Wiley-Blackwell.
- Hughes, A., Trudgill, P., & Watt, D. (2005). *English accents and dialects: An introduction to social and regional varieties of English in the British Isles* (4th ed.). London: Hodder Arnold.
- Nasution, A., Syarfina, T. (2023) Highest and Lowest Pitch and Vowel Formants Measurement by Using PRAAT Application. *Journal of English Language and Education*. pp. 196 <https://doi.org/10.31004/jele.v8i2.445>
- R Core Team. 2025. R: A Language and Environment for Statistical Computing [Version 4.5.0].
- Schreier, D. ‘Historical Phonology and Koinéization’, in P. Honeybone & J. Salmons (Eds.), *The Oxford Handbook of Historical Phonology* (2015; online edn, Oxford Academic, 16 Dec. 2013), <https://doi.org/10.1093/oxfordhb/9780199232819.013.037>
- Smith, J., Stuart-Smith, J., Macdonald, R., & Jamieson, E. (2024). Scots and Scottish Standard English. In S. Fox (Ed.), *Language in Britain and Ireland* (pp. 151–177). Cambridge: Cambridge University Press.

Turton, D., & Baranowski, M. (2021). Not quite the same: The social stratification and phonetic conditioning of the foot–strut vowels in Manchester. *Journal of Linguistics*, 57(1), pp. 163–201.
<https://doi.org/10.1017/S002226720000122>

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Wieling, M., Margaretha, E., Nerbonne, J. (2011) Inducing Phonetic Distances from Dialect Variation. *Computational Linguistics in the Netherlands Journal*. Figure 3.