NURSE Vowels in Scottish Standard English: Still Distinct or Merged?

Zeyu Li¹, Ulrike Gut¹, and Ole Schützler²

Abstract
While nearly all dialects on the British Isles have undergone the nurse merger, a process which merged the Middle English vowels /ɪ ɛ ʊ/ into the vowel /ə/ (which was later lengthened to /ɜː/) in pre-rhotic positions, Scottish Standard English (SSE) is traditionally described as having retained a three-way distinction in these contexts. However, the gradual loss of this contrast has been observed in some varieties of Scottish English. This study investigates phonetic realizations within the nurse lexical set in SSE speech. 1227 tokens of the nurse vowel produced by ninety-two speakers were drawn from broadcast news, broadcast talks, legal presentations, non-broadcast talks, and unscripted speeches from the Scottish component of the International Corpus of English (ICE Scotland). The first two formants (F1 and F2) were measured, transformed into Bark and normalized. A Bayesian linear mixed-effects regression model showed that in purely acoustic terms, the vowels in fir, fern, and fur are not merged and have a distinct F1 and F2. However, the pre-rhotic items are distinct from the reference categories kit, dress, and strut in being more centralized, and in some genres fir and fern are more strongly drawn towards the center of the vowel space (and each other) than fur is. While the social variables age and gender do not influence realizations of the nurse vowels in formal Scottish English at this general level, orthography and the realization of the following /r/ have a clear effect. Inspection of individual speakers further shows that several types of partial merger of these vowels exist; it is argued that this perspective is needed to understand variation within the SSE nurse lexical set.

Keywords
Scottish Standard English, nurse vowel, vowel merger, orthography, /r/ variant

¹University of Münster, Germany
²University of Leipzig, Germany

Corresponding Author:
Zeyu Li, University of Münster, Johannisstraße 12-20, Münster, 48143, Germany.
Email: zeyu.li@uni-muenster.de
1. Introduction

Nearly all dialects in the British Isles have undergone the nurse merger (Wells 1982:407), a process which merged the Middle English vowels /ɪ ɛ ʊ/ before syllable coda /r/ into the vowel /ə/, which was later lengthened to /ɜː/ (Cruttenden 2008:131). The only modern accents of English that have retained three separate vowels in this context are some varieties of Irish English and Scottish English.¹ This means that, in many forms of Scottish English, for words such as *fir, fern, and fur*, the vowels are phonetically distinct, namely [i], [ɛ], and [ʌ]. This three-way distinction was also characteristic of most traditional Scots dialects and has been maintained in some varieties of Scottish Standard English, as well as in non-standard varieties of Scottish English (e.g., Grant 1913:50-51, 54-56, 62-63; Jones 2002:27; Dyer 2002:103). However, scholars have claimed that this three-way distinction has been lost in some varieties of Scottish English, especially in middle-class speech in the so-called “Central Belt,” sometimes with a particular emphasis on Edinburgh accents (Wells 1982:407; Lawson, Scobbie & Stuart-Smith 2013). Possible reasons for this potential merger of /ɪr/, /ɛr/, and /ʌr/ into /ɜːr/ are the promotion of Received Pronunciation (RP) phonology as a prestigious pronunciation target by middle-class speakers (probably inspired by Southern British accents) in early twentieth century Scotland (McAllister 1938:179), or a natural coarticulation effect resulting from the change from traditional trilled or tapped realizations of /r/ to the incoming standard realization as a bunched approximant [ɹ] (Aitken 1979:111; Lawson, Scobbie & Stuart-Smith 2013:208). At the same time, there is a dearth of empirical studies on the current status of variation in the standard variety of Scottish English and possible factors conditioning the underlying dynamics. In order to provide empirical evidence for the nurse merger in Scottish Standard English (SSE), the present study explores the realization of the nurse lexical set produced by SSE speakers and potential linguistic and social factors influencing the acoustic properties of the vowel by drawing on the International Corpus of English (ICE) Scotland (Schützler, Gut & Fuchs 2017), a large-scale phonologically annotated corpus on SSE.

The article is organized as follows. Section 2 provides a review of the literature describing the pronunciation of the nurse vowels in Scottish English and empirical evidence for a nurse merger. The aim and research questions of this study are outlined in full at the end of this section. Section 3 presents the method and data used. This includes an overview of the corpus data (section 3.1), the extraction and acoustic analysis of the nurse vowels (section 3.2), and the auditory analysis of the following /r/ variant (section 3.3). In section 4, we present our findings on the acoustic properties of the nurse lexical set and their interplay with various linguistic and social variables (section 4.1). Four distinct realization types of nurse merger are discovered upon inspecting patterns produced by individual speakers (section 4.2). Section 5 draws the analyses together to discuss the current status of a nurse merger in formal Scottish English and gives an outlook on future research.

2. Nurse Merger in Scottish English

While there is agreement that /ɪr/, /ɛr/, and /ʌr/ vary in Scottish English speech between their traditional distinct realizations and vowels approaching a full merger upon the
lexical set defined by Wells (1982) as nurse, less is known about the mechanisms underlying this variation. For example, Wells (1982:407) describes the three-way pre-rhotic vowel distinction in prestigious Scottish accents as typical only of speakers in the west of Scotland. On the other hand, he describes Edinburgh middle-class speakers as having a complete merger of the vowels in the nurse lexical set, while working-class Clydeside speakers are claimed to have a partial merger, producing words like fir and fur with the vowel /ʌ/, while retaining words in the fern sub-class as a distinct category. Along the same lines, Giegerich (1992:63) distinguishes Scottish English speakers with a complete and partial merger of the vowels in the nurse lexical set from those that produce three distinct vowels in this context, but does not provide details about possible regional or social effects. Likewise, Stuart-Smith (2008:57-58) describes variation across speakers of Scottish Standard English: she suggests that some have either /ɪ/ or /ʌ/ for the entire nurse lexical set, others make a distinction between /e/ and /ʌ/, and a third group, Scottish Standard speakers, has the three-way distinction of /ɪ/, /ɛ/, and /ʌ/.

However, despite such descriptions in the literature, empirical evidence for the observed variation or nurse merger in Scottish English is rare. Lawson, Scobbie, and Stuart-Smith (2013) analyze word lists containing a total of 223 tokens of the nurse lexical set read by eight working class and seven middle-class adolescents aged 12-13 years. Comparing the acoustic properties of the vowels in these words with the acoustic properties of the vowels /i/, /ɪ/, /e/, /ɛ/, /ɔ/, /o/, and /u/ produced in non-pre-rhotic position, they find that the middle-class adolescents show a greater centralizing tendency for /ɪ/, /ɛ/, and /ʌ/ in pre-rhotic position than the working-class adolescents do. However, an actual merger of the vowels of the nurse lexical set is observed for neither group. They explain the centralizing tendency of the middle-class adolescents as coarticulatory pressure exerted by the speakers’ bunched variants of the postvocalic /r/, which contrasts with a tongue front-raised realization of /r/ produced by most of the working-class adolescents. Indeed, ultrasound images of the tongue position for the articulation of pre-rhotic /ɪ/, /ɛ/, and /ʌ/ show that they are similar in location and configuration for the middle-class adolescents, suggesting that these speakers neutralize the distinction between the three vowels in pre-rhotic context by anticipating the articulatory tongue setting of the following /r/.

Apart from the realization of the following syllable coda /r/, two more factors emerge in Lawson, Scobbie, and Stuart-Smith’s (2013) study as potential influencing factors of the realization of nurse vowels. The first is orthography, i.e., the vowel that appears in the spelling of the words the adolescents read out. For about half of the working-class participants, <i> in words such as fir is variably pronounced as /ʌ/ or as /ɪ/. Three speakers produce /ɪ/ when first reading the word fir but produce /ʌ/ at the second reading. It is thus possible that the speakers’ first rendition of the vowel is influenced by the orthography of the word, which is then “overridden” when the word is repeated. A second factor suggested by Lawson, Scobbie, and Stuart-Smith (2013) is the use of a careful speech style, elicited by asking the adolescents to read out word lists. They propose that these working-class speakers have an underlying phonological representation of separate vowels for the nurse lexical set which might be accessed in formal speaking contexts.
In sum, very little empirical evidence substantiates claims of either merged or separate vowels for words such as *fir, fern, and fur* in Scottish English: the *nurse* merger and social stratification in the realization of pre-rhotic vowels in Scottish English are understudied. The present investigation aims to begin to fill this gap with a large-scale analysis of SSE that examines both the realization of the individual vowels in the *nurse* set as well as the factors that potentially have an influence on it. The following research questions will be addressed:

I. What are the acoustic properties of the *nurse* lexical set in SSE? Are different vowels produced or are they merged?
II. What effect do phonological (realization of following /r/), orthographic, and social factors (here, gender and age) have on realizations of the *nurse* vowels?
III. How different or similar are the patterns produced by individual speakers, and with what frequency do different types of *nurse*-merger occur?

### 3. Data and Method

#### 3.1. Corpus Data

The data for this study were drawn from ICE-Scotland, the Scottish component of the International Corpus of English (Schützler, Gut & Fuchs 2017). Corpora from the ICE-family follow a uniform sampling scheme (see Nelson 1996; Nelson, Wallis & Aarts 2002), with 600,000 words collected from fifteen spoken genres (or text types) and 400,000 words sampled from seventeen written genres.

In our analyses we include material from five spoken genres: four categories of scripted speech, including broadcast news (“bnew”), broadcast talks (“btal”; e.g., scripted speeches given in the Scottish parliament by invited speakers), legal presentations (“leg”; e.g., legal sentencing) and non-broadcast talks (“nbtal”; e.g., scripted speeches given at university such as inaugural lectures on any topic), as well as unscripted speeches (“unsp”; e.g., conference presentations in any field). If the general design of ICE is to lean intentionally towards the standard end of language use by including only highly educated speakers (e.g., comments in Greenbaum 1996:177), these genres could be said to be particularly formal and most likely to evoke the use of the standard variety of Scottish English, as compared to face-to-face conversations or telephone calls, for example. All speakers in this study are users of SSE, including university lecturers, journalists, lawyers, university students, and ministers. It should be noted that, although they are considered to be middle-class speakers due to their current occupations, their original social class is largely unknown. This of course has consequences for the generalizability of our data, as we will discuss in the final section.

The total number of speakers is ninety-two; the breakdown by gender and genre is shown in the left-hand part of Table 1 (dataset A). For most of the analyses, a smaller dataset was derived from this dataset A, one that includes only those speakers for
whom information regarding both gender and age was available (dataset B). Further, a third subset of the data was used to inspect the relationship between the acoustic quality of the vowel and the phonetic realization of the following /r/ (dataset C). Since this analysis was based on a time-consuming auditory analysis, a smaller dataset was chosen (see section 3.3). Rhoticity is a highly sensitive feature that is subject to a series of language-internal and external factors; dataset C is based solely on broadcast talks rather than a cross-section of genres to rule out potential effects of speech rate, text type, and orthography on the realization of /r/. As Table 1 shows, the genre of broadcast talks is strongly overrepresented in our sample in terms of speaker numbers, particularly so in dataset B. Moreover, in all datasets the gender distribution within most genres is uneven.

### Table 1. Number of Speakers per Genre

| Genre                  | Dataset A |                  | Dataset B |                  | Dataset C |                  |
|------------------------|-----------|------------------|-----------|------------------|-----------|------------------|
|                        | All speakers | Speakers with known gender and age | Broadcast speakers |
|                        | Man | Woman | Total | Man | Woman | Total | Man | Woman | Total |
| Broadcast news         | 10  | 4     | 14    | 2   | 0     | 2     | 11  | 19    | 30    |
| Broadcast talks        | 19  | 26    | 45    | 16  | 23    | 39    |      |       |       |
| Legal presentations    | 8   | 3     | 11    | 7   | 3     | 10    |      |       |       |
| Non-broadcast talks    | 3   | 3     | 6     | 2   | 2     | 4     |      |       |       |
| Unscripted speeches    | 9   | 7     | 16    | 3   | 7     | 10    |      |       |       |
| Total                  | 49  | 43    | 92    | 30  | 35    | 65    | 11  | 19    | 30    |

The total number of target words containing the vowels we consider here was 64,616 in dataset A, 51,576 in dataset B, and 23,640 in dataset C. We include word as a predictor in our statistical analyses.

### 3.2. Extraction and Acoustic Analysis of Vowels

ICE Scotland contains automatic phonemic transcriptions that were created with WebMAUS (Schiel 2004). This is a web service that creates phonemic transcriptions via forced alignment of orthographic transcriptions and sound files. These transcriptions were subsequently corrected manually for both phoneme transcription and location of phoneme boundaries in Praat (Boersma & Weenink 2017). For the present study, we created TextGrids that were based on the automatically generated annotations and contained twelve vowel categories in the stressed position (three nurse vowels and nine additional reference vowels). The first three categories target the three pre-rhotic short vowels with the potential for nurse-merger: /ɪr/ as in bird, /ɛr/ as in earth, and /ʌr/ as in nurse.² We include the following (phonological) rhotic in the label to avoid confusion with the corresponding non-pre-rhotic vowels. The nine additional vowels were mea-
sure to provide context and points of reference within the vowel plots: /i/ as in fleece, /u/ as in goose, /e/ as in face, /o/ as in goat, /u/ as in kit, /a/ as in bath, /ʌ/ as in strut, /ɔ/ as in lot, and /ɛ/ as in dress. Moreover, the first eight of these—/i u e o ɪ a ʌ ɔ/—were used as normalizer vowels (we describe the procedure below). Since identification of stressed environments was based on manual inspection of the transcripts, analyses of longer recordings were terminated around four minutes into the file for the additional vowels; for the three targeted contexts /ɪ r ɛ r/ all instances were analyzed. The total number of analyzed tokens is documented in Table 2.

Table 2. Total Number of Vowels Extracted

| Vowel | N  | Vowel | N  | Vowel | N  | Vowel | N  |
|-------|----|-------|----|-------|----|-------|----|
| ɪr    | 255| i     | 2061| o     | 1341| a     | 2236|
| ɛr    | 491| u     | 1778| ɪ     | 4047| ʌ     | 1743|
| ʌr    | 481| e     | 2316| ɛ     | 3212| ɔ     | 2608|

Using Praat, the F1 and F2 values of all target vowels were extracted automatically in Bark (Traunmüller 1990), averaging across a section of the vowel that extended from 20-70 percent of its total duration. Based on Sönning and Schützler (2018), we then applied an adapted form of Lobanov’s (1971) z-score normalization procedure. Based on the raw data, we first determined the midpoint of the acoustic vowel space for each speaker by calculating median values of F1 and F2 for each of the eight normalizer categories for that speaker, and by calculating their mean, again separately for each formant. Next, the respective mean was subtracted from each individual vowel measurement, creating values centered around zero. For all vowels, these negative and positive values of F1 and F2—effectively deviations (in Bark) from the midpoint of the space—were then divided by the standard deviation of the sixteen centered median values of normalizers for the respective speaker across both F1 and F2. While the original Lobanov normalization procedure measures formants in Hz and divides centered values by a standard deviation of normalizers that is different for F1 and F2, using a single standard deviation in both dimensions preserves relative positions and distances between vowels and thus respects the psychoacoustic scaling introduced by the Bark-scale.

3.3. Auditory Analysis of the /r/ Variant

For the thirty speakers in dataset C (see Table 1), an auditory analysis was carried out to determine the realization of the /r/ variant (if present) in all /ɪr/, /ɛr/, and /ʌr/ words. Two phonetically trained raters assessed whether /r/ was realized as a tap/trill ([ɾ r]), an approximant ([ɹ V]), or whether no constriction was perceived (coded as Ø). These three categories were employed in order to capture variation along a continuum ranging from the “traditional” Scottish tap/trill realization via the modern standard realization as an approximant to a vocalized, unconstricted variant (following Schützler 2013, 2015:59). Unlike /r/-vocalization in working-class speech, the latter is interpreted as possibly due to anglicization in middle-class speech.
We refrained from using a more fine-grained distinction of rhotic realizations because it would have made it difficult to achieve sufficient interrater reliability (as in Dickson & Hall-Lew 2017). Following this method, we achieved good interrater reliability of about 89 percent between the first and the second rater for identification of /r/ variants. In case of disagreement between the two raters, a third rater, also phonetically trained, made the final decision. The auditory analysis was complemented with a visual spectrographic analysis using Praat (Boersma & Weenink 2017) for less clear tokens or when in doubt.

### 3.4. Statistical Modeling

For the statistical analysis, Bayesian linear (Gaussian) mixed-effects regression models with speaker and word as random intercepts were used (for further information about this type of analysis, see, e.g., Jackman 2009; Kruschke 2015; McElreath 2016). Models were fitted using the R-package `{brms}` (Bürkner 2019), which in turn uses the software Stan (Stan Development Team 2018). Table 3 lists all variables that were used.

| Table 3. Variables in the Statistical Models |
|---------------------------------------------|
| **Class** | **Variable** | **Values/levels** | **Comment** |
| Outcome | Z | Median = 0.28; 95% CI[−1.70, 1.05] | |
| Predictor | CATEGORY | /r/, /ə/ | reference: /r/ |
| | FORMANT | F1, F2 | reference: F1 |
| | GENDER | −1 (‘man’), 1 (‘woman’) | |
| | AGE | Median = 12; range[−23, 35] | centered on age = 40 |
| | GENRE | btal, bnew, leg, nbtal, unsp | reference: btal |
| | SPELLING | <ir>, <er>, <ear>, <ur>, <or> | reference: <ir> |
| | RHOTICITY | [ɾ] [ɹ] [ɾ]Ø | reference: [ɾ V'] |
| Grouping/random | SPEAKER | Model A: N = 92 levels | |
| | | Model B: N = 65 levels | |
| | | Model C: N = 30 levels | |
| | WORD | Model A: N = 271 levels | |
| | | Model B: N = 220 levels | |
| | | Model C: N = 100 levels | |

As summarized in Table 4, three models were fitted to datasets A, B, and C; they were analogously labeled “Model A,” “Model B,” and “Model C.” Model A and Model B differ only in the latter having AGE as an additional predictor and including fewer speakers. All three models were run with 5000 iterations (warmup = 1000) and four chains, resulting in 16,000 posterior samples: Bayesian regression models establish probability distributions for the specified (e.g., predictor) coefficients and then sample from, or iterate across, these distributions. This procedure yields a large sample of likely values but also a number of less likely values, which are then transformed into an equally large number of actual predicted values for the outcome. The distribution of predicted values can then be shown as an average value with uncertainties, as in our plots. The sampling process is partitioned into parallel chains (in our case four chains...
of 4000 valid samples each), and the R-hat diagnostic checks if chains converge, i.e., if they are in reasonable agreement. R-hat values of 1.0 indicate that the specified model is able to “cope with” the data, i.e., the individual chains come to similar conclusions concerning parameter values—they converge. In cases of overcomplex (or ill-specified) models or when data are extremely sparse, this would not be the case, as each chain would be “groping in the dark” and come to an erratic solution.

Table 4. Model Syntax for Linear Mixed-effects Regression Models

| Model A                                    | $Z \sim ((\text{gender} + \text{genre}) \ast \text{category} + \text{spel}) \ast \text{formant}$ |
|--------------------------------------------|----------------------------------------------------------------------------------|
|                                            | + (\text{category}\ast\text{formant} \mid \text{speaker})                       |
|                                            | + (\text{formant} \mid \text{word})                                             |
| Model B                                    | $Z \sim ((\text{gender} \ast \text{age} + \text{genre}) \ast \text{category} + \text{spel}) \ast \text{formant}$ |
|                                            | + (\text{category} \ast \text{formant} \mid \text{speaker})                     |
|                                            | + (\text{formant} \mid \text{word})                                             |
| Model C                                    | $Z \sim \text{category} \ast \text{formant} \ast \text{gender} \ast \text{rhoticity}$ |
|                                            | + (\text{category} \ast \text{formant} \mid \text{speaker})                     |
|                                            | + (\text{formant} \mid \text{word})                                             |

The outcome variable in all models is the formant value $Z$, which is a z-score calculated as described in the previous section. In the fixed part of Model A, the variables \text{category} and \text{formant} necessarily interact not only with each other but also with the two socio-stylistic predictors, \text{gender} and \text{genre}: prediction only makes sense if it can be applied to a specific vowel and a specific formant. \text{gender} and \text{genre} are specified as not interacting with each other, because we have little reason to expect that the effect of our five genres will differ systematically between genders. In Model B, the additional predictor \text{age} interacts with \text{gender}, so that it is possible, for instance, to assess the special roles played by young women and men, rather than young speakers in general—a perspective taken in many classic sociolinguistic studies. In both models, the predictor spelling only interacts with the predictor F2: an interaction with \text{category} would entail perfect collinearities (e.g., <ur> spellings must always belong in the category /ʌr/), and we have no reason to expect that a speaker’s age or gender should have an effect on the impact of this factor; we thus treat spelling as a truly language-internal factor. In Model C, full interactions for all four fixed predictors are specified.

Concerning the random part of the models, two nesting structures were specified: language-external nesting by speaker and language-internal nesting by lexeme captured by the variable word. The slope of F2 varies randomly across levels of word, while the interaction category\ast F2, varies randomly across speaker. Full regression models for the fixed effects can be found in the Appendix.

4. Results

4.1. Realization of nurse Vowels

In the following discussion of the results, predictors that are not of interest in a particular display are constrained to take their normal/average values. For example, if there is
no focus on genre differences, predictor values of the four predicted categories of GENRE were set to 0.2 to simulate a balanced scenario in which all five genres are represented equally. For SPELLING, those controlled predictor values were set to reflect the relative frequency of spellings in our raw data: of all /ɛr/ cases, 76 percent were <er> (e.g., service) and 24 percent were <ear> (e.g., early); of all /ʌr/ cases, 31 percent were <ur> (e.g., turn) and 69 percent were <or> (e.g., word). Accordingly, predictor values for the respective categories were set to 0.76, 0.24, 0.31, and 0.69 to generate more realistic estimates that take into account the actual frequencies of different spellings in our sample. For GENDER, the controlled predictor value was zero (“between man and woman”). Finally, scenarios in which age differences do not play a role used a predictor value of AGE = 0, since this, too, is a centered variable, with a predictor value of zero corresponding to an assumed “middle” age of 40 years.

Figure 1 displays the estimates of the speaker means of F1 and F2 for /ɪr/, /ɛr/, and /ʌr/, based on dataset A. The thicker, central parts of the crosses placed on each vowel position represent 50 percent uncertainty intervals, while their thinner and longer parts represent 95 percent uncertainty intervals. The high threshold of confidence was chosen because of the large number of speakers (N = 92) and because GENRE was modeled as a fixed effect. Figure 1 also shows the other vowels of the speakers’ system that were measured. The resulting vowel system agrees very well with descriptions of the Scottish vowel system elsewhere (e.g., Abercrombie 1979; Giegerich 1992; Stuart-Smith 2008; Schützler 2015).

As can be seen from Figure 1, in SSE the nurse vowels are not acoustically merged, at least not from this general, averaging perspective: /ɪr/, /ɛr/, and /ʌr/ are clearly distinct. However, it can also be seen that in comparison with /ɪɛʌ/ in non-pre-rhotic contexts, /ɪr/ and /ɛr/ centralize strongly. /ɪr/ is most clearly mid-central and differs markedly in both F1 and F2 from its reference vowel /ɪ/. While we focus on the question of whether or not vowels have merged upon a reduced number of acoustic qualities, another relevant aspect is the question of how far the pre-rhotic cases are removed.

**Figure 1.** Position of /ɪr/, /ɛr/, and /ʌr/ in the Acoustic Vowel Space (Dataset A)
from their assumed non-rhotic reference values. For example, we can see that the vowel in \(/\text{ɪr}/\) has to be relatively far away from its non-pre-rhotic counterpart to reach the central area in the vowel space in which a merger can take place. This is much less the case for the other two vowels, in particular pre-rhotic \(/\text{ʌr}/\), and one could thus argue that the contribution of \(/\text{ɪr}/\) to a perceived merger is greater than the contributions of the other two categories.

While Figure 1 presents the mean F1 and F2 values of \(/\text{ɪr}/\), \(/\text{ɛr}/\), and \(/\text{ʌr}/\) averaging across all speakers in dataset A on the basis of the statistical model, Figure 2 zooms into the realizations of \(/\text{ɪr}/\), \(/\text{ɛr}/\), and \(/\text{ʌr}/\) by the individual speakers. Token numbers differ between the three panels of the plot because we excluded speakers for whom the respective vowel category was represented by less than two tokens (thus, \(N = 57\) for \(/\text{ɪr}/\), \(N = 73\) for \(/\text{ɛr}/\), and \(N = 70\) for \(/\text{ʌr}/\)). Individual speakers’ average positions for the three categories form “clouds” in the vowel space that are centered roughly on the average positions shown in Figure 1. However, considerable inter-speaker variation can be observed for all three nurse vowels. We can therefore conclude that the high degree of certainty for the average positions of the three vowels in Figure 1 above is due to the large number of speakers; it does not mean that all speakers produce vowels that are actually close to those averages. Together with the finding that the areas occupied by the three vowel categories overlap considerably, this suggests that there may in fact be mergers, or partial mergers, at the level of individual speakers. Concrete patterns of this kind are neither visible in Figure 1 nor in Figure 2, but we return to this issue below, where we argue that the speaker-based (non-averaging) perspective is crucial for a better understanding of the phenomenon at hand.

![Figure 2. Positions of \(/\text{ɪr}/\), \(/\text{ɛr}/\), and \(/\text{ʌr}/\) by Individual Speaker (Dataset A)](image)

Figure 3 presents estimated mean transformed F1 and F2 values for the different gender and age groups, based on dataset B: women are shown in the top panel (using 50 percent and 90 percent uncertainty intervals) and men are shown in the bottom panel. For the respective older groups, estimates are based on the assumption of sixty-year-old speakers; for younger groups, speakers are assumed to be twenty years old.

The figure shows that there are no substantial differences between groups established on the basis of age and gender. However, women produce realizations of \(/\text{ɪr}/\) and \(/\text{ɛr}/\) that are acoustically somewhat closer to the positions of the non-rhotic reference
vowels. If anything, this means that the men are the ones who tend to centralize those vowels into the space where merger is possible, but this is only a relatively weak tendency. It is noteworthy that the men in the sample generally produce /ʌr/ with a lower F2, in an area of the vowel space that is acoustically closer to /o/ and /ɔ/ than to /ʌ/. This may in fact suggest that there is a tendency among those speakers to avoid the incipient merger of /ɪr/ and /ʌr/: lowered and retracted realizations of /ɪr/ occupy the mid-central area of the acoustic space, and /ʌr/ moves out of the way and takes an even further retracted position, not unlike push-chain effects in sound changes such as the Great Vowel Shift (e.g., Krug 2012). As a result, this age-group related data appears to show that /ʌr/ may be becoming more distinct (more retracted in acoustic space) from /ɪr/ and /ɛr/ in the speech of younger men only.

Concerning age patterns, results are inconclusive. For both men and women, differences between age groups are moderate with regard to most pre-rhotic vowels. If we assume an ongoing change in the direction of centralization and/or merger, this is only partially supported, e.g., by the tendency of younger men to centralize the vowel in /ɪr/ more than older men, or a similar (albeit very slight) tendency of the same kind that affects /ʌr/ between younger and older women. However, these tendencies appear to be sporadic and are not part of a more general pattern. For example, it is the younger women and the younger men who produce acoustically more peripheral variants of /ɛr/, and /ʌr/ is realized further towards the back with a lower F2 by younger men compared to their older counterparts, while younger women have a slightly increasing

![Figure 3. Position of /ɪr/, /ɛr/, and /ʌr/ by Age and Gender (Dataset B)](image-url)
F2 than older women do. In other cases, there is no correlation between age and the centralization of vowels, for example concerning the realization of /ɪr/ by women.

Again, the considerable scatter of individual speakers’ vowel positions for all three nurse categories in Figure 2 can serve as a warning not to overinterpret the weak tendencies observed in Figure 3 and other visualizations of this kind. If we assume that individual accents are characterized by different types of (partial) merger, it would come as no surprise if averaging across such patterns for groups of speakers produces relatively indiscriminate patterns.

Figure 4 presents differences in vowel realization across the four genres broadcast talks (btal), non-broadcast talks (nbtal), legal presentations (leg), and unscripted speeches (unsp), again showing 50 percent and 90 percent uncertainty intervals. The text category broadcast news (bnew) was not plotted since data are rather sparse and included no tokens from women (see dataset B in Table 1). As discussed above, predictors that are not in focus are constrained to take their “normal” values (age, gender, spelling).

Figure 4 shows that in the category non-broadcast talk, arguably the least formal of the categories analyzed here, /ɪr/, /ɛr/, and /ʌr/ are most similar, with /ɪr/ and /ɛr/ close to merging and /ʌr/ being further apart. No difference between scripted and unscripted speech can be seen. The same caveat as above holds with regard to this analysis as well: if we are interested in the average positions of the three vowels for groups of speakers (in this case arranged by the genres in which language was produced), we arrive at results that are only subtly different and not always easy to interpret. The approach followed thus far produces idealized abstractions based on the potentially more radically different and therefore more revealing underlying individual patterns.
Figure 5 shows the realizations of /ɪr/, /ɛr/, and /ʌr/ according to the orthography of the word. In contrast to all previous analyses, the predictor speling was not held constant, but specific spellings were targeted when estimating formant values (see discussions above).

The figure shows that while the realization of /ɛr/ is not influenced by the spelling of either <er> or <ear>, spelling has a marked effect on the realization of /ʌr/: the vowel in words spelled with <or> such as in world is closer to /o/ and /ɔ/, while the vowel in words spelled with <ur> such as in burn is much more central and thus closer to the strut vowel /ʌ/. It is important to note that all the words spelled with <or> in the sample are preceded by /w/, while the <ur> words are not.

Figure 6 presents the vowel realizations, based on dataset C and model C (broadcast talk only; see Tables 1 and 4) according to the realization of the following /r/. It offers two perspectives on the same data: the three panels at the top show differences within each of the three vowel categories, depending on which variant of /r/ follows. The three panels at the bottom show complete constellations of /ɪr/, /ɛr/, and /ʌr/ in combination with each of the three /r/-variants. There is a clear tendency for vowels followed by approximants to be grouped more closely together in the mid-central area of the acoustic vowel space than vowels followed by the traditional tap/trill and vocalized variants of /r/. In the top half of Figure 6, a following approximant always correlates with the most central vowel realization. Furthermore, the realization of the following /r/ as a tap/trill promotes the partial merger of /ɪr/ and /ʌr/, which, by contrast, are shown to be more distinct when followed by a vocalized /r/ variant. In fact, if we look at the effect of zero-variants of /r/, we see that the preceding vowel either has a quality intermediate between the ones associated with tapped or approximant /r/ (for /ɛr/ and /ʌr/), or it takes a different position altogether (for /ɪr/). In other words, there is no evidence in these data that the historical trajectory of change affecting /r/ ([ɾ]→[ɹ]→Ø), which may be assumed at least in middle-class Scottish accents,
correlates with a continuum of variation and change between less centralized and more centralized (and eventually merged) qualities of nurse vowels in a straightforward way. Finally, the variable realizations of coda /r/ display a more pronounced effect on realization of /ɪr/ than of /ɛr/ and /ʌr/. More specifically, realizations of /ɪr/ are more distinct when followed by the three different /r/ variants, while for /ɛr/ and /ʌr/ variants conditioned by the different types of /r/ are closer together, notwithstanding the centralizing effect of the approximant.

Figure 6. Relationships between Tap/Trill (Black), Approximant (Dark Gray), and Vocalized or Deleted (Light Gray) Variants of Coda /r/ and the Acoustics of /ɪr/, /ɛr/, and /ʌr/ (Dataset C)

4.2. Types of nurse Merger

As indicated in Figure 2, there is not only considerable inter-speaker variation in the position of each of the three categories /ɪr/, /ɛr/, and /ʌr/ but also substantial overlap of the areas occupied by each of them. On the basis of our results thus far, it therefore seems perfectly possible for individual speakers to have merged at least two of the three categories, which our earlier perspective on the data had not been able to reveal. Based on vowel plots for thirty-eight speakers in dataset A who produced each of the nurse vowels at least two times, we identified four different types of nurse merger. Figure 7 presents three representative examples of each type. The plots are based on the raw data.
Type 1: Unmerged. Three separate vowel spaces for /ɪr/, /ɛr/, and /ʌr/ are observed. The traditional three-way distinction of nurse vowels is maintained and there is no tendency for nurse merger.

Type 2: Merger of /ɪr/ and /ʌr/. The realizations of /ɪr/ and /ʌr/ are centralized and merged, while /ɛr/ is realized close to its non-pre-rhotic counterpart and retained as a distinct category.

Type 3: Merger of /ɪr/ and /ɛr/. The vowel spaces of /ɪr/ and /ɛr/ overlap, while /ʌr/ is further apart to the back and close to /ʌ/ and /ɔ/.

Type 4: Full merger. The realizations of /ɪr/, /ɛr/, and /ʌr/ are centralized and merged in terms of F1/F2 space.

Table 5 shows the classification of the speakers and information regarding their gender and age, if applicable. The individual vowel spaces for all thirty-eight speakers can be found in the Appendix. As Table 5 shows, full merger is the most frequent type (N = 13) among the speakers, while the unmerged type is the least frequent pattern (N = 7). Upon inspecting the speakers’ gender, the unmerged category seems more frequent among women, while the full merger type displays the opposite. However, any generalizations based on such a limited number of speakers should be treated with caution: all types occur across different age groups and genders. Patterned differences between groups of Scottish speakers may become more obvious when socio-economic background is considered. The social class of the speakers included in ICE Scotland is not available, however.
| Type 1 (Unmerged) | Type 2 (Merger of /ɪr/ and /ʌr/) | Type 3 (Merger of /ɪr/ and /ɛr/) | Type 4 (Full merger) |
|-------------------|---------------------------------|---------------------------------|---------------------|
| **Speaker**       | **Gender** | **Age** | **Speaker** | **Gender** | **Age** | **Speaker** | **Gender** | **Age** | **Speaker** | **Gender** | **Age** |
| unsp_08           | w         | 20-40   | unsp_25    | w         | 35      | leg_07    | m         | 63      | unsp_05    | w         | 47      |
| unsp_07           | w         | 35      | unsp_26    | w         | 50+     | btal_29   | w         | 50-60    | leg_16     | m         | 61      |
| leg_04            | w         | 60+     | btal_44    | m         | 65+     | btal_23   | m         | 52      | nbtal_04   | w         | 45-55   |
| leg_11            | w         | 60+     | btal_43    | m         | 45-60   | btal_24   | w         | 40-50    | nbtal_14   | m         | 70+     |
| btal_34           | w         | 63      | btal_40    | m         | 60+     | btal_04   | m         | 60-75    | leg_13     | m         | 63      |
| btal_22           | m         | 50      | btal_36    | m         | 47      | unsp_16   | m         | n/a      | leg_14     | m         | 61      |
| unsp_04           | m         | n/a     | btal_21    | m         | 63      | unsp_03   | m         | n/a      | btal_42     | w         | 50-55   |
|                   |           |         | btal_19    | w         | 50      | bnew_09   | w         | n/a      | btal_27     | w         | 73      |
|                   |           |         | btal_03    | w         | 60      | unsp_23   | m         | n/a      | btal_18     | w         | 48      |
|                   |           |         |           |           |         |          |           |          |          | btal_45     | m         | n/a     |
|                   |           |         |           |           |         |          |           |          |          | btal_47     | m         | n/a     |
|                   |           |         |           |           |         |          |           |          |          | btal_31     | m         | 40-50   |
|                   |           |         |           |           |         |          |           |          |          | btal_39     | m         | 51      |
5. Discussion and Conclusion

This study investigated the acoustic properties of the nurse lexical set in SSE with a view of determining whether separate vowels are produced or whether speakers merge them. The global perspective we took at the outset suggests that, while /ɪ/, /ɛ/, and /ʌ/ in pre-rhotic position are, on average, acoustically more centralized than /ɪ/, /ɛ/, and /ʌ/ are in non-pre-rhotic position, they remain acoustically distinct. Thus, in this approach there is no support for the variously proposed loss of the three-way distinction in the realization of the nurse vowels in this variety, although clear tendencies are visible. Our results thus confirm Lawson, Scobbie, and Stuart-Smith’s (2013) findings for adolescent Edinburgh speakers—the only other empirical data available so far—who do not display a complete nurse merger either. However, we found considerable inter-speaker variation of all three vowel categories. This led us to the inspection of individual speakers’ vowel spaces. From this complementary perspective, the following, rather different, picture emerged: four different merger types seem to be current among SSE speakers of different genders and age groups. While the merger of /ɪr/ and /ʌr/ is largely in line with descriptions in previous literature (Giegerich 1992; Stuart-Smith 2008), we also found patterns in which /ɪr/ and /ɛr/ were merged, while /ʌr/ was retained as a distinct category. This is an important contribution to our understanding of the realization and distribution of the nurse lexical set in SSE, as it reveals even more diversity than previously assumed.

The second research aim was to explore possible social and linguistic constraints of the realization of the nurse lexical set in SSE. Once again averaging across groups of speakers based on the output of our statistical model, and purely in terms of merger vs. non-merger of /ɪr/, /ɛr/, and /ʌr/, the results show relatively little effect of either speaker age or gender. However, there seems to be some tendency for men to have more space between the vowels in /ɪr/ and /ʌr/ in particular, and this trend is more pronounced in younger speakers. What these speakers display could be described as a partial return to a more traditional vowel system in which the three contexts under investigation increase once again in distinctness. This could be argued to coincide with unmerged qualities that also exist in working-class speech, but on the basis of our data it is not possible to explore this as a potential source of variation and change. Some evidence that would lead us to suspect more centralization and potentially merger of these vowels in SSE is found only for the categories /ɪr/ and /ʌr/, and only for women. However, the speaker-based perspective, in which individual vowel spaces are compared, once again qualifies these findings: when different kinds of (relatively categorical) partial merger in F1/F2 space are available to speakers of various age and gender groups, these different patterns are likely to neutralize each other and result in intermediate patterns.

Our results further show that in SSE the nurse lexical set is systematically constrained by spelling. The /ʌr/ vowel in particular has rather distinct acoustic properties corresponding to the spellings <or> and <ur>, respectively. The preceding phonetic
contexts of /ʌr/ might account for this distinction, since all of the <or> words produced as /ʌr/ occur after /w/ (e.g., work, word, world), while <ur> does not. It would also be interesting to investigate further when and how this factor becomes established, for example, by analyzing the speech of non-literate, i.e., pre-school children. As Lawson, Scobbie, and Stuart-Smith’s (2013) study suggests, even the pronunciation of adolescents aged 12 to 13 can still be unstable and be variably influenced by spelling in a word reading task. For the adults analyzed in this study, however, reading out or speaking freely does not influence the realization of the nurse vowels: no difference in their realization was found between scripted and unscripted speech. Further investigation of spelling would therefore benefit from a variety of lexical types and a relatively balanced number of tokens in each orthographic category. We suggest that this kind of research would partly have to be supported by experimental, rather than observational work.

The only genre effect found in this study was a more pronounced centralization and near-merger of /ɪr/ and /ɛr/ in non-broadcast talks. The tendency for a more closely spaced array of /ɪr/, /ɛr/, and /ʌr/ to emerge in the arguably less formal genre non-broadcast talks harmonizes with Wells’ (1982:407) association between the distinctness (of vowel categories) and more prestigious accents, particularly in the west of Scotland. We would argue that our findings reflect this tendency at a purely stylistic level, but the pattern will need to be investigated more rigorously across a broader range of genres.

A relatively strong effect on the realization of the nurse lexical set in SSE found in this study is the following /r/ variant: when speakers produce an approximant /r/, the preceding vowels are more mid-central and less distinct from each other. When the following /r/ is realized as a tapped or trilled variant, /ɪr/ and /ʌr/ are largely merged, whereas a vocalized realization of /r/ does not promote the tendency of /ɪr/-/ʌr/ merger and does not seem to pattern in a meaningful way relative to the other two types. Moreover, the realization of /ɪr/ is more strongly influenced by the variable realizations of the following /t/ variant than /ʌr/ and /ɛr/, which are less sensitive to the following phonetic contexts.

Our finding that /ɪr/ and /ʌr/ before a tapped or trilled /r/ are quite retracted and very close to each other fits one of the scenarios described by Stuart-Smith (2008:57-58), who says that, apart from the full three-way distinction and the mid-central merger, speakers may realize the contexts under investigation with a single vowel /ɪ/ or /ʌ/, or make a two-way distinction between /ɛ/ and /ʌ/. The latter seems to describe our findings quite well in combination with tapped or trilled realizations of /r/. Possibly, then, the relevant reference points for our data are not the three traditional vowels, but the two-way distinction between /ɪr/ and /ʌr/. These findings have potentially far-reaching consequences for assumptions concerning the history of vowel changes in the contexts under investigation: it is quite possible that in middle-class (or standard) accents the first target of the change was a two-way distinction as described above, with a later
(incipient) merger in /ɜ/. At the moment, however, this can be no more than an interesting speculation.

Our comments in the preceding paragraph need to be tested in further research. We would like to suggest, however, that the underlying coherence of phonological features within individual speakers of SSE can to some extent account for the relationship between variants of /r/ and vowel realizations. At least for middle-class speakers of SSE, the realization of /r/ as a tap or a trill can be considered “traditional,” while the approximant [ɹ] has been established as the current standard variant. Likewise, there are traditional distinctions made between historically different members of present-day NURSE—be it as two or three categories, which contrast with allegedly more recent (partial) mergers. It would therefore be expected that a speaker will be “traditional” or “innovative” across both features, and there is some limited evidence in our data to this effect. At the same time, preservation of more peripheral vowel qualities is articulatorily facilitated by /r/ articulations that involve the flexible tongue tip (taps or trills), while bunched variants of approximant [ɹ] with a primary constriction formed by the tongue body will exert a strongly centralizing coarticulatory pressure upon the preceding vowel. As such, it is therefore assumed that the two variables are likely to cohere not only socio-stylistically but also articulatorily. Previous studies on the covariation of linguistic variables have yielded contradictory results as to its existence (e.g., Becker [2016] on New York City English; Guy [2013] and Hinskens & Guy [2016] on Brazilian Portuguese). It would therefore be interesting to further investigate the cohesion of different linguistic features across individual Scottish English speakers in order to contribute to our understanding of the interrelationship between various phonological variables.

In conclusion, our study has demonstrated the benefits of combining large-scale corpus-based quantitative data with fine-grained qualitative analyses of individual speakers. We were thus able to show that four distinct types of NURSE realization patterns are in use across SSE speakers of all ages and genders. A corpus-based study, in contrast to classic sociophonetic study designs, has the advantage of including more speakers and thus being able to generalize from a firmer basis. Moreover, corpus data avoids the Observer’s Paradox effect ingrained in sociolinguistic interviews. On the downside, however, speakers in a corpus do not always produce sufficient tokens of the target items, leading to the exclusion of speakers in some contexts. Likewise, the exact (original) social background of speakers is not always provided: only information about their current professions was collected, which made the investigation of possible effects of social class challenging. Maybe in the future, our combined methodological approach can be further extended to include more controlled experimental tasks such as wordlists or providing cues that increase the chance of producing certain words; it would also be helpful to consider more detailed background information of individual speakers. We expect this to advance our understanding not only of the NURSE vowels but of phonetic and phonological features of SSE more generally.
### Appendix

**Model A. Fixed Effects Only**

| Parameter                        | Coeff. | 95% CI       | Eff. sample | R-hat |
|----------------------------------|--------|--------------|-------------|-------|
|                                  |        | Lower | Upper | Bulk | Tail |       |
| Intercept                        | 0.44   | 0.34  | 0.53  | 5194 | 8742 | 1.00  |
| GENDER woman                     | 0.02   | -0.04 | 0.08  | 4178 | 7681 | 1.00  |
| GENRE bnew                       | -0.10  | -0.28 | 0.09  | 5497 | 8559 | 1.00  |
| GENRE leg                        | 0.05   | -0.12 | 0.23  | 5020 | 8417 | 1.00  |
| GENRE nbtal                      | -0.00  | -0.27 | 0.26  | 5526 | 8741 | 1.00  |
| GENRE unsp                       | -0.00  | -0.16 | 0.15  | 4602 | 7871 | 1.00  |
| CATEGORY et                      | 0.01   | -0.57 | 0.57  | 6007 | 8365 | 1.00  |
| CATEGORY αf                      | 0.30   | -0.10 | 0.70  | 5778 | 8602 | 1.00  |
| SPEL <er>                        | -0.04  | -0.60 | 0.53  | 6018 | 8237 | 1.00  |
| SPEL <et>                        | 0.04   | -0.52 | 0.61  | 6054 | 8418 | 1.00  |
| SPEL <ur>                        | -0.16  | -0.57 | 0.24  | 5815 | 8898 | 1.00  |
| SPEL <or>                        | -0.22  | -0.62 | 0.18  | 5765 | 8730 | 1.00  |
| F2                               | -0.74  | -0.90 | -0.57 | 5449 | 8125 | 1.00  |
| GENDER woman : CATEGORY et       | -0.01  | -0.08 | 0.06  | 5246 | 7546 | 1.00  |
| GENDER woman : CATEGORY αf       | 0.03   | -0.03 | 0.09  | 5445 | 8914 | 1.00  |
| GENRE bnew : CATEGORY et         | -0.02  | -0.23 | 0.20  | 6387 | 8886 | 1.00  |
| GENRE leg : CATEGORY et          | 0.02   | -0.18 | 0.22  | 6220 | 9352 | 1.00  |
| GENRE nbtal : CATEGORY et        | -0.03  | -0.32 | 0.26  | 6219 | 8721 | 1.00  |
| GENRE unsp : CATEGORY et         | 0.16   | -0.02 | 0.34  | 5283 | 8265 | 1.00  |
| GENRE bnew : CATEGORY αf         | -0.03  | -0.24 | 0.19  | 8376 | 11,279 | 1.00 |
| GENRE leg : CATEGORY αf          | -0.12  | -0.28 | 0.05  | 7118 | 9121 | 1.00  |
| GENRE nbtal : CATEGORY αf        | -0.15  | -0.40 | 0.10  | 6737 | 9324 | 1.00  |
| GENRE unsp : CATEGORY αf         | -0.03  | -0.17 | 0.12  | 6981 | 9640 | 1.00  |
| GENDER woman : F2                | 0.10   | 0.01  | 0.20  | 4060 | 7155 | 1.00  |
| GENRE bnew : F2                  | 0.05   | -0.24 | 0.33  | 5645 | 7652 | 1.00  |
| GENRE leg : F2                   | -0.14  | -0.40 | 0.12  | 5222 | 8146 | 1.00  |
| GENRE nbtal : F2                 | 0.24   | -0.14 | 0.61  | 6351 | 8673 | 1.00  |
| GENRE unsp : F2                  | 0.08   | -0.16 | 0.32  | 4179 | 7209 | 1.00  |
| CATEGORY et : F2                 | 0.40   | -0.18 | 0.99  | 5781 | 8777 | 1.00  |
| CATEGORY αf : F2                 | -0.34  | -0.85 | 0.15  | 6092 | 8816 | 1.00  |
| SPEL <er> : F2                   | 0.24   | -0.34 | 0.81  | 5633 | 8371 | 1.00  |
| SPEL <et> : F2                   | 0.17   | -0.42 | 0.74  | 5537 | 8706 | 1.00  |
| SPEL <ur> : F2                   | -0.08  | -0.57 | 0.41  | 6054 | 9009 | 1.00  |
| SPEL <or> : F2                   | -0.55  | -1.04 | -0.06 | 6065 | 8748 | 1.00  |
| GENDER woman : CATEGORY et : F2  | 0.03   | -0.11 | 0.18  | 4823 | 7246 | 1.00  |
| GENDER woman : CATEGORY αf : F2  | 0.01   | -0.11 | 0.13  | 6212 | 8718 | 1.00  |
| GENRE bnew : CATEGORY et : F2    | -0.07  | -0.47 | 0.33  | 5192 | 7201 | 1.00  |
| GENRE leg : CATEGORY et : F2     | -0.21  | -0.60 | 0.18  | 5315 | 7896 | 1.00  |
| GENRE nbtal : CATEGORY et : F2   | -0.44  | -0.95 | 0.05  | 6234 | 9036 | 1.00  |
| GENRE unsp : CATEGORY et : F2    | -0.31  | -0.66 | 0.05  | 4323 | 8129 | 1.00  |
| GENRE bnew : CATEGORY αf : F2    | -0.02  | -0.40 | 0.37  | 7797 | 10,563 | 1.00 |
| GENRE leg : CATEGORY αf : F2     | 0.17   | -0.16 | 0.50  | 7190 | 9801 | 1.00  |
| GENRE nbtal : CATEGORY αf : F2   | 0.02   | -0.42 | 0.44  | 7360 | 9889 | 1.00  |
| GENRE unsp : CATEGORY αf : F2    | 0.17   | -0.12 | 0.46  | 6785 | 9220 | 1.00  |
### Model B. Fixed Effects Only

| Parameter                        | Coeff. | 95% CI  | Eff. sample |  |
|----------------------------------|--------|---------|-------------|---|
| Intercept                        | 0.48   | 0.34 | 0.62 | 8074 | 11,281 | 1.00 |
| Gender woman                     | -0.02  | -0.16 | 0.11 | 3958 | 6486  | 1.00 |
| Age                              | -0.00  | -0.01 | 0.01 | 6920 | 9893  | 1.00 |
| Genre bnew                       | -0.67  | -1.10 | -0.21 | 8547 | 11,241 | 1.00 |
| Genre leg                        | 0.06   | -0.12 | 0.24 | 7253 | 10,273 | 1.00 |
| Genre nbtal                      | -0.14  | -0.45 | 0.18 | 8202 | 10,527 | 1.00 |
| Genre unsp                       | -0.04  | -0.24 | 0.16 | 6362 | 9721  | 1.00 |
| Category ɛr                      | -0.06  | -0.65 | 0.53 | 9509 | 10,078 | 1.00 |
| Category ʌr                      | 0.29   | -0.10 | 0.70 | 8222 | 9877  | 1.00 |
| Spel 〈ɛr〉                       | -0.07  | -0.64 | 0.52 | 9597 | 9992  | 1.00 |
| Spel 〈ɔr〉                        | 0.01   | -0.56 | 0.60 | 9662 | 9993  | 1.00 |
| Spel 〈ur〉                        | -0.15  | -0.54 | 0.24 | 8077 | 9910  | 1.00 |
| Spel 〈ɔɾ〉                        | -0.23  | -0.62 | 0.15 | 7999 | 9582  | 1.00 |
| F2                               | -0.79  | -1.03 | -0.56 | 8049 | 10,477 | 1.00 |
| Gender woman : Age               | 0.00   | -0.00 | 0.01 | 4154 | 7195  | 1.00 |
| Gender woman : Category ɛr       | 0.02   | -0.14 | 0.17 | 4615 | 7247  | 1.00 |
| Gender woman : Category ʌr       | 0.08   | -0.06 | 0.21 | 4050 | 6335  | 1.00 |
| Age : Category ɛr                | 0.01   | -0.00 | 0.01 | 7877 | 9832  | 1.00 |
| Age : Category ʌr                | 0.00   | -0.01 | 0.01 | 7877 | 9832  | 1.00 |
| Genre bnew : Category ɛr         | 0.04   | -0.43 | 0.52 | 10,760 | 11,314 | 1.00 |
| Genre leg : Category ɛr          | -0.02  | -0.25 | 0.21 | 8953 | 10,835 | 1.00 |
| Genre nbtal : Category ɛr        | 0.08   | -0.29 | 0.43 | 9653 | 10,668 | 1.00 |
| Genre unsp : Category ɛr         | 0.27   | 0.02  | 0.52 | 7987 | 9749  | 1.00 |
| Genre bnew : Category ʌr         | -0.44  | -1.00 | 0.11 | 15,060 | 12,889 | 1.00 |
| Genre leg : Category ʌr          | -0.15  | -0.34 | 0.04 | 8841 | 10,768 | 1.00 |
| Genre nbtal : Category ʌr        | -0.03  | -0.34 | 0.28 | 8934 | 10,823 | 1.00 |
| Genre unsp : Category ʌr         | -0.03  | -0.22 | 0.17 | 7137 | 10,133 | 1.00 |
| Gender woman : F2                | 0.21   | -0.00 | 0.43 | 4229 | 7891  | 1.00 |
| Age : F2                         | 0.00   | -0.01 | 0.02 | 6516 | 9463  | 1.00 |
| Genre bnew : F2                  | -0.17  | -0.77 | 0.43 | 11,876 | 10,843 | 1.00 |
| Genre leg : F2                   | -0.20  | -0.50 | 0.10 | 7775 | 9834  | 1.00 |
| Genre nbtal : F2                 | 0.10   | -0.36 | 0.55 | 8985 | 11,346 | 1.00 |
| Genre unsp : F2                  | 0.08   | -0.23 | 0.40 | 7073 | 9878  | 1.00 |
| Category ɛr : F2                 | 0.54   | -0.07 | 1.13 | 10,531 | 10,786 | 1.00 |
| Category ʌr : F2                 | -0.35  | -0.87 | 0.17 | 8791 | 9933  | 1.00 |
| Spel 〈ɛr〉 : F2                  | 0.31   | -0.27 | 0.89 | 10,591 | 10,471 | 1.00 |
| Spel 〈ɔr〉 : F2                   | 0.21   | -0.37 | 0.80 | 10,788 | 10,513 | 1.00 |
| Spel 〈ur〉 : F2                   | -0.12  | -0.62 | 0.37 | 8974 | 10,361 | 1.00 |
| Spel 〈ɔɾ〉 : F2                   | -0.48  | -0.98 | 0.01 | 8702 | 9696  | 1.00 |
| Gender woman : Age : Category ɛr | -0.00  | -0.01 | 0.01 | 4968 | 7965  | 1.00 |
| Gender woman : Age : Category ʌr | -0.01  | -0.01 | 0.00 | 4427 | 7247  | 1.00 |
| Gender woman : F2                | -0.01  | -0.02 | 0.00 | 4372 | 7802  | 1.00 |
| Gender woman : Category ɛr : F2  | -0.13  | -0.41 | 0.14 | 4935 | 7879  | 1.00 |
| Gender woman : Category ʌr : F2  | -0.02  | -0.25 | 0.21 | 5141 | 8729  | 1.00 |
| Age : Category ɛr : F2           | -0.01  | -0.03 | 0.01 | 7136 | 9910  | 1.00 |
| Age : Category ʌr : F2           | 0.00   | -0.01 | 0.01 | 8278 | 10,335 | 1.00 |
| Genre bnew : Category ɛr : F2    | 0.24   | -0.45 | 0.94 | 13,460 | 11,796 | 1.00 |
| Genre leg : Category ɛr : F2     | -0.17  | -0.59 | 0.25 | 8606 | 10,097 | 1.00 |
| Genre nbtal : Category ɛr : F2   | -0.39  | -0.97 | 0.18 | 9960 | 11,246 | 1.00 |

(continued)
### Model B. (continued)

| Parameter | Coeff. | 95% CI | Eff. sample | R-hat |
|-----------|--------|--------|-------------|-------|
| GENRE unsp : CATEGORY et : F2 | -0.46 | -0.89 -0.02 | 8376 10,935 | 1.00 |
| GENRE bnew : CATEGORY af : F2 | 0.31 | -0.44 1.06 | 17,925 12,618 | 1.00 |
| GENRE leg : CATEGORY af : F2 | 0.21 | -0.13 0.54 | 10,634 10,707 | 1.00 |
| GENRE nbtal : CATEGORY af : F2 | 0.07 | -0.42 0.56 | 11,062 11,752 | 1.00 |
| GENRE unsp : CATEGORY af : F2 | 0.15 | -0.19 0.50 | 9779 10,231 | 1.00 |
| GENDER woman : AGE : CATEGORY et : F2 | 0.01 | -0.01 0.03 | 5013 8535 | 1.00 |
| GENDER woman : AGE : CATEGORY af : F2 | 0.00 | -0.01 0.02 | 5368 8621 | 1.00 |

### Model C. Fixed Effects Only

| Parameter | Coeff. | 95% CI | Eff. sample | R-hat |
|-----------|--------|--------|-------------|-------|
| Intercept | 0.58 | 0.39 0.77 | 3662 6782 | 1.00 |
| CATEGORY et | -0.10 | -0.32 0.11 | 3789 7978 | 1.00 |
| CATEGORY af | 0.00 | -0.23 0.24 | 3689 6643 | 1.00 |
| F2 | -0.94 | -1.32 -0.57 | 3207 5521 | 1.00 |
| GENDER woman | 0.06 | -0.09 0.21 | 3792 7306 | 1.00 |
| RHOTICITY deleted | -0.20 | -0.45 0.05 | 4001 8123 | 1.00 |
| RHOTICITY tap | -0.15 | -0.55 0.25 | 3982 7403 | 1.00 |
| CATEGORY et : F2 | 0.72 | 0.27 1.17 | 3333 6788 | 1.00 |
| CATEGORY af : F2 | -0.31 | -0.76 0.14 | 3508 5995 | 1.00 |
| CATEGORY et : GENDER woman | -0.07 | -0.24 0.12 | 4391 8349 | 1.00 |
| CATEGORY af : GENDER woman | -0.01 | -0.21 0.18 | 3963 6861 | 1.00 |
| F2 : GENDER woman | 0.05 | -0.18 0.29 | 3855 6967 | 1.00 |
| CATEGORY et : RHOTICITY deleted | 0.13 | -0.18 0.44 | 3955 8277 | 1.00 |
| CATEGORY af : RHOTICITY deleted | 0.16 | -0.13 0.46 | 4006 8282 | 1.00 |
| CATEGORY et : RHOTICITY tap | 0.07 | -2.04 2.21 | 10,295 12,078 | 1.00 |
| CATEGORY af : RHOTICITY tap | 0.03 | -0.42 0.48 | 4294 7613 | 1.00 |
| F2 : RHOTICITY deleted | 0.45 | 0.06 0.84 | 3897 6534 | 1.00 |
| F2 : RHOTICITY tap | -0.18 | -0.76 0.40 | 4784 8164 | 1.00 |
| GENDER woman : RHOTICITY deleted | 0.02 | -0.23 0.27 | 3686 6720 | 1.00 |
| GENDER woman : RHOTICITY tap | 0.20 | -0.19 0.58 | 4401 7870 | 1.00 |
| CATEGORY et : F2 : GENDER woman | 0.07 | -0.25 0.37 | 4391 8490 | 1.00 |
| CATEGORY af : F2 : GENDER woman | 0.06 | -0.25 0.36 | 4149 7671 | 1.00 |
| CATEGORY et : F2 : RHOTICITY deleted | -0.20 | -0.71 0.32 | 4200 8230 | 1.00 |
| CATEGORY af : F2 : RHOTICITY deleted | -0.57 | -1.03 -0.11 | 4132 6922 | 1.00 |
| CATEGORY et : F2 : RHOTICITY tap | 0.51 | -1.63 2.68 | 10,758 11,564 | 1.00 |
| CATEGORY af : F2 : RHOTICITY tap | 0.06 | -0.60 0.72 | 4952 8495 | 1.00 |
| CATEGORY et : GENDER woman : RHOTICITY deleted | -0.02 | -0.32 0.28 | 4251 7486 | 1.00 |
| CATEGORY af : GENDER woman : RHOTICITY deleted | -0.02 | -0.31 0.27 | 3916 7609 | 1.00 |
| CATEGORY et : GENDER woman : RHOTICITY tap | -0.08 | -2.19 2.08 | 10,568 10,820 | 1.00 |
| CATEGORY af : GENDER woman : RHOTICITY tap | -0.15 | -0.59 0.29 | 4394 8163 | 1.00 |
| F2 : GENDER woman : RHOTICITY deleted | -0.30 | -0.69 0.09 | 3637 6726 | 1.00 |
| F2 : GENDER woman : RHOTICITY tap | -0.06 | -0.60 0.48 | 4569 8602 | 1.00 |

(continued)
## Model C. (continued)

| Parameter | Coeff. | 95% CI | Eff. sample | R-hat |
|-----------|--------|--------|-------------|-------|
| CATEGORY $\varepsilon_r$: F2: gender woman: rhoticity deleted | 0.42 | -0.08 | 0.93 | 4276 | 7529 | 1.00 |
| CATEGORY $\alpha_r$: F2: gender woman: rhoticity deleted | 0.28 | -0.17 | 0.74 | 3870 | 7478 | 1.00 |
| CATEGORY $\varepsilon_r$: F2: gender woman: rhoticity tap | -0.54 | -2.71 | 1.66 | 10,028 | 11,525 | 1.00 |
| CATEGORY $\alpha_r$: F2: gender woman: rhoticity tap | -0.03 | -0.64 | 0.61 | 4644 | 7722 | 1.00 |

### Individual Vowel Spaces (see Table 5)

![Individual Vowel Spaces](image)
Acknowledgments
We are grateful for the very helpful comments we received from our two reviewers and the editors on an earlier version of this article.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the DFG (German Research Council; GU 548/13-1 and SCHU 3250/1-1).

ORCID iD
Zeyu Li https://orcid.org/0000-0003-3980-6581

Notes
1. We here used the term “Scottish English” as an umbrella term to refer to all forms of English spoken in Scotland ranging from Scots to Scottish Standard English.
2. All pre-rhotic nurse vowels were categorized according to their modern spelling, i.e., <ir> as /ɪr/, <er> and <ear> as /ɛr/, and <ur> and <or> as /ʌr/, following e.g., the general categorization in Wells (1982:407). Seventeen items with the spelling <our> were excluded, since the pre-rhotic vowels in words like tournament or Luxembourg may historically belong in a different phonemic category, namely /ur/. We are aware that the historical vowel may not match the modern spelling in all cases.
3. Uncertainties for /ɛr/ in combination with tapped/trilled variants of /r/ are extremely high, due to very low token numbers; they are therefore indicated in the plot using an asterisk rather than a cross with 50 percent and 95 percent uncertainty intervals, as they would have resulted in clutter and reduced readability. The absolute numbers of variants by vowel category are the following: /ɪr/ ($N = 116$ [10 taps, 84 approximants, 22 deleted]); /ɛr/ ($N = 150$ [2 taps, 94 approximants, 54 deleted]); /ʌr/ ($N = 228$ [30 taps, 110 approximants, 88 deleted]).

Corpora and Software
Boersma, Paul & David Weenink. 2017. Praat: doing phonetics by computer [Computer program]. Version 6.0.31, retrieved September 2018 from http://www.praat.org/ICE Scotland. https://www.ice-corpora.uzh.ch/en/joinice/Teams/icesco.html. May 2021.
Stan Development Team. 2018. Stan modeling language users guide and reference manual. Version 2.18.0, retrieved December 2018 from http://mc-stan.org

References
Abercrombie, David. 1979. The accents of standard English in Scotland. In Adam J. Aitken & Tom McArthur (eds.), Languages of Scotland, 68-84. Edinburgh: Chambers.
Aitken, Adam J. 1979. Scottish speech: A historical view with special reference to the Standard English of Scotland. In Adam J. Aitken & Tom McArthur (eds.), Languages of Scotland, 85-118. Edinburgh: Chambers.

Becker, Kara. 2016. Linking community coherence, individual coherence, and bricolage: The co-occurrence of (r), raised bought and raised bad in New York City English. Lingua 172-173. 87-99.

Bürkner, Paul-Christian. 2019. brms. Bayesian Regression Models using ‘Stan.’ R-package version 2.8.0, retrieved October 2019 from https://cran.r-project.org/web/packages/brms/brms.pdf

Cruttenden, Alan. 2008. Gimson’s pronunciation of English. London: Hodder Education.

Dickson, Victoria & Lauren Hall-Lew. 2017. Class, gender, and rhoticity: The social stratification of non-prevocalic /r/ in Edinburgh speech. Journal of English Linguistics 45(3). 229-259.

Dyer, Judy. 2002. “We all speak the same round here”: Dialect levelling in a Scottish-English community. Journal of Sociolinguistics 6(1). 99-116.

Giegerich, Heinz J. 1992. English phonology: An introduction. Cambridge: Cambridge University Press.

Grant, William. 1913. The pronunciation of English in Scotland. Cambridge: Cambridge University Press.

Greenbaum, Sidney (ed.). 1996. Comparing English worldwide: The International Corpus of English. Oxford: Clarendon.

Guy, Gregory R. 2013. The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? Journal of Pragmatics 52. 63-71.

Hinskens, Frans & Gregory R. Guy (eds.). 2016. Coherence, covariation and bricolage: Various approaches to the systematicity of language variation. Special issue of Lingua 172-173. 1-146.

Jackman, Simon. 2009. Bayesian analysis for the social sciences. Chichester: Wiley.

Jones, Charles. 2002. The English language in Scotland: An introduction to Scots. East Linton: Tuckwell Press.

Krug, Manfred. 2012. The Great Vowel Shift. In Alexander Bergs & Laurel Brinton (eds.), Historical linguistics of English: An international handbook, 756-776. Berlin: Mouton de Gruyter.

Kruschke, John K. 2015. Doing Bayesian data analysis. A tutorial with R, JAGS and Stan. Amsterdam: Academic Press.

Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2013. Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. Journal of Phonetics 41(3-4). 198-210.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. Journal of the Acoustical Society of America 49(2B). 606-608.

McAllister, Anne Hutcheson. 1938. A year’s course in speech training. London: University of London Press.

McElreath, Richard. 2016. Statistical rethinking. A Bayesian course with examples in R and Stan. Boca Raton, FL: CRC Press.

Nelson, Gerald. 1996. The design of the corpus. In Sidney Greenbaum (ed.), Comparing English worldwide. The international corpus of English, 27-35. Oxford: Clarendon.

Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. Exploring natural language: Working with the British component of the International Corpus of English. Amsterdam: John Benjamins.
Schiel, Florian. 2004. MAUS goes iterative. In Maria Teresa Lino (ed.), *Proceedings of the IV international conference on language resources and evaluation*, 1015-1018. Lisbon: University of Lisbon.

Schützler, Ole. 2013. The sociophonology and sociophonetics of Scottish Standard English (r). In Peter Auer, Javier Caro & Göz Kaufmann (eds.), *Language variation—European perspectives IV*, 215-228. Amsterdam: John Benjamins.

Schützler, Ole. 2015. *A sociophonetic approach to Scottish Standard English*. Amsterdam: John Benjamins.

Schützler, Ole, Ulrike Gut & Robert Fuchs. 2017. New perspectives on Scottish Standard English. Introducing the Scottish component of the International Corpus of English. In Sylvie Hancil & Joan Beal (eds.), *Perspectives on northern Englishes*, 273-301. Berlin: Mouton de Gruyter.

Sönning, Lukas & Ole Schützler. 2018. A normalization procedure for auditory vowel descriptions: Method and application. Paper presented at ISLE5, University College London, London.

Stuart-Smith, Jane. 2008. Scottish English: Phonology. In Bernd Kortmann & Clive Upton (eds.), *Varieties of English: The British Isles*, 48-70. Berlin: Mouton de Gruyter.

Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America* 88(1). 97-100.

Wells, John C. 1982. *Accents of English*. Vol. 1. Cambridge: Cambridge University Press.

**Author Biographies**

Zeyu Li is a research fellow at the Department of English Linguistics at the University of Münster. She received her PhD in English Linguistics from the University of Münster in 2020. Her research focuses primarily on phonetic, phonological, and sociolinguistic aspects of language variation and change and second language speech learning.

Ulrike Gut holds the Chair for English Linguistics at the University of Münster. Her main research interests are phonetics and phonology, second and third language acquisition, corpus linguistics, and varieties of English.

Ole Schützler holds a doctoral and a postdoctoral degree from the University of Bamberg and is currently Professor for Varieties of English at the University of Leipzig. He is broadly interested in synchronic and diachronic language variation and change at all linguistics levels and thus applies both sociophonetic and corpus-linguistic methodologies in his work.