A comparison of acoustic and articulatory parameters for the GOOSE vowel across British Isles Englishes

Eleanor Lawson,[1,a)] Jane Stuart-Smith,[2] and Lydia Rodger[1]
[1]*Clinical Audiology Speech and Language Research Centre, Speech and Hearing Sciences, Queen Margaret University, Edinburgh EH21 6UU, United Kingdom*
[2]*Glasgow University Laboratory of Phonetics/English Language and Linguistics, 12 University Gardens, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

(Received 23 April 2019; revised 1 November 2019; accepted 7 November 2019; published online 19 December 2019)

This study quantifies vocalic variation that cannot be measured from the acoustic signal alone and develops methods of standardisation and measurement of articulatory parameters for vowels. Articulatory-acoustic variation in the GOOSE vowel was measured across 3 regional accents of the British Isles using a total of 18 speakers from the Republic of Ireland, Scotland, and England, recorded with synchronous ultrasound tongue imaging, lip camera, and audio. Single co-temporal measures were taken of tongue-body height and backness, lip protrusion, F1, and F2. After normalisation, mixed-effects modelling identified statistically significant variations per region; tongue-body position was significantly higher and fronter for Irish and English speakers. Region was also significant for lip-protrusion measures with Scottish speakers showing significantly smaller degrees of protrusion than English speakers. However, the region was only significant for acoustic height and not for frontness. Correlational analyses of all measures showed a significant positive correlation between tongue-body height and acoustic height, a negative correlation between lip-protrusion and acoustic frontness, but no correlation between tongue-body frontness and acoustic frontness. Effectively, two distinct regional production strategies were found to result in similar normalised acoustic frontness measures for GOOSE. Scottish tongue-body positions were backer and lips less protruded, while English and Irish speakers had fronter tongue-body positions, but more protruded lips. https://doi.org/10.1121/1.5139215

[SF] Pages: 4363–4381

I. INTRODUCTION

Given the ease and sophistication of acoustic recording and analysis of vowels, for example, advances in automation of grapheme-to-phoneme conversion, segment-to-signal time alignment, normalisation, analysis, and plotting (Adank *et al.*, 2004; Bigi, 2015; Fabricius and Watt, 2002; Labov *et al.*, 2013), it is unsurprising that acoustic studies of accent variation in vowels dominate, while articulatory analyses are comparatively rare. It might seem that little is to be gained from instrumental articulatory analysis, where recording is generally more time-consuming and difficult (Narayanan *et al.*, 2011; Scobbie and Pouplier, 2010; Stone, 2005; Wrench and Hardcastle, 2000), and where smaller numbers of speakers and tokens are generally obtained. However, with a purely acoustic approach, it is possible that significant performative variation is not identified. For coda /r/ in English, articulatory analysis has previously identified the presence of delayed and covert lingual gestures where auditory coding identified segment deletion (Lawson *et al.*, 2018). It has also revealed radically different articulatory variants of coda /r/ that had gone unnoticed despite decades of auditory and detailed acoustic analysis (Lawson *et al.*, 2011, 2014), most likely due to the fact that the articulatory variation affected higher formants (F4 and F5) that were not routinely studied (Zhou *et al.*, 2008; Lawson *et al.*, 2018). In both cases, this apparently covert variation was socially stratified rather than idiosyncratic, and, therefore, meaningful in the speech communities studied. In the present paper, we consider articulatory-acoustic variation in vowel production and study regional performative variation that is masked when considering acoustic measures alone. Specifically, we study fronting of the GOOSE vowel [representing a set of lexical items in English that contain the vowel /u(ː)/; Wells, 1982a] in different regional accents of British Isles English from the corpus collected for the audio-articulatory Dynamic Dialects Web resource between 2012 and 2014.[1] Motor equivalence in the GOOSE vowel had been studied before using electromagnetic articulography (EMA). Perkell *et al.* (1993) showed that the objective of articulatory variation for American /uː/ was a stable acoustic target by revealing articulatory trade-offs between lip rounding and tongue-body raising (to form a velo-palatal constriction) for /uː/, both of which served to lower second formant (F2) values. We study a different scenario, where F2 of GOOSE vowels have converged across regional varieties due to the effects of distinct sound-change processes.

The current study further develops methods to standardise articulatory ultrasound tongue imaging (UTI) between

---

[a)]Electronic mail: elawson@qmu.ac.uk

speakers, adapts UTI vowel-normalisation methods set out in Scobbie *et al.* (2012), and sets out a method to measure and normalise lip protrusion between speakers. It uses these methods to identify and compare patterns of regional variation (Scotland, the Republic of Ireland, and England) across co-temporal acoustic and articulatory measures. We focus on the GOOSE vowel as it has already been shown that this vowel's formant values are altered by both tongue-body position and lip rounding, and it is difficult to tease apart the contributions made by these two articulatory parameters from the acoustic signal alone; see Secs. I B and I C.

Our study was motivated by three factors: (i) the apparent similarity of F2 measures for GOOSE across accents of the British Isles in large-scale acoustic surveys, such as Ferragne and Pellegrino (2010), despite (ii) very different diachronic trajectories of GOOSE fronting in different regional accents, (iii) accounts and observations of different lip postures for the GOOSE vowel in different regional varieties of British English and accounts by early phoneticians of performative variation in the GOOSE vowel in regional accents; see McAllister (1938). More will be said about these motivations in Sec. I A.

A. GOOSE fronting in the British Isles

GOOSE[2] acts as a keyword for the set of lexical items in English that contain the vowel phoneme /u(ː)/[3] (Wells, 1982a). Fronting of the GOOSE vowel in English is not a new phenomenon. In the 1980s, Wells noted that a GOOSE vowel with a back, rather than a central, quality was a feature of conservative varieties of English (Wells, 1982a, Sec. 2.2.15). However, studies of fronted GOOSE have become increasingly common over the past couple of decades. A large number of studies focus on GOOSE fronting as a change in progress in the south of England with apparent-time studies showing fronting over at least the past five decades; see Przedlacka (2001), Hawkins and Midgley (2005), Fabricius (2007), Harrington *et al.* (2011), and Harrington *et al.* (2008). There are also reports of fronted GOOSE in the Republic of Ireland (Hickey, 2016), an area where high-back GOOSE vowel variants were prevalent (Wells, 1982a, Sec. 2.2.15). In a recent broad-based acoustic survey of British Isles vowel systems, Ferragne and Pellegrino (2010) found fronted GOOSE variants in a majority of accents surveyed. Additionally, GOOSE fronting has been identified in varieties of English worldwide, e.g., in New Zealand English (Gordon *et al.*, 2004), South African English (Mesthrie, 2010), and North American English (Boberg, 2011).

The current research, focussed on GOOSE fronting, as well as Ferragne and Pellegrino's acoustic vowel survey, could give the impression that the same sound change is occurring throughout the British Isles in the same way. For example, Ferragne and Pellegrino found comparable degrees of F2 frontedness after normalisation and Bark transformation in Glasgow, East Anglia, Birmingham, Standard Southern British English (SSBE), Liverpool, and Lancashire accents (Ferragne and Pellegrino, 2010), although they did also identify differences in GOOSE height between Glasgow and Ulster accents and other British Isles accents. Historical evidence, on the other hand, shows that GOOSE fronting has had very different diachronic trajectories in different parts of the British Isles. For example, we know that while fronting of the GOOSE vowel is comparatively recent in southern England, in northern English varieties, including Scottish English, centrality or frontness of GOOSE can be attributed, in part, to sound changes that occurred in the late 13th century, beginning with the fronting (and unrounding) of ME *ọ̄* (Johnston, 1997). This early northern sound change resulted in a range of front monophthongal and diphthongal older rural and dialectal reflexes for the GOOSE lexical set; Scottish English [gys/gʏs] *goose*; north-west Midlands [gɪʊs/gʏːs], Yorkshire [gʊɪs] (Johnston, 1997, Sec. 3.3.1.2); see also the *Linguistic Atlas of England* (Orton *et al.*, 1978), maps Ph138–142, which record vowel variants [iɤ, iə, iu, y, ui] for reflexes of Middle English *ō* (*moon, goose, boots*, etc.) in the North of England, but [u] in the Midlands and further south. There are also majority forms with vowels /y/ and /ɪ/ recorded for *moon*, *spoon*, *roof*, *tooth*, etc. in the *Linguistic Atlas of Scotland* (Mather and Speitel, 1986, pp. 360 and 368), which recorded older, rural variants.

Early impressionistic phonetic accounts of GOOSE vowel variation capture some of the performative variation present in the Scottish GOOSE vowel that is not represented in present-day acoustic studies, drawing attention to variation in both tongue and lip positions. McAllister (1938) gives an account of exolabial versus endolabial lip rounding between Central Scottish FOOT/GOOSE and Standard English GOOSE in the 1930s, respectively:

> "The change in lip rounding makes even a more marked difference in the vowel quality than the change in tongue posture. The local [Central Scottish] pronunciation of **u** in (d*o*) is produced with the lips closely rounded against the teeth, the centre of the upper lip being drawn downwards to the lower lip. For the standard [Anglo-English] vowel, the lip rounding should be full and loose, the lips being protruded well forward beyond the teeth, the centre of the upper lip turned, upward and outward, away from the lower lip and kept free from contact with the upper teeth…" (McAllister, 1938, note ii).

It could be argued that important articulatory details such as these have been overlooked to some extent since the advent of speech spectrography in the 1940s (Joos, 1948; Delattre *et al.*, 1952). Today, the majority of studies of diatopic vowel variation involve acoustic analysis focussed on the F2-F1 plane. Few studies consider vocalic variation and change from an articulatory perspective, or try to separate out the effect of the tongue and the lip positions for rounded vowels. However, there are vocal-tract modelling studies that consider the impact of separate articulatory parameters on acoustic output, which we will discuss in Sec. II B.

B. Articulatory-acoustic relations in modelled vowel systems

As Scobbie *et al.* (2012) point out, since the early acoustic work of Joos (1948), Cooper *et al.* (1952), etc., the first and second formants have been considered key perceptual correlates of the height and front-back dimensions of the articulatory vowel space, respectively; see also, Bladon and Fant (1978) and Savariaux *et al.* (1995). However, the similarity between traditional articulatory-auditory-based vowel space diagrams (Bell, 1887; Jones, 1909) and formant plots (Joos, 1948) could give the impression that there is a one-to-one mapping of tongue-body frontness to F2 and tongue-body height to an inverse of F1, overlooking the effects of lip protrusion on the first two formants. Early three-parameter (tongue constriction location, constriction size, and lip constriction ratio) vocal-tract modelling studies, and later acoustic-articulatory comparisons, have shown that F1 and F2 are altered by lip constrictions (Lindblom and Sundberg, 1971; Stevens and House, 1955; Stevens, 1998), particularly F2 (Fant, 1992, Fig. 5; Savariaux *et al.*, 1995, Fig. 1).

It is often assumed that lip position effects are recoverable from variation in F3. An intrinsically normalised F3-F2 measure has been suggested to us as a potential acoustic correlate of lip protrusion (see Sec. II H); however, articulatory-acoustic models show that there is no straightforward correlation between F3 and lip rounding across the vowel space. The effects of lip rounding on F3 vary depending on other articulatory parameters such as constriction location. Lindblom and Sundberg's articulatory-acoustic study of Swedish vowels, using x-ray-based vocal-tract models, quantified the effects of independent variation of articulatory parameters on derived formant frequencies, showing that F3 lowering was dependent on tongue-body shape variation: neutral and with palatal, velar, and pharyngeal bunching. Lip rounding resulted in greater degrees of F3 lowering when there was a palatal tongue constriction (Lindblom and Sundberg, 1971). These findings are also supported by Fant's vocal-tract nomograms (Fant, 1992, Fig. 5).

One pertinent finding of many vocal-tract modelling studies that consider the relative contribution of tongue-body and lip positions in /u/ production is that different strategies involving these two articulators can be used to achieve characteristic lowered F2 values, so-called "motor equivalence" (Perkell *et al.*, 1993). The electric vocal-tract analogue by Stevens and House predicted that more than one vocal-tract configuration could produce F1 and F2 formant frequencies associated with the [u] vowel. Their vocal-tract analogue was set up to allow variation of (1) constriction location along the length of the vocal-tract tube, (2) tube radius at the constriction, and (3) a ratio measure of aperture area and length for the lip tube (the lower the value, the more constricted the lip tube, or the longer the lip tube). Using average formant data from 33 adult American male speakers (Peterson and Barney, 1952), Stevens and House found that average F1 and F2 formant values for [u] could be obtained using 2 different vocal-tract parameter settings, one with a fronter vocal-tract tube constriction and more constricted/longer lip-tube setting and the other with a backer vocal-tract tube constriction and less constricted/shorter lip-tube setting; see Stevens and House (1955, Fig. 7).

Below, we report on some studies where articulatory analysis and perceptual methods have been used to study the roles of multiple articulators in the production of the GOOSE vowel in different accents of English. The present study continues such an approach, aiming to determine whether articulatory analysis can provide a more detailed picture of regional variation for this widely studied vowel.

C. Articulatory-acoustic studies of the GOOSE vowel

Harrington *et al.* (2011) realised that the F2 raising associated with auditory GOOSE vowel fronting in SSBE could be due to either tongue-body fronting, lip unrounding, or a combination of both. They developed a range of techniques to identify the contributions of tongue-body movement and lip movement to auditory fronting of /u/, including acoustic (spectral centre of gravity) analysis of the coarticulatory effect of /u/ on the preceding /s/, compared with the effect of an unrounded vowel /i/. Harrington *et al.* hypothesised that if the fronting of /u/ was due to unrounding, /s/ before /i/ and /u/ would be more acoustically similar in a younger GOOSE-fronting speaker group than in an older speaker group with lesser degrees of GOOSE fronting. An (audio)-visual perception experiment was also carried out, where video recordings were made of young SSBE speakers producing /uː/. The videos were presented to German speakers with an /iː/ overdubbed or no audio signal. Harrington *et al.* hypothesised that if lip rounding were still present in /uː/, interaction between vision and hearing (McGurk and MacDonald, 1976) would lead the German speakers to classify the vowel as "front rounded." Finally, Harrington *et al.* used direct articulatory evidence from EMA to determine the position of the lips and the tongue for /uː/ in relation to other vowels in the SSBE system. These experiments each provided evidence that present-day SSBE /u:/ was produced with a fronted tongue-body and lip rounding (Harrington *et al.*, 2011).

A preliminary UTI study of the GOOSE vowel in eastern Central Scotland was carried out by Scobbie *et al.* (2012) with the additional aim to address some of the fundamental issues relating to the articulatory measurement of the two-dimensional (2-D) midsagittal tongue-body position during vowel production, such as defining "horizontal" and "vertical" for articulatory measures and normalising frontness and height measures. The articulatory-acoustic study by Scobbie *et al.* used a single token of the GOOSE vowel per speaker ($N = 15$) in the word "boom." Measurements were taken at a single time point from the highest point of the tongue (Jones, 1917) for each GOOSE-vowel token and the full set of monophthongal stressed Scottish vowels (Scobbie *et al.*, 2012). Lip position was not recorded. Comparison of speakers' GOOSE vowels with other vowels in the system, particularly the FACE vowel, led Scobbie *et al.* to assert that the Central Scottish GOOSE vowel was a "truly front" vowel (Scobbie *et al.*, 2012). Regarding the height of the GOOSE vowel, it was found to be articulatorily lower than that of FACE; however, there was a mismatch between the results of the articulatory and acoustic analysis, as mean F1s (in Bark) for FACE and GOOSE were found to be similar.

Blackwood Ximenes *et al.* (2017) also studied the GOOSE vowel, alongside the main monothongal vowels in Australian (four speakers) and American (five speakers) English, using EMA and acoustic measures. Articulatory and acoustic measures were taken at the tangential minimum velocity of a coil placed on the tongue dorsum, i.e., capturing the time point of the constriction maximum of the vowel target. An articulatory measure was taken from the tongue-dorsum flesh point and the upper and lower lip flesh points. F1 and F2 were also measured at this time point and all articulatory and acoustic measures underwent extrinsic normalisation using *Z* scoring (Lobanov, 1971). Using data from all vowels, Blackwood Ximenes *et al.* (2017) found a strong inverse correlation between normalised tongue height and normalised F1, and a strong positive correlation between tongue fronting and normalised F2. However, in both American and Australian English, there was a mismatch for the GOOSE vowel between the horizontal tongue dorsum position and F2. In each variety, the tongue dorsum position relative to other vowels in the system appeared to be fronter than it was in acoustic space, overlapping in some cases with the positions of KIT and FLEECE vowels. GOOSE vowels in both varieties were found to have the greatest degree of lip protrusion of all the vowels studied, suggesting that lip protrusion was lengthening the front cavity and lowering F2; however, while GOOSE can be described as "back" in acoustic space in American English, it is acoustically "central" in Australian English, despite similar degrees of lip rounding. They attribute this mismatch between the data of the two varieties to potential differences in the posterior tongue surface that cannot be recorded with the EMA technique, namely a potentially larger pharyngeal cavity for Australian speakers' GOOSE vowels.

Finally, Savariaux *et al.* (1995) used a perturbation experiment, involving a lip tube, with mid-sagittal x ray to investigate the impact of compensatory strategies on the F1-F2 space used to produce French /u/. In the perturbation trials, 7 of their 11 speakers moved their tongue backward to maintain F2 values observed in the non-perturbed trials. Savariaux *et al.* measured all three formants, but concentrated on F1 and F2 (Savariaux *et al.*, 1995, p. 2433).

These findings pertaining to /u/-vowel variants, based on vocal-tract modelling and, later, articulatory-acoustic studies, highlight the complexity of the relationship between articulatory movement and the acoustic signal produced for the GOOSE vowel set across English (and also in French). They provide justification for undertaking the current study using both articulatory and acoustic data. In the current study, we further assess the relationship between articulatory and acoustic parameters, while also developing UTI measurement and normalisation methods pioneered in Scobbie *et al.* (2012), and we develop methods for lip measurement and normalisation from profile lip video. Our research questions are as follows:

1. How can we standardise and normalise inter-speaker tongue-body measures, recorded with UTI, and lip measures recorded with the profile lip camera?
2. Do we see regional patterns of articulatory variation that are distinct from regional patterns of acoustic variation?
3. What correlations do we find between articulatory measures of tongue-body height and frontness and lip protrusion, and acoustic F1, F2 measures?

We will propose methods of articulatory standardisation, relating to the UTI recording technique and the lip-camera data we have collected. We hypothesise that while similar F2 measures have been obtained for different regional accent groups within the British Isles (Ferragne and Pellegrino, 2010), these values might be achieved through different production strategies involving the tongue-body position and the lips, reflecting the different sound-change trajectories of the GOOSE-fronting processes. We also investigate GOOSE lowering, which has received less attention than GOOSE fronting to date, though, see Scobbie *et al.* (2012) and Stuart-Smith *et al.* (2017). We hypothesise that we will find a correlation between tongue-body height and F1, but that the relationship between tongue-body frontness and F2 will be more complex due to the impact that lip protrusion is known to have on the second resonance.

II. METHODOLOGICAL ISSUES AND METHODOLOGY

In this study, we use UTI to study tongue position associated with vowels, along with the profile lip camera. One main advantage of the UTI technique in speech analysis is the fact that it is not invasive and has shorter set-up times than other techniques such as EMA (Blackwood Ximenes *et al.*, 2017; Fant, 1992; Lee *et al.*, 2016) and EPG (Scobbie and Wrench, 2003). However, there are a few challenges associated with UTI and lip video, namely, stabilisation of the ultrasound probe, establishing vertical and horizontal axes for the physical vowel space across multiple speakers, standardising measurement locations, establishing protocols for lip measurement, and normalisation of measures. In Secs. II A–II K, the study's dataset is initially described, and thereafter these challenges are discussed, and methods to meet the challenges are detailed.

A. Speaker corpus

Data used in this study were not collected specifically to study GOOSE vowel variation. We made use of a subset of a pre-existing audio-ultrasound speech corpus: Dynamic Dialects, recorded between January 2012 and January 2014 in Edinburgh, U.K.[1] In the present study, there were 18 speakers from the British Isles: 9 females and 9 males with most speakers aged between 20 and 35 years old, and one speaker aged 48 years old; see Table I. All speakers in the study self-identified as middle class. Although there are a small number of speakers in the dataset compared to most acoustic studies, this study has a greater number of speakers than most articulatory-based accent studies, e.g., 12 speakers, 1 token per speaker (Scobbie *et al.*, 2012), 9 speakers (Blackwood Ximenes *et al.*, 2017), and 5 speakers (Harrington *et al.*, 2011). Speakers in the current subcorpus come from three regions of the British Isles: England (seven speakers), the Republic of Ireland (three speakers) and Scotland (eight speakers). Initially, there were 20 speakers in the subcorpus; however, 1 male and 1 female speaker were excluded from the study. A male speaker from Kent was excluded because we could not obtain a clear ultrasound image of his FLEECE vowel tokens, which were needed for articulatory normalisation. One female

TABLE I. List of speakers used in the present study, along with demographic information.

Region	Speaker number	Location	Age	Gender
Scotland	S19	Orkney	31	Female
	S18	Inverness-shire	21	Female
	S6	Aberdeenshire	48	Female
	S2	Perthshire	23	Male
	S9	Fife	22	Male
	S8	West Lothian	29	Male
	S3	South Lanarkshire	20	Female
	S10	Renfrewshire	35	Male
Republic of Ireland	S13	County Monaghan	23	Female
	S14	Dublin	26	Female
	S23	County Tipperary	25	Female
England	S7	Newcastle	21	Male
	S21	North Yorkshire	24	Female
	S22	Sheffield	30	Male
	S24	Sheffield	22	Male
	S4	Greater Manchester	23	Female
	S25	London	25	Male
	S5	Southampton	20	Male

FIG. 1. British Isles map showing the location where study participants spent the majority of their lives.

speaker from County Antrim in Northern Ireland was excluded as she had markedly different allophones of the GOOSE vowel in different Scottish-vowel-length-rule (SVLR) contexts (Aitken, 1981); short [ɵ] before voiceless consonants in *goose*, *hoop*, *root*, etc., and diphthongal [əy] or [əʉ] before voiced consonants, a morpheme boundary, or in open syllables, e.g., in *choose*, *brewed*, *Sue*. We investigated the possibility that there could be qualitative differences between GOOSE vowels in long and short SVLR contexts in the Scottish cohort, but we did not identify any significant differences in quality based on the articulatory and acoustic measures taken in the study.

As audio-visual recordings (showing the lower portion of the face) of these speakers is available online, to avoid identifying speakers, their location is referred to by city if the speaker came from a large city, and by county if they came from a smaller town or village. Speaker numbers are small and geographical coverage of the British Isles is uneven; see Fig. 1. However, this is the first articulatory study of a variety of regional British Isles Englishes, where tongue and lip movements are available alongside audio recordings. Speaker S10 from Renfrewshire in Scotland was the only speaker to have been recorded over two sessions, and only tokens from the first session are included in the articulatory and acoustic analysis. Lip data were not available for the speaker from Orkney due to a recording equipment malfunction.

In this study, we examine speakers by region of the British Isles: England, the Republic of Ireland, and Scotland. The English phonology of the Republic of Ireland most closely matches that of Anglo-English due to the English settlement of Ireland since the early middle ages (Wells, 1982b, Sec. 5.3.1), whereas Northern Irish English is phonologically closer to Scottish English (Wells, 1982b). Wells stated that the conservative nature of Irish English had resulted in better preservation of a truly back GOOSE vowel quality, while most urban British Isles varieties used a centralised variant (Wells, 1982a, Sec. 2.2.15); however, as already mentioned, Hickey has identified GOOSE fronting in young Irish speech (Hickey, 2016). While fronted monophthongal and diphthongal variants of GOOSE have been evident in older, rural speech in the North of England, as mentioned in Sec. I A, these variants are evanescent and, generally, less evident in middle-class speech; see Wells (1982b, Sec. 4.4.4). Wells described the Standard Scottish GOOSE vowel as central [ʉ], or centralised front [ʏ] (Wells, 1982b, Sec. 5.2.3), and there is evidence of further fronting and lowering of this vowel in Central Scotland throughout the 20th century (Stuart-Smith *et al.*, 2017). Therefore, we see distinct realisations of GOOSE in England, the Republic of Ireland, and Scotland, and there is evidence also of ongoing change, not necessarily toward the same target.

B. Word list

The word list containing a total of 106 items was not specifically designed to capture examples of the GOOSE vowel. Only one repetition of the full word list was collected for each participant. There were 13 target word-list items containing the GOOSE vowel: *goose*; *smooth*; *choose*; *brewed*; *hoop*; *coop*; *brood*; *sue*; *(this) room*; *root*; *this (shoe)*, although not all items were produced by each speaker or were measureable. A mean of 11 (standard deviation, s.d., 1.5) acoustic and articulatory tokens from these words were analysed for each speaker. "Root" and "room" can be pronounced by some speakers with an [ʊ] vowel, and "do" can also be pronounced in a reduced manner—articulatorily centralised with reduced lip rounding, even in citation form; however, we used the random factor *prompt* in our mixed effects modelling to account for variation attributable to pronunciation of particular stimuli.

C. Ultrasound recording scenario and probe stabilisation

UTI allows the imaging of most of the sagittal tongue surface and automatic identification and measurement of any point on that tongue surface, rather than sampling three or four points on the tongue as with EMA, where identification of sagittal portions of the tongue are approximate and likely to vary between speakers (Blackwood Ximenes *et al.*, 2017; Fant, 1992; Lee *et al.*, 2016).

Ultrasound recordings were made using a Sonix RP medical ultrasound machine (Ultrasonix, Vancouver, Canada), operating at 120 scans per second, located in a purpose-built sound studio. All noise-making equipment, such as the ultrasound machine and personal computer (PC) hard drive, were located in room adjacent to sound studio. Audio was recorded using an Audio-Technica AT803D clip-on condenser microphone (Audio-Technica, Tokyo, Japan), attached near the speaker's mouth and clipped to the ultrasound probe's stabilising headset. Audio recordings were sampled at 22 kHz.

Stabilisation of the ultrasound probe and reduction of pitch (sagittal rotation), yaw (axial rotation), and roll (coronal rotation) of the probe is essential for obtaining a coherent dataset to quantify the articulatory vowel space. Pitch movements of the probe result in clockwise and anticlockwise rotation of the 2-D midsagittal tongue surface. If pitch movements are not reduced or eliminated during the recording session, they can be detected and corrected *post hoc* to make data useable (see Mielke *et al.*, 2005). Yaw movements result in the probe no longer recording the midline of the tongue, and are likely to produce discontinuities in the imaged tongue surface and result in misleading and unusable data. Roll movements lead to imaging of either side of the midline of the tongue and usually result in an indistinct image of the tongue surface (the tip and root, in particular). The further the probe shifts from the midline, the more likely it is that multiple surfaces will be imaged due to reflections from both the grooved midline of the tongue and the tongue surface to the left and right of the midline. Methods of stabilisation other than headset-stabilisation exist; see Mielke *et al.* (2005).

In Scobbie *et al.* (2012) and the current study, an aluminum probe-stabilisation headset was used, with 13 adjustable sections to allow it to be fitted to different sized heads. The probe is held in place underneath the chin by the headset, which is stabilised against the top of the speaker's head, their cheekbones, and the sides and back of their head. The headset prevents roll and yaw movements of the probe and greatly reduces pitch (sagittal rotation) movement (Scobbie *et al.*, 2008). An added advantage of using a stabilising headset is the possibility of using headset-mounted micro cameras to film lip movement. In the current study, a profile micro camera was fitted to a bracket extending from the right-hand side of the headset. A front-facing lip camera was added later in the project, and not all participants were recorded with the front-facing camera; therefore, measurements were taken using the profile lip camera only.

D. Imaging the occlusal plane

The occlusal plane, i.e., the speaker's bite plane, is an axial plane passing through the occlusal (biting) surfaces of the teeth. Imaging and recording the position of the occlusal plane in each recording session improves interpretation of tongue position and inter-speaker comparison. A sagittal trace of the occlusal plane can be achieved with UTI using a plastic bite plate, or other flat surface (e.g., see Blackwood Ximenes *et al.*, 2017), placed in the speaker's mouth and gripped between the incisors, premolars, and molars. In UTI studies, speakers are asked to press their tongue against the underside of the bite plate, which results in their tongue bulging upward at the back edge of the bite plate. The quasi-horizontal image of the occlusal plane becomes visible in the UTI image; see Fig. 2. At the beginning of each UTI recording in the corpus, images of the speakers' occlusal planes were obtained and the probe-to-chin angle was adjusted and set using the stabilising headset, so that the image of the occlusal plane was observed to be parallel to the upper and lower edges of the video pane.

E. Establishing 2-D axes for UTI data

Scobbie *et al.* (2012) posed a key methodological question for studies involving measures of tongue-body position: "What is an appropriate horizontal axis for the articulatory vowel space?" In acoustic analysis, the primary axes for plotting vowel position are the continua along which the F1 and F2 values vary, the "horizontal" axis being the F2 continuum. This question of physical articulatory axes is particularly important where quantification methods that involve single-point measurements from the tongue surface (Jones, 1917) are used. Changing the rotation of a tongue surface in a 2-D space results in measurements of different points on the tongue's surface, affecting raw measures and also normalised values that involve measures from other vowel tongue positions. Rotation of midsagittal tongue surfaces is a particular problem for UTI recordings where the ultrasound probe is set at slightly, or radically, different angles relative to the cranium for each recording session.

We measured the effect of tongue-surface rotation on tongue-body position by comparing the x and y distances between the highest points of FLEECE and GOOSE at different rotations. Mean tongue surface contours were created for the FLEECE and GOOSE vowels for one near received pronunciation (RP) speaking Anglo-English female speaker using Articulate Assistant Advanced (AAA) (Wrench, 2012). The tongue surface contours were then rotated at 10°

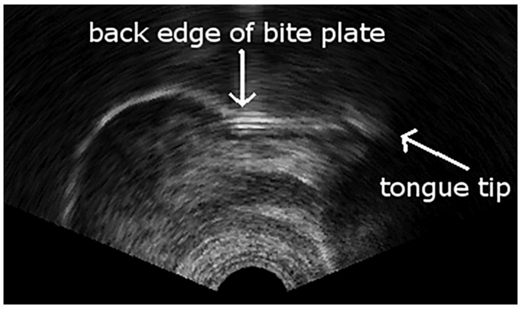

FIG. 2. Bite plane image after probe-to-chin angle adjustment. The horizontal flat section of the tongue in the image shows the area where the tongue is pressed against the bite plate.

increments to different angles, relative to the occlusal plane, and then their Cartesian coordinates were exported for plotting and obtaining a highest point of the tongue measurement using *R* (R Core Team, 2018). Figure 3 shows plots of the tongue surfaces at different rotations relative to the occlusal plane: 90° rotation shows the tongue-surface splines when the ultrasound probe is positioned at right angles to the speaker's occlusal plane; for lesser degrees of rotation (60°–80°), the probe would be angled more toward the speaker's throat; and for greater degrees of rotation (100°–110°), the probe would be angled more toward the speaker's chin. The highest point of each tongue curve, automatically identified using an *R* script, is marked on each contour plot with an "×" (grey for FLEECE and black for GOOSE). The raw horizontal and vertical distances between the highest point of the tongue for FLEECE and GOOSE are graphed across each rotation in Fig. 4.

Figure 4 shows that while the effect of spline rotation in the 2-D vowel space on vertical distance measures GOOSE (y) to FLEECE (y) is almost non-existent—less than 1 mm between rotations 60° and 110°; horizontal variation in GOOSE (x) to FLEECE (x) distance measures is more striking—between -0.5 mm and 4.1 mm, i.e., GOOSE (x) to FLEECE (x), for this speaker, is sometimes a positive value and sometimes a negative value, depending on rotation, due to the intersection of the tongue surfaces. Rotation will have an even greater effect on automatic articulatory measures if normalisation is carried out using corner vowels. Regardless of the speaker or vowels chosen, changing the rotation of tongue surfaces will affect any distance measures made using automatic highest point of the tongue measures. We cannot say that there is a "correct" angle of rotation, but limiting rotational variation in probe position between recording sessions is required, and standardisation on a particular probe rotation arguably results in more comparable measures of tongue-body location than, e.g., placement of an EMA coil on a speaker's tongue, the location of which will vary from speaker to speaker depending on tongue shape, strength of gag reflex, and changes in location between the extended-protruded and relaxed tongue when EMA coils are attached.

Scobbie *et al.* suggest two possible approaches to the inter-speaker standardisation of the rotation of the articulatory vowel space: method 1, occlusal plane, standardising on the speaker's (approximated) occlusal plane, i.e., having each speaker's bite plane as the horizontal axis; and method 2, common tangent, drawing a tangent from the tongue surface of the two high corner vowels (in the study by Scobbie *et al.*, these were FLEECE and GOAT) and using this tangent as the horizontal axis; see Scobbie *et al.* (2012, Fig. 3).

With both vowel space rotation standardisation techniques by Scobbie *et al.*, tongue-body position can be quantified using highest point of the tongue measures, although different values are inevitably obtained with each method. One potential issue associated with the common tangent technique is the possibility for inter-speaker variation relating to the position of the highest point of the tongue for the corner vowels due to accent variation or variation in speech rate. In the present study, each speaker's tongue surface was

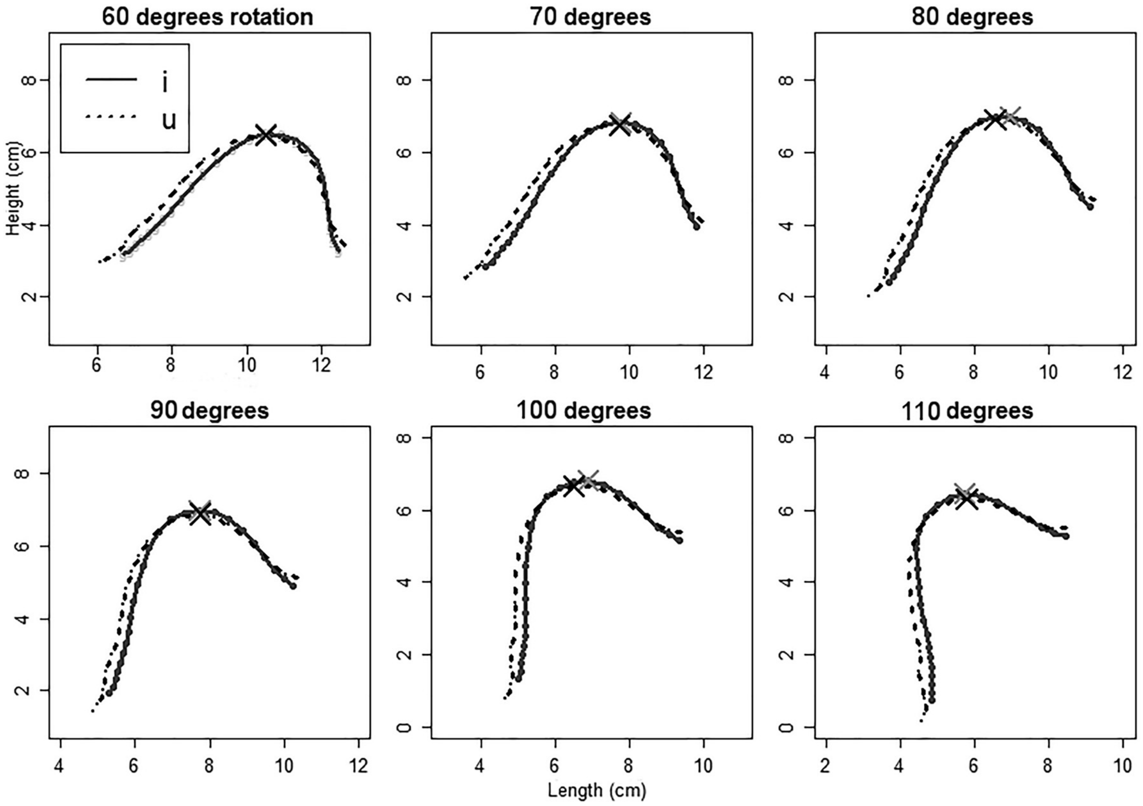

FIG. 3. Mean tongue surfaces for FLEECE (solid line) and GOOSE (broken line) vowels for speaker S4 Manchester female. In each panel, the two tongue surfaces are rotated at six different angles (between 60° and 110°) and the highest points of each tongue surface are automatically assigned and marked with crosses, grey ×, FLEECE; black ×, GOOSE.

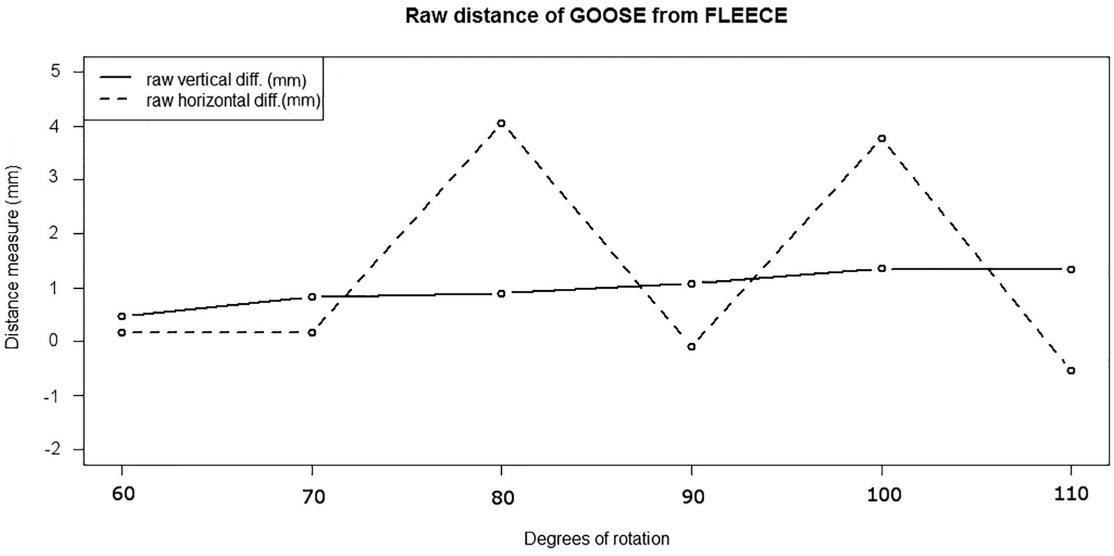

FIG. 4. Raw horizontal (broken line) and vertical (solid line) distances between the highest points of the mean FLEECE and GOOSE tongue curves of speaker S4 from Manchester, showing the impact of tongue surface rotation on raw horizontal and vertical distance measures.

rotated to a circa 90° probe-to-occlusal plane angle using the speaker's imaged occlusal plane.

F. Tongue distance measures

The highest point of the tongue has been an established descriptive tool in phonetics since the work by Jones (1909, 1917), based on earlier work by Bell (1887). Jones' vowel-description system was partly based on articulatory description (for cardinal vowels 1 and 5) and partly on the concept of auditory equidistance (Jones, 1947). Although a single-point measure of the highest point of the tongue traditionally has been considered an appropriate way of representing 2-D tongue-body movement, some researchers have suggested that this measure is less successful in capturing variation in the front-back dimension than variation in height, e.g., Ladefoged (1964); Lindau (1978). From the highest point of the tongue measures in Fig. 3, we can see that, in some rotations (80°,100°), the highest point of the tongue measure for FLEECE is fronter than that of GOOSE and, in others (60°,70°,90°,110°), their location is almost identical. Despite this variation in the highest point of the tongue measure, it is clear that, in all rotations, the tongue-body for GOOSE is *less* front than that of FLEECE, which will impact on the length of the resonating cavities. Highest point of the tongue measures do not capture variation in the position of the posterior part of the tongue surface.

In the current study, we use two measures: (1) a y axis measure taken from the highest point of the tongue to represent tongue-body height, and (2) an x axis measure from the back of the tongue, halfway up the pharyngeal cavity, to capture tongue-body frontness. This latter measure is taken halfway up the back of the tongue in order to avoid measurement of the position of the tongue root, which can move independently of the tongue-body. We suggest that these two measures are more likely to capture variation in tongue-body position that affects pharyngeal and oral cavity lengths and constriction locations.

G. Articulatory measurement and normalisation

1. Finding articulatory corner vowels across accents

Articulatory and acoustic measurements in this study were taken from words in the corpus that contained the GOOSE vowel (around 12 tokens per speaker). Following Scobbie *et al.* (2012), we measured GOOSE vowels relative to the FLEECE anchor vowel. To normalise the raw GOOSE-to-FLEECE articulatory measure, we expressed it as a proportion of the extent of the front-back and high-low vowel space using corner vowels: FLEECE (around 12 tokens) and TRAP (around 12 tokens). Finding a high-back corner vowel that worked for all varieties of British Isles English was difficult. The GOAT vowel works well for Scottish varieties of English, where GOAT is a monophthongal high-back vowel, but not for most other varieties, as the GOAT vowel in other varieties is often diphthongal and neither truly high nor back, e.g., [əʊ] in RP (Wells, 1982b), [əɪ]/[əʏ] for some young southern speakers (Kerswill and Williams, 2005), and [ɵː] in some northeastern English speakers (Watt and Milroy, 1999). For articulatory normalisation, we opted to use the semi-vowel [w] (mean, 5.6 tokens; s.d., 0.87) as the high-back corner vowel, as it was more stable and consistent across varieties and occupied a high-back position in the vowel space; see Sec. II H for information on how normalisation was carried out for acoustic measures.

2. Articulatory measurements from corner vowels and the GOOSE vowel

For each articulatory GOOSE vowel and corner vowel produced by each speaker, a single temporal midpoint was manually annotated during a steady state of the vowel, avoiding any initial diphthongal changes in the formants. We opted to use single-point measures in this study, after Scobbie *et al.* (2012), and also after Blackwood Ximenes *et al.* (2017), who took EMA measures of tongue coil positions at articulatory velocity minima (Blackwood Ximenes *et al.*, 2017, Sec. II D). In the future, we hope to consider these data using dynamic articulatory and acoustic measures.

Using AAA v2.16.12 (Wrench, 2012), a spline was fitted automatically to the midsagittal tongue surface in the scan-sequence image closest to the temporal annotation and hand-corrected where necessary. For the corner vowels FLEECE, TRAP, and [w], mean tongue splines were created for each speaker from multiple vowel tokens. Mean tongue surfaces were created by averaging the distances where individual splines intersect each of the 42 radial axes of the superimposed fan-shaped grid. Individual tongue surfaces were fitted and extracted for GOOSE vowels. Mean tongue surface splines of corner vowels, individual tongue-surface splines for GOOSE vowels, and occlusal-plane splines were transferred to a workspace in AAA for rotation where necessary; see Sec. II E. Thereafter, all splines were exported as sets of Cartesian coordinates for automatic measurement using *R* version 3.3.2 (R Core Team, 2018).

An *R* script automatically identified the y value of the highest point of the tongue's surface and the x value of a point halfway up the back of the tongue's surface for mean corner-vowel tongue contours and individual tongue surface contours of GOOSE vowels. The script also plotted the tongue surface contours and measurement points to allow eyeballing of measurement locations. Seven of the individual GOOSE-vowel tokens in total from four speakers were discarded after eyeballing as irregularities in the tongue-surface spline caused automatic measures to be taken from the wrong locations.

For each speaker, raw vertical and horizontal distances were obtained between the highest and backest points of the GOOSE tongue surface, relative to the same two measurement points on the mean FLEECE tongue surface. Proportional normalisation was carried out following Scobbie *et al.* (2012) by expressing the raw GOOSE-to-FLEECE measures as proportions of the full horizontal and vertical articulatory vowel space, based on corner-vowel measures (see Sec. I B), giving us two lingual articulatory dependent variables

Normalised tongue-body frontness:

$$\frac{\mathrm{FLEECEx} - \mathrm{GOOSEx}}{\mathrm{FLEECEx} - /\mathrm{w}/\mathrm{x}}, \tag{1}$$

Normalised tongue-body height:

$$\frac{\mathrm{FLEECEy} - \mathrm{GOOSEy}}{\mathrm{FLEECEy} - \mathrm{TRAPy}}. \tag{2}$$

Other researchers have used the *Z*-scoring normalisation method (Lobanov, 1971) for both articulatory and acoustic data, e.g., see Blackwood Ximenes *et al.* (2017), and we also present our articulatory and acoustic data with Lobanov normalisation for comparison; see Sec. III A.

H. Lip protrusion measurement

The use of a stabilising headset permitted use of a micro camera, located in a fixed position relative to the speaker's lips. The micro camera was mounted on a bracket protruding from the right side of the headset at a fixed distance from the midline of the speaker's head and enabled us to film lip movement in profile orientation. The camera recorded video in greyscale in National Television System Committee format, circa 29.97 frames per second. Lip video was synchronised with audio and UTI data using a SynchBrightUp unit (Articulate Instruments, Edinburgh, UK), which acts like a clapperboard, and through which the audio and video signals pass, adding a bright square to the video signal [see Fig. 5(b)] and a tone and pulses to the audio signal at the beginning of each recording. These signals are then aligned by AAA in a post processing stage, ensuring that all the video, audio, and ultrasound signals are aligned, and re-establishing the lip-video frame rate.

For the British Isles section of the *Dynamic Dialects* corpus, which was the first part of the corpus to be recorded, only profile lip video is available. Subsequently, a front-facing camera arm was designed and fitted to the stabilisation headset, permitting vertical and horizontal measures of the lip aperture. Analysis of lip-protrusion measures permits only a restricted comparison with acoustic data. Most vocal-tract modelling studies include a protrusion length (l) to aperture area (A) ratio in their models (Fant, 1992, Sec. 4; Stevens and House, 1955), capturing the effect of both protrusion and compression on acoustic impedance.

Using AAA, we first created a scaled horizontal fiducial line to act as a ruler along which protrusion could be measured; see Fig. 5(a). We then positioned the ruler fiducial so that it intersected the corner of the speaker's mouth. The position of the ruler remains constant in all subsequent video

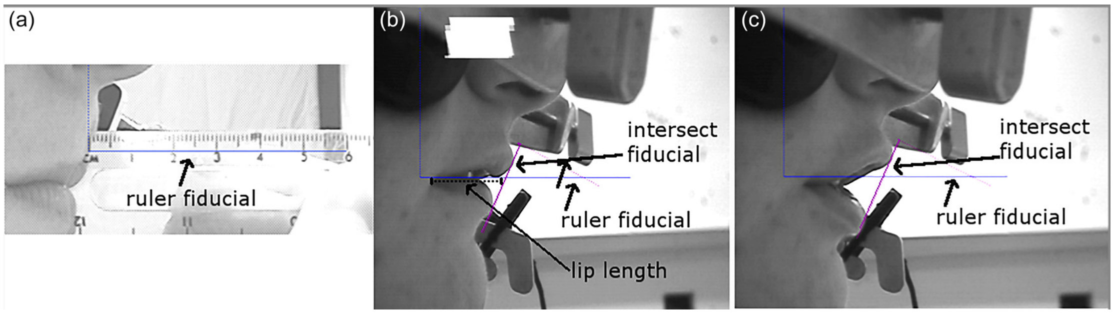

FIG. 5. (Color online) (a) Initial scaling of the lip-ruler fiducial using a physical ruler in the video frame, positioned in the middle of the philtrum. (b) Annotation of the "neutral" lip position before speech, using an intersect fiducial to measure the position of the lip edges relative to the lip ruler. (c) Annotation of the position of the lip edges relative to the lip ruler at the vowel "midpoint" using an intersect fiducial.

frames of each vowel token. We measured protrusion distances using a quasi-vertical fiducial, positioned to touch the edges of the upper and lower lips, set to intersect the ruler fiducial; see Figs. 5(b) and 5(c). Three measurements of lip protrusion were taken: one from a frame when the lips were in a neutral position before speech, another at the midpoint measure (where other articulatory and acoustic measures were taken), and another at the maximum point of lip protrusion during the GOOSE vowel segment; see Fig. 8(c). Normalisation was carried out by measuring lip length from one token in the recording when the lips were in a neutral position. Lip length was determined to be the distance from the corner of the mouth to the intersect fiducial when the lips were in a neutral position. All raw protrusion measurements were thereafter expressed as a proportion of this lip length measure,

$$\text{Normalised lip protrusion: } \frac{\text{mid\_prot} - \text{neutral\_prot}}{\text{neutral lip length}}. \tag{3}$$

Raw and normalised lip protrusion measures were highly correlated, $r_P = 0.97$, $p < 0.001$.

I. Acoustic measures and normalisation

Automatic acoustic measures of F1, F2, and F3 for GOOSE, FLEECE, and TRAP, respectively, were made using the Praat (version 6.0.23) Burg spectral analysis (Boersma and Weenink, 2013) with a 25 ms window length, 6 dB pre-emphasis above 50 Hz, assuming five formants per frequency range and adjusting the frequency range between 0 and 5 kHz (male) and 0 and 6 kHz (female). Mean F1 and F2 values were calculated for each vowel produced by each speaker. Instead of measuring [ɯ] as a corner vowel, which could have varying degrees of lip rounding that would affect formant measures, we followed the practice of Fabricius and Watt (2002) and used F1 of FLEECE as the F1 and F2 of a hypothetical high-back corner vowel u'. Where there is no suitable high-back corner vowel available for extrinsic normalisation processes that involve determining the extent of the acoustic vowel space, Fabricius and Watt's method establishes hypothetical lower limits of F1 and F2. The method assumes that the F1 of FLEECE is the minimum F1 of the acoustic space and assigns the same value to F2, as F2 cannot, by definition, have a lower value than F1.

Thereafter, the acoustic F1 and F2 distances of GOOSE from FLEECE were measured and proportionally normalised in the same way as the articulatory data to give us our two acoustic variables

$$\text{Normalised acoustic frontness: } \frac{\text{FLEECE}_{\text{F2}} - \text{GOOSE}_{\text{F2}}}{\text{FLEECE}_{\text{F2}} - u'_{\text{F2}}}, \tag{4}$$

$$\text{Normalised acoustic height: } \frac{\text{FLEECE}_{\text{F1}} - \text{GOOSE}_{\text{F1}}}{\text{FLEECE}_{\text{F1}} - \text{TRAP}_{\text{F1}}}. \tag{5}$$

Initially, we included an F3-F2 measure in our study, which was suggested to us as a potential acoustic correlate of lip rounding. F3-F2 is often used as an alternative measure of acoustic frontness (see Syrdal and Gopal, 1986); however, we did not find evidence that this measure captured variation in lip rounding. As already mentioned, vocal-tract modelling studies show that the lowering effect of lip rounding on F3 varies depending on other articulatory variables such as location of lingual constriction (Lindblom and Sundberg, 1971; Fant, 1992, p. 810). While a significant negative correlation was found between F3-F2, and normalised lip protrusion $r_P = -0.26$ $p < 0.05$, there was a much stronger correlation between normalised acoustic frontness and lip protrusion, $r_P = -0.35$, $p < 0.001$. Therefore, we did not consider F3-F2 to be a source of additional information on the effects of lip protrusion on the acoustics of the GOOSE vowel.

J. Lobanov normalisation

Proportional articulatory and acoustic normalisation, as described in Secs. II G and II I, involve the use of different high-back corner vowels: (i) the tongue-body position for [ɯ] for the articulatory data, and (ii) a hypothetical high-back corner vowel based on the F1 of /i/ for the acoustic data, which follows Fabricius and Watt (2002). This method was employed because it was felt that use of [ɯ] in acoustic normalisation would reduce comparability with articulatory measures where only the tongue-body was considered. However, this approach leaves open the possibility that statistically significant differences in acoustic and articulatory fronting between geographical regions in the study are attributable to different methods of normalisation of the articulatory and acoustic data. For this reason, following the methods of Blackwood Ximenes *et al.* (2017), we also present for comparison Lobanov-normalised (Lobanov, 1971) measures of the articulatory and acoustic data, using [ɯ] as a high-back corner vowel for both data types. Raw articulatory measures were tongue-body height and backness, as described in Sec. II G, and raw acoustic measures were F1 and F2. [u] tokens were Lobanov normalised for each speaker in *R* using measures from all individual tokens of [i], [a], [ɯ], and the norm.lobanov function of the vowels package (Kendall and Thomas, 2018). Variation in Lobanov-normalised measures is presented in Sec. III, alongside the proportionally normalised data, and is statistically analysed using mixed-effects modelling, as described in Sec. II K.

K. Statistical analysis

Mixed-effects modelling was carried out in *R* 3.3.2 (R Core Team, 2018). The following fixed factors were included in the models: (1) REGION with levels (i) English, (ii) Irish, and (iii) Scottish, and (2) SEX with levels (a) male and (b) female on the five dependent measures: (i) *normalised tongue-body height*, (ii) *normalised tongue-body frontness*, (iii) *normalised lip protrusion*, (iv) *normalised acoustic height*, and (v) *normalised acoustic frontness*. Lobanov-normalised articulatory and acoustic dependent variables were also analysed: (vi) *normalised tongue-body height*, (vii) *normalised tongue-body* *frontness*, (viii) *normalised acoustic height*, and (ix) *normalised acoustic frontness*. We did not test for interactions between the fixed factors. Random intercepts tested for all models were SPEAKER and PROMPT. Only random intercepts were tested, as testing of by-speaker random slopes for either fixed factor resulted in non-convergence. The step( ) function in the LmerTest package (Kuznetsova *et al.*, 2017) was used to find models that best fit the data. Both SPEAKER and PROMPT were found to be significant for dependent measures (i) *tongue-body height*, (iv) *acoustic height*, and (v) *acoustic frontness*, (vi) *Lobanov-normalised tongue body height*, (viii) *Lobanov-normalised acoustic height*, and (ix) *Lobanov-normalised acoustic frontness*. Only SPEAKER was significant for the dependent measure (iii) *lip protrusion*, and no random factors were significant for the dependent measures (ii) *tongue-body frontness* and (vii) *Lobanov-normalised tongue-body frontness*. For these two variables, therefore, the stepAIC( ) function of the MASS package (Venables and Ripley, 2002) was used to find linear models that best fit the data. The "lsmeans" package (Lenth, 2016) was used to carry out Tukey *post hoc* tests.

We also carried out Spearman's correlational analyses with Bonferroni corrections on all articulatory and acoustic measures, primarily in order to identify correlational relationships between articulatory parameters and acoustic output, although we report on all correlations.

III. RESULTS

In this section, we present the findings of the mixed-effects modelling concerning the effects of REGION and SEX on normalised articulatory and acoustic height and frontness, and lip protrusion. The effects of the fixed factors are illustrated using boxplots of normalised dependent measures and scatterplots of speaker means for the normalised tongue-body and acoustic measures. We then present the results of the correlational analysis.

A. Statistical analysis of variation in articulatory (tongue-body position) and acoustic space

Boxplots showing the effect of REGION on tongue-body and acoustic height and frontness based on proportionally normalised data are presented in Fig. 6. As all GOOSE-vowel-token measures were made relative to the FLEECE anchor vowel, the closer the measure values are to 0, the higher and fronter the GOOSE vowel token was in

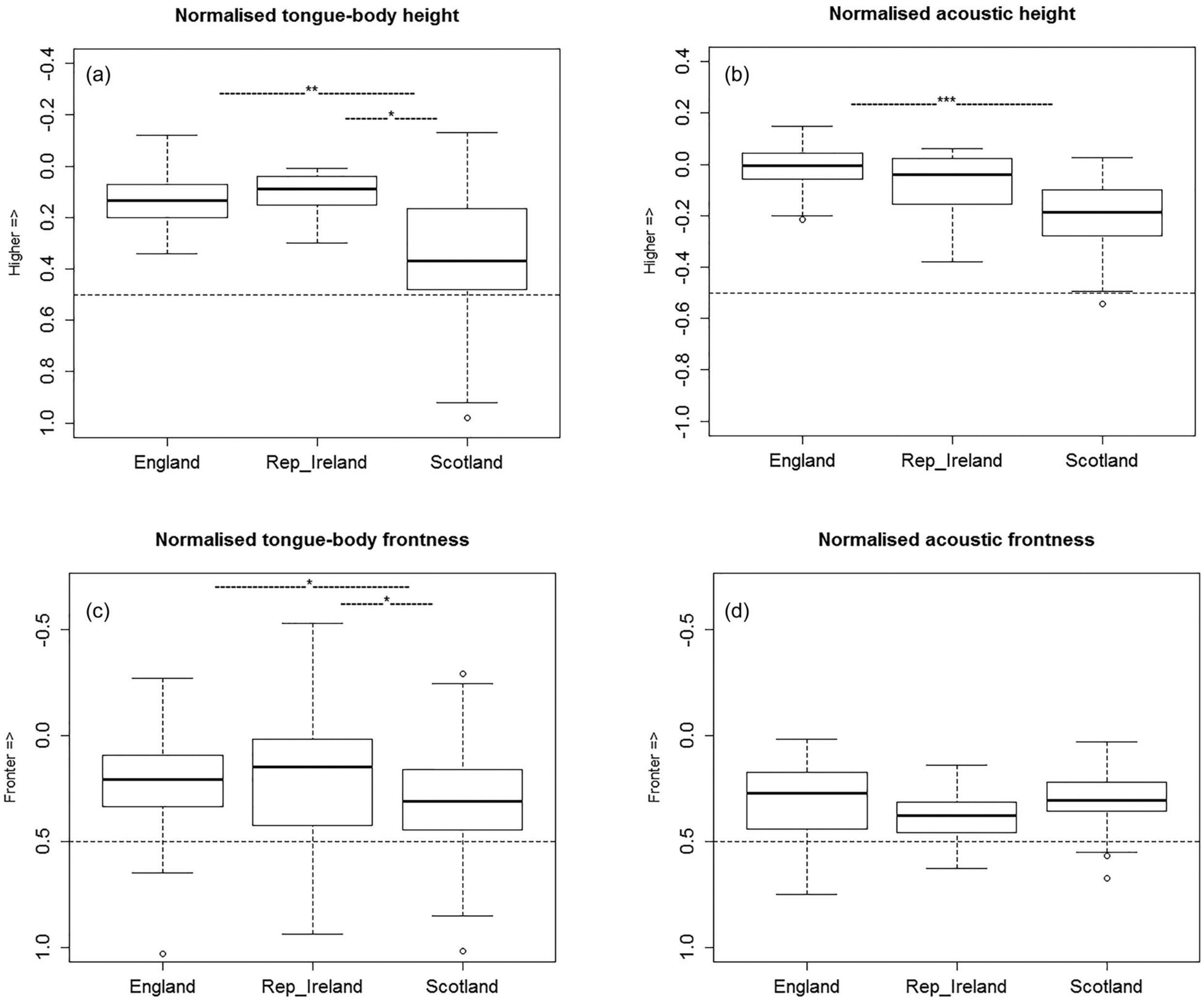

FIG. 6. Boxplots of (a) normalised tongue-body height, (b) normalised acoustic height, (c) normalised tongue-body frontness, and (d) normalised acoustic frontness for GOOSE vowel tokens, categorised by region. $N = 200$. Significant differences, marked on the figures using asterisks, relate to the outcome of the linear (mixed effects) regression analysis.

articulatory or acoustic space; therefore, the y axes in all but the acoustic height plot have been reversed to present the data more intuitively. Negative values in the *normalised tongue-body frontness* boxplots [Fig. 6(c)] occurred when the backest point of the GOOSE-vowel tongue surface was in a fronter position than the backest point of the mean FLEECE vowel tongue surface. Horizontal broken lines at 0.5 in Figs. 6(a), 6(c), and 6(d) and −0.5 in Fig. 6(b) represent the midpoints of the front-back and high-low dimensions of the articulatory and acoustic vowel spaces. Statistically significant variation between regions is marked with asterisks and based on the results of the mixed-effects modelling.

Figures 7(a) and 7(b) present mean articulatory (tongue-body) and acoustic height and frontness values based on proportionally normalised data for each speaker as points on a 2-D scatterplot, labelled by geographical location. Again, the axes have been reversed in these scatterplots in order to present results in a familiar way, similar to the commonly used F2-by-F1 plot. Again, broken lines represent the midpoints in the horizontal and vertical dimensions of articulatory/acoustic space. Y axes on the articulatory and acoustic scatterplots have the same normalised scale; however, there is a slight difference in the scaling of the x axes of each plot in order to avoid crowding the data in the articulatory plot.

Lobanov-normalised mean articulatory (tongue-body) and acoustic height and frontness values are presented for comparison in Figs. 8(a) and 8(b). For the Lobanov-normalised data, both the acoustic and articulatory plots have the same scaling.

1. Comparison of tongue-body and acoustic height

Boxplots of tongue-body and acoustic height in Figs. 6(a) and 6(b), respectively, show that English and Irish GOOSE vowels can be described as high vowels in both acoustic space and in terms of tongue-body height. This is not the case for all Scottish GOOSE vowels as has already been shown by Scobbie *et al.* (2012) and Stuart-Smith *et al.* (2017).

The final model for *tongue-body height* showed that the fixed factor REGION was significant, $F(2,18) = 9.97$, $p = 0.0012$. *Post hoc* Tukey tests showed that there were significant differences between the regional accent groups: England and Scotland $t(18) = 3.71$, $p = 0.0036$, and Ireland and Scotland $t(18) = 2.985$, $p = 0.0186$. For the Lobanov-normalised data, REGION was also significant, $F(2,18) = 6.41$, $p = 0.0078$. *Post hoc* Tukey tests showed that there were significant differences between the regional accent groups: England and Scotland $t(18) = 3.41$, $p < 0.0083$, only.

The final model for *acoustic height* showed that REGION was also significant, $F(2,18) = 20.184$, $p < 0.001$; *post hoc* Tukey tests showing significant differences between England and Scotland, $t(18) = 5.62$, $p < 0.0001$ only, with Scottish speakers' GOOSE vowels located significantly lower in acoustic space than English speakers' GOOSE vowels; see Fig. 6(b). The fixed factor SEX was also significant, $F(1,18) = 6.29$, $p = 0.0218$, with female speakers' GOOSE vowels located significantly higher in acoustic space than those of the male speakers. For the Lobanov-normalised data, REGION was significant, $F(2,17) = 18.595$, $p < 0.001$. *Post hoc* Tukey tests showed that there were significant differences between the regional accent groups: England and Scotland $t(17) = 5.839$, $p = 0.0001$, and between Ireland and Scotland, $t(17) = 3.09$, $p = 0.0171$, with Scottish speakers having acoustically lower GOOSE vowels than both English and Irish speakers. There was also a significant effect of SEX, $F(1,17) = 6.331$, $p = 0.0221$, with female speakers' GOOSE vowels located significantly higher in acoustic space than those of the male speakers.

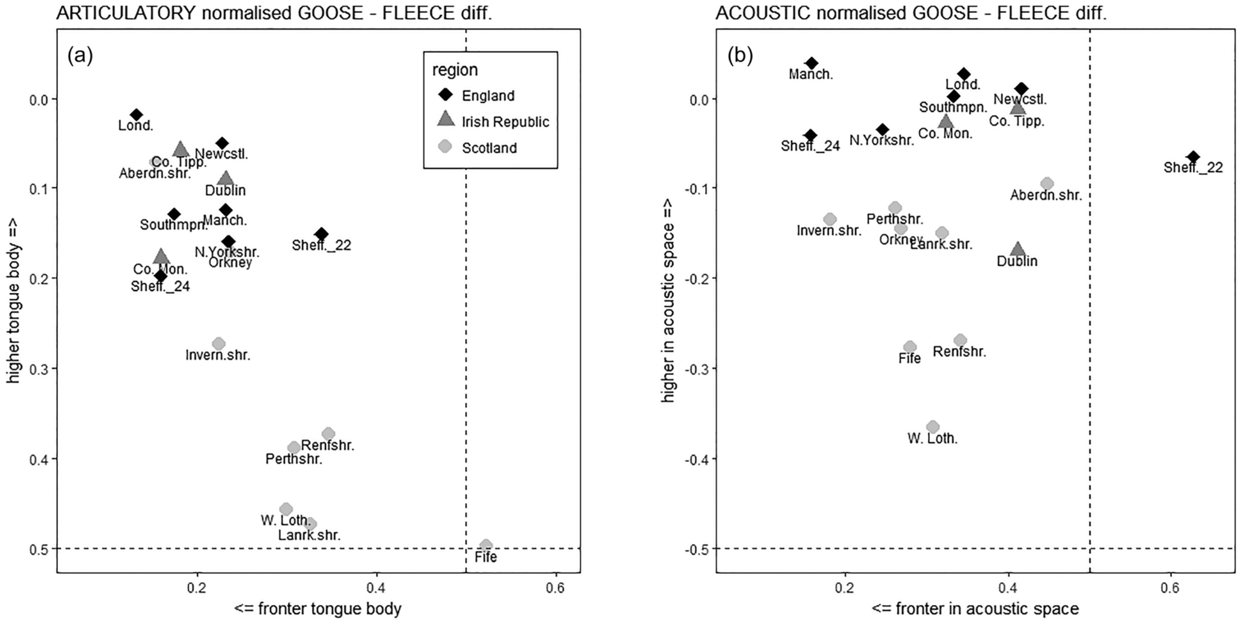

FIG. 7. Scatterplots of speaker means for (a) normalised tongue-body height and frontness and (b) normalised acoustic height and frontness for the GOOSE vowel. Some axis scales have been reversed in order to present these data in an F2-by-F1 plot style. The closer the datapoint is to the top left corner of the plot, the higher and fronter the mean GOOSE vowel produced by the speaker is. Broken lines indicate the midpoints of the articulatory and acoustic vowel spaces. Speaker labels indicate the location where speakers have lived longest.

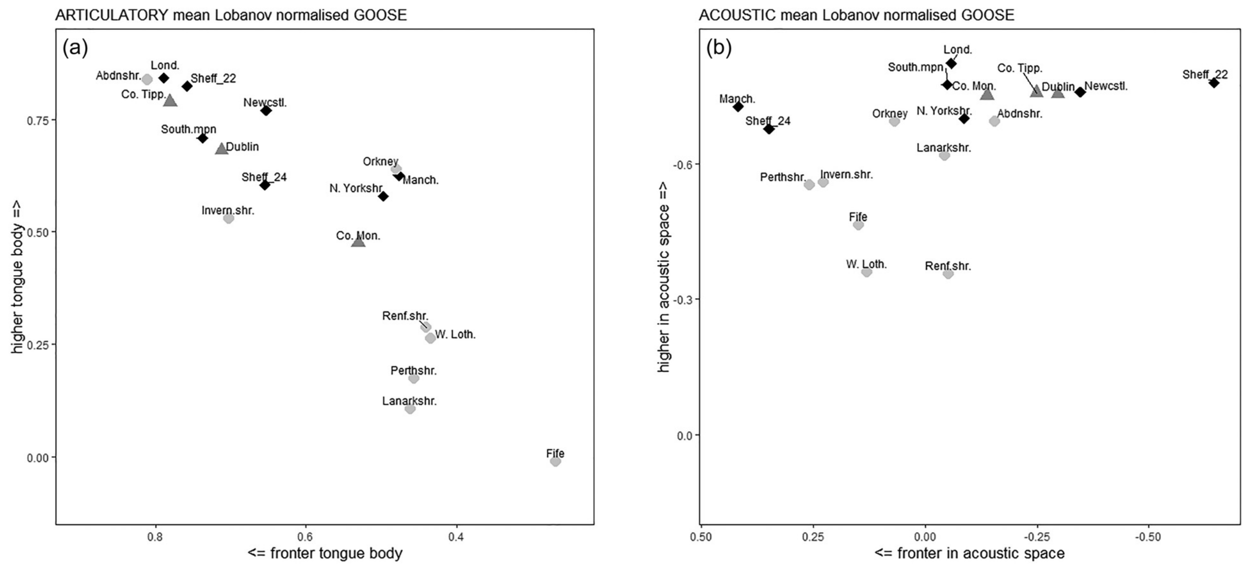

FIG. 8. Scatterplots of speaker means for (a) Lobanov-normalised tongue-body height and frontness and (b) Lobanov-normalised acoustic height and frontness for the GOOSE vowel. Speaker labels indicate the location where speakers have lived longest. Speakers' regional identities are coded by datapoint shape and colour as shown in the key for Fig. 7.

Inspection of Figs. 7(a) and 8(a) reveals that the comparatively large spread of tongue-body height values for the Scottish regional group can be attributed to subregional phonetic variation; Northern Scottish speakers (Aberdeen, Inverness-shire, and Orkney) have higher mean tongue-body positions, while Scottish Central Belt speakers (Fife, S. Lanarkshire, W. Lothian, and Renfrewshire) have more central tongue-body positions, close to the vertical midline of the articulatory vowel space. Additionally, a large tongue-body-height range appears to map onto a more restricted acoustic height range.

2. Comparison of tongue-body and acoustic frontness

Boxplots of tongue-body and acoustic frontness in Figs. 6(c) and 6(d) show that truly back tokens of the GOOSE vowel are rare in this dataset; the majority of tokens occur beyond the midline of the horizontal vowel space. In Figs. 7(a) and 7(b), tongue-body and acoustic frontness measures also place the majority of the mean GOOSE vowel datapoints in the central-to-front region of the horizontal vowels space.

The final linear model for *tongue-body frontness* showed the fixed factor REGION was significant, $F(2,18) = 4.43$, $p = 0.0131$, with *post hoc* tests showing significant variation between England and Scotland, $t(18) = 2.481$, $p = 0.0369$, and Ireland and Scotland, $t(18) = 2.405$, $p = 0.0449$. In both cases, Scottish speakers had significantly backer tongue-body positions for the GOOSE vowel. The Lobanov-normalised data for tongue-body frontness also show that the fixed factor REGION was significant, $F(2,18) = 6.409$, $p = 0.0079$, with *post hoc* tests showing significant variation between England and Scotland, $t(18) = 3.412$, $p = 0.0083$, and, again, Scottish speakers had significantly backer tongue-body positions for the GOOSE vowel than English speakers.

However, in the final model for *acoustic frontness*, no fixed factors were significant. This was also the case for the Lobanov-normalised data. Figures 7(a) and 8(a) show subregional phonetic variation in the Scottish group with Northern Scots having fronter tongue-body positions for GOOSE than Central Scots. This plot suggests that the GOOSE vowel in some Central Scottish speech is now a lax vowel.

Figures 7(a) and 8(a) perhaps show most clearly that the GOOSE vowel for Central Scottish English speakers is articulated with a centralised tongue-body position, while acoustic analysis, based on F1 and F2 measures in Figs. 7(b) and 8(b), shows only that this vowel is lowered compared with those of speakers from other regions.

3. Comparison of tongue-body and acoustic position

Based on the proportionally normalised data, Fig. 9 shows differences in mean GOOSE vowel position between normalised articulatory space (tongue-body position) and normalised acoustic space, and arrows move from articulatory to acoustic. Speakers from each region are presented separately to avoid crowding.

Figure 9 shows that, across all three regions, almost all speakers' GOOSE vowels are higher in normalised acoustic space than in normalised articulatory space (based on tongue-body position). This difference is most pronounced for the Scottish speakers, many of whom produce GOOSE with a central tongue-body position. Scobbie *et al.* (2012) found a similar mismatch between tongue-body height and Bark-transformed F1 measures, whereby the GOOSE vowel was found to be higher in acoustic space than in articulatory space.

For the front-back dimensions in articulatory and acoustic space, we see regional patterns of variation. English and Irish speakers have backer GOOSE values in normalised acoustic space than in normalised articulatory (tongue-body position) space, while Scottish speakers show fronting or little difference in front-back position between articulatory and acoustic space. Figure 9 suggests that another articulatory

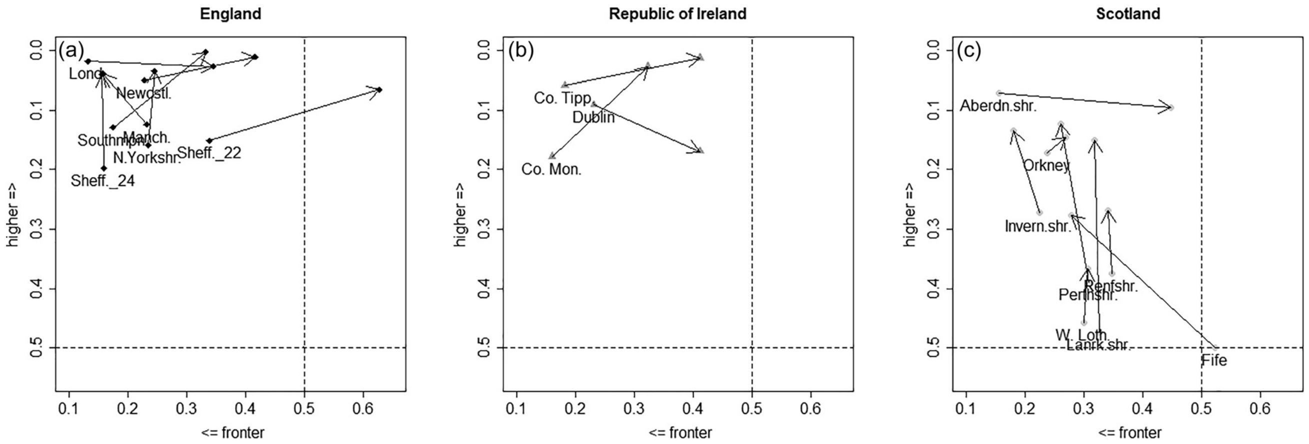

FIG. 9. 2-D plots comparing normalised tongue-body positions of the GOOSE vowel with its position in normalised acoustic space with separate plots for each region of the British Isles. Arrows show the direction of difference from articulatory (tongue-body position) to acoustic.

parameter is affecting F2 measures for the GOOSE vowel. Below, we consider regional variation in the lip protrusion measure.

B. Statistical analysis of regional variation in lip protrusion

Figure 10 below shows boxplots of normalised lip protrusion from measures taken at the same time point as tongue-body and acoustic measures. Maximum lip protrusion during the GOOSE vowel was also measured and found to be highly correlated to the midpoint lip protrusion measure, $r_P = 0.98$, $p < 0.001$, and temporal measurement locations of maximum lip protrusion were similar to those of the midpoint measures, as was found in Mayr (2010).

The final mixed-effects model for lip protrusion showed a significant effect for REGION, $F(2,17) = 3.65$, $p = 0.048$, with *post hoc* tests showing significant variation between England and Scotland with only $t(17) = 2.619$, $p = 0.0443$. Figure 11 shows individual speakers' normalised mean lip protrusions for GOOSE vowels with standard deviations. The bars are colour-coded by region and ordered from the lowest degree of normalised lip protrusion to the greatest degree of lip protrusion.

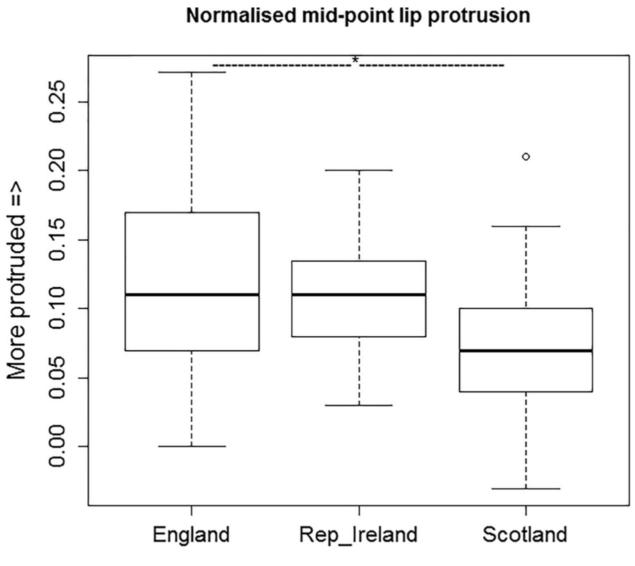

FIG. 10. Boxplots of normalised lip protrusion from the GOOSE vowel "midpoint" measure, organised by region of the British Isles $N = 187$. The significant difference, marked on the figure using an asterisk, relates to the outcome of the linear mixed effects regression analysis.

Scottish speakers tended to have smaller degrees of lip protrusion for GOOSE, perhaps confirming the persistence of exolabial lip rounding in GOOSE, as described in McAllister's articulatory-phonetic account of Central Scottish English (McAllister, 1938); see Sec. 1 A. However, particularly for Central Belt Scottish speakers, lip positions often looked neutral rather than exhibiting exolabial rounding. One English speaker, Sheffield_24 exhibited low levels of lip protrusion. Interestingly, he is one of the few English speakers to show no fronting difference between articulatory and acoustic space in Fig. 9(a). However, as mentioned in Sec. II H, we have only lip protrusion data and lack information about lip aperture area, so we cannot fully model the relationship between lip position and acoustics for individuals.

Speakers with the greatest and smallest degrees of lip protrusion were from Newcastle (England) and Fife (Scotland), respectively. S9_Fife's production of "room" resulted in a slightly negative normalised value of −0.03 (−0.5 mm raw measure), while S7_Newcastle's production of "shoe" resulted in a large positive normalised value of 0.27 (7 mm raw measure). Although both measures were taken during the vowel, doubtless, the secondary labial articulation on [ʃ] had a coarticulatory effect that emphasised lip protrusion in the vowel of "shoe."[4] For four tokens of the GOOSE vowel, S9_Fife had a lip protrusion maximum that was marginally less protruded than his neutral lip position. These negative values were too small to indicate lip spreading, but certainly indicate a lack of any kind of lip protrusion in these tokens of the GOOSE vowel.

C. Correlational analysis of articulatory and acoustic measures

Spearman's correlation tests with Bonferroni corrections were undertaken for all dependent measures for proportionally normalised data, and taken at the same temporal-point of each vowel token: (1) *proportionally normalised tongue-body height*, (2) *proportionally normalised tongue-body*

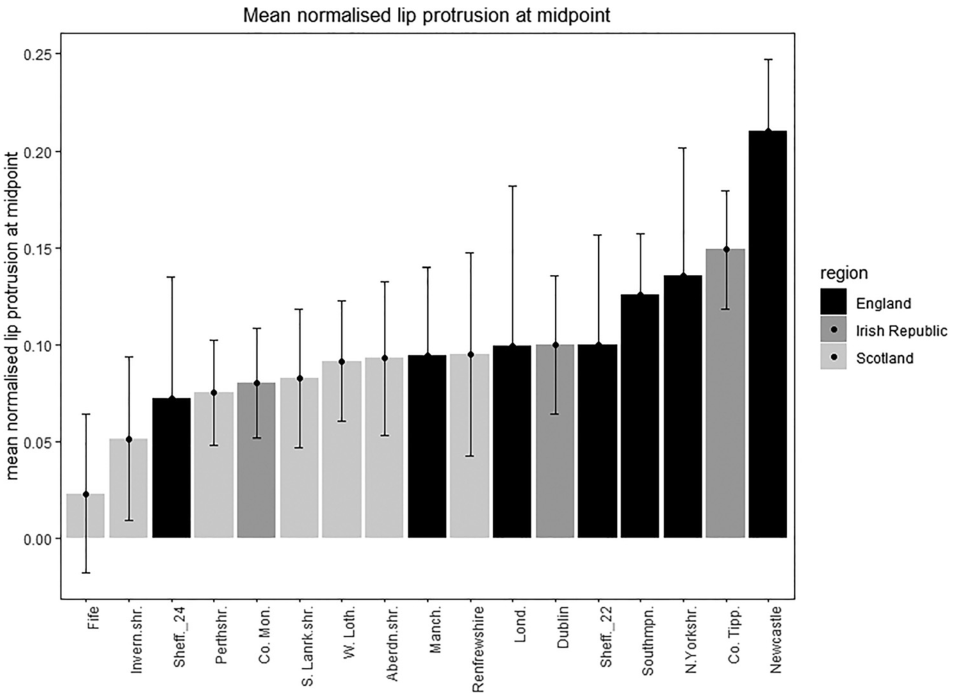

FIG. 11. Barplot of normalised mean lip protrusion in the GOOSE vowel, labelled by speaker and coloured by region of the British Isles, $N = 187$.

frontness, (3) *normalised lip protrusion*, (4) *normalised acoustic height*, and (5) *normalised acoustic frontness*. Measures were transformed before carrying out the correlational analysis to improve interpretability. Up until this point, axes have been reversed on plots in order to present articulatory and acoustic measures in a more conventional and intuitive way. By reversing axes, we have also set an expectation of how the data should be interpreted. We anticipated that performing a correlational analysis on the untransformed data would cause the reader some difficulty in interpreting the direction of some correlations; we therefore decided to transform the data as follows so that the direction of correlations would be easier to interpret. Normalised tongue-body frontness and height values and normalised acoustic frontness values were multiplied by -1 so that the greater the value, the higher or fronter the tongue-body or the fronter the GOOSE vowel in acoustic space. Acoustic height and lip protrusion were left untransformed and the higher their values, the higher the GOOSE vowel in acoustic space and the more protruded the lips. Transformation of the data does not affect the r_P values or significance levels, only whether the r_P value was positive or negative.

Table II presents the results of the correlational analysis with r_S values and asterisks indicating levels of significance. The four strongest correlations are plotted in Fig. 12.

Some articulatory measures were found to correlate with one another, and there were also correlations between articulatory and acoustic measures.

For articulatory measures, we see positive correlations between normalised tongue-body height and normalised tongue-body frontness, $r_S = 0.40$, $p < 0.001$ [see Fig. 12(a)], and between normalised tongue-body height and normalised lip protrusion, $r_S = 0.30$, $p < 0.001$ [see Fig. 12(b)]. As can be seen from the regional datapoint coding in Figs. 12(a) and 12(b), these correlations reflect regional performative variation in the dataset. English and Irish speakers who have higher tongue-body positions also tend to have fronter tongue-body positions and more protruded lips for GOOSE vowel productions, while Scottish speakers have lower and backer tongue-body positions and lesser degrees of lip protrusion.

For articulatory and acoustic measures, we see that the strongest correlation was a positive one, normalised tongue-body height and normalised acoustic height $r_S = 0.55$, $p < 0.001$ [see Fig. 12(c)], and there was no significant correlation between normalised tongue-body frontness and normalised acoustic frontness. There was, however, a significant negative correlation between normalised lip protrusion and normalised acoustic frontness, $r_S = -0.35$, $p < 0.001$ [see Fig. 12(d)], and greater lip protrusion is associated with a reduction in acoustic frontness. Two further significant positive correlations between articulatory and acoustic measures are probably attributable to the regional performative variation described above,with normalised acoustic height and normalised lip protrusion, $r_S = 0.37$,

TABLE II. Correlation matrix for all articulatory and acoustic measures, showing r_S values and levels of significance using asterisks. **, $p < 0.01$; ***, $p < 0.001$.

	Normalised tongue-body frontness	Normalised lip protrusion	Normalised acoustic height	Normalised acoustic frontness
Normalised tongue-body height	0.40***	0.30***	0.55***	−0.11
Normalised tongue-body frontness		0.11	0.28***	0.18
Normalised lip protrusion			0.37***	−0.35***
Normalised acoustic height				0.03

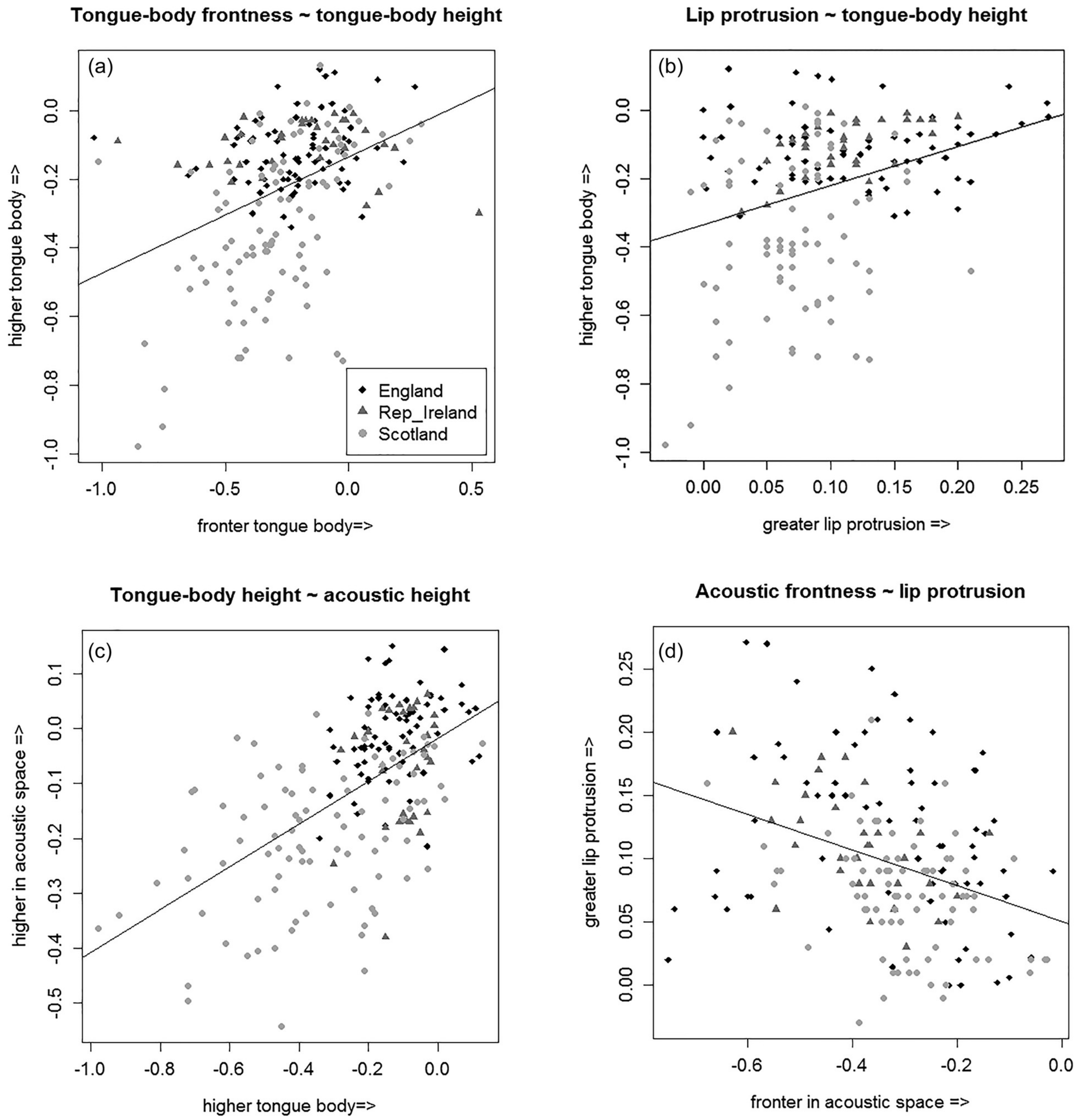

FIG. 12. Correlation scatterplots with regression lines (a) normalised tongue-body frontness and normalised tongue-body height, (b) normalised lip protrusion and normalised tongue-body height, (c) normalised tongue-body height and normalised acoustic height, and (d) normalised acoustic frontness and normalised lip protrusion. Datapoints from each region in the British Isles are coded by shape and colour; see (a).

$p < 0.001$,[5] and normalised acoustic height and normalised tongue-body frontness, $r_S = 0.28$, $p < 0.001$. Their correlation plots are not shown as they closely resemble Figs. 12(a) and 12(b).

IV. DISCUSSION

Acoustic analysis of vowel variation is, to date, much quicker and more suitable for large quantities of data than articulatory analysis; however, articulatory analysis is worth undertaking if we want to avoid overlooking significant social or diatopic variation, which is not easily recoverable from the acoustic signal due to motor equivalence (Blackwood Ximenes *et al.*, 2017; Lawson *et al.*, 2011, 2015). For the GOOSE vowel, in particular, it is possible to achieve similar normalised F2 values for vowels using different articulatory strategies, e.g., backer tongue-body position with lesser degrees of lip rounding, or fronter tongue position with greater degrees of lip rounding (Harrington *et al.*, 2008; Stevens and House, 1955; Savariaux *et al.*, 1995; Lindblom and Sundberg, 1971).

In the British Isles, we have a situation where similar normalised F2 values are reported in present-day studies for markedly different varieties. It might be assumed that similarity in the frontness of the GOOSE vowel across the British Isles is the result of accent levelling or sound-change diffusion; however, we know that the fronting of GOOSE in Scottish English results, in part, from a 13th century fronting process, affecting the antecedents of the GOOSE lexical set in northern dialects of the British Isles (Johnston, 1997). Additionally, there are phonetic descriptions from the 1930s of performative variation between Central Scottish and Southern English /u(ː)/ (McAllister, 1938) noting systematic variation in lip posture.

The present study investigated the relationship between the articulatory parameters of tongue-body height and frontness, lip position, and acoustic height and frontness based on F1 and F2 measures and aimed to answer the research questions:

(1) How can we standardise and normalise inter-speaker tongue-body measures recorded with UTI and lip measures recorded with the profile lip camera?
(2) Do we see regional patterns of articulatory variation that are distinct from regional patterns of acoustic variation?
(3) What correlations do we find between articulatory measures of tongue-body height and frontness and lip protrusion and acoustic F1, F2 measures?

In answer to research question (1), we have suggested a preliminary methodology to help minimise inter-speaker variation during UTI recording by introducing a method of probe-to-cranium standardisation (the bite plate) and suggesting normalisation methods for tongue and lip measurements, particularly addressing the issue of missing corner vowels for articulatory data.

In answer to research question (2), articulatory measures show significant regional variation across all three parameters measured: tongue-body height, tongue-body frontness, and lip protrusion, but regional variation is only apparent in acoustic height (F1) measures. Irish and English speakers, on one hand, and Scottish speakers, on the other hand, use different production strategies involving tongue and lip positions that result in similar acoustic frontness (F2) measures. While Irish and English speakers have fronter tongue-body positions, they also use greater degrees of lip protrusion. Scottish speakers, particularly those from Central Scotland, have backer, technically more centralised, tongue-body positions and weakly protruded or neutral lip positions. These different production strategies cannot be considered to be an example of trading relations (Perkell *et al.*, 1993), where a stable acoustic target is achieved using different articulatory strategies, as the sound changes that resulted in fronted GOOSE in each regional variety occurred centuries apart, and the differences in acoustic height between English/Irish GOOSE and Scottish GOOSE mean that they remain auditorily distinct.

In answer to question (3), correlation tests confirmed that while there was a strong positive correlation between tongue-body height and acoustic height in our data and a negative correlation between lip protrusion and acoustic frontness, there was no significant correlation between tongue-body frontness and acoustic frontness. We suggest that the lack of correlation between tongue-body position and acoustic frontness is due to the impact of lip protrusion on the acoustics of this vowel.

Our study confirms the persistence of older Scottish GOOSE-vowel phonetic variants, identified over 80 years ago (McAllister, 1938), although it seems that some Central Scottish speakers might be using a neutral lip posture rather than an exolabial one today. Our study confirms the articulatory findings of Scobbie *et al.* (2012), namely that some Central Belt Scottish speakers produce GOOSE with a centralised tongue-body position, although extreme tongue-body lowering does not result in equally extreme lowering in acoustic space. In addition, we have shown that this lowering is likely to be a feature of Central Belt speech, rather than Scottish speech, in general, as speakers from northern Scotland were found to have higher, fronter tongue-body positions. This intra-regional variation supports the real-time acoustic results of Stuart-Smith *et al.* (2017), whose acoustic study of variation in the BOOT vowel in the Central Belt Scottish city of Glasgow over the last century identified a gradual diachronic process of lowering (Stuart-Smith *et al.*, 2017).

Our study also supports the findings of Harrington *et al.* (2011) in showing that Anglo-English speakers tend to produce GOOSE with high-front tongue-body positions, and they preserve lip protrusion. Our study confirms that other strategies, e.g., backer tongue-body positions with smaller degrees of lip protrusion, can be used to produce acoustically fronted GOOSE.

Regarding previous work on Irish speech, our study supports the findings of Hickey (2016) that GOOSE fronting is evident in young female Irish speech, although the Irish speakers in the present study had GOOSE vowels closer to the middle than the front of acoustic space.

V. CONCLUSION

This study contributes to articulatory-acoustic mapping by presenting novel methods to help address issues of variability in recording settings and inter-speaker variation in articulatory analysis of vowels. Using these methods, we have shown that the GOOSE vowel can be performatively different in British Isles English varieties, where GOOSE fronting results from different sound-change histories. In short, we have shown that not all GOOSE fronting is "the same"; similar degrees of acoustic fronting can be achieved by tongue-body fronting or a reduction in lip protrusion. Our articulatory study represents a step toward an improved understanding of articulatory variation in different regional vowel systems as well as increasing our understanding of the relationship between articulation and acoustics. In the future, methods set out in this paper could be employed in the study of a larger UTI-based accent dataset in order to characterise vocalic variation from an articulatory perspective with the addition of information about lip aperture ratio in order to model the impact of lip position on acoustics in more detail.

ACKNOWLEDGMENTS

This work was supported by the Arts and Humanities Research Council, U.K., Grant No. AH/L010380/1, and an award from the Carnegie Trust for the universities of Scotland, "An online ultrasound tongue imaging resource for phonetics, linguistics, and speech therapy teaching at Scottish universities."

[1] www.dynamicdialects.ac.uk (Last viewed 11 April 2019).

[2] In Scottish English, we could more accurately use the keywords FOOT/GOOSE (sometimes researchers use the keyword BOOT; Stuart-Smith *et al.*, 2017), as there is no phonemic /ʊ/-/u/ split in Scottish English (Wells, 1982b,a).

[3] In Scots vernacular and some dialectal northern English, the MOUTH lexical set are also produced with monopthongal /u(ː)/, preserving an older Anglian vowel quality, e.g., *mouth* [mʉθ/mʉf]; *house* [hʉs], *cow* [kʰʉː] (Chirrey, 1999; Corbett and Stuart-Smith, 2012; Stuart-Smith, 1999; Watt and Milroy, 1999; Wells, 1982b). However, in the current paper, all participants spoke forms of Standard English; therefore, there are no instances of monophthongal /u(ː)/ for MOUTH vowels.

[4] While [ɹ] often has tertiary lip rounding in English, our observation is that lip rounding on [ɹ] in Scottish English is not as common as in other varieties.

[5] Increased lip protrusion lowers F1 (Stevens and House, 1955), increasing acoustic height, which may also contribute to the significant correlation between these measures.

Adank, P., Smits, R., and Van Hout, R. (**2004**). "A comparison of vowel normalization procedures for language variation research," J. Acoust. Soc. Am. **116**(5), 3099–3207.

Aitken, A. J. (**1981**). "The Scottish vowel length rule," in *So Meny People Longages and Tonges: Philological Essays in Scots and Mediaeval English Presented to Angus McIntosh*, edited by M. Benskin and M. S. Samuels (Michael Benskin & M. L. Samuels, Edinburgh), pp. 131–157.

Bell, A. M. (**1887**). *The Elocutionary Manual: The Principles of Elocution with Exercises and Notations for Pronunciation, Intonation, Emphasis, Gesture and Emotional Expression*, 5th ed. (Parker, Washington, D.C.).

Bigi, B. (**2015**). "SPPAS—Multi-lingual approaches to the automatic annotation of speech," The Phonetician **111-112**, 54–69.

Blackwood Ximenes, A., Shaw, J. A., and Carignan, C. (**2017**). "A comparison of acoustic and articulatory methods for analyzing vowel differences across dialects: Data from American and Australian English," J. Acoust. Soc. Am. **142**(1), 363–377.

Bladon, R. A. W., and Fant, G. (**1978**). "A two-formant model and the cardinal vowels," Royal Institute of Technology, Speech Transmission Laboratory, Stockholm Quarterly Progress and Status Report, Vol. 19(1), pp. 1–8.

Boberg, C. (**2011**). "Reshaping the vowel system: An index of phonetic innovation," Univ. Pa. Work. Pap. Linguist. **17**(2), 20–29.

Boersma, P., and Weenink, D. (**2013**). "Praat: Doing phonetics by computer," (version 6.0.23) [computer software], http://www.praat.org/ (Last viewed 11 April 2019).

Chirrey, D. (**1999**). "Edinburgh: Descriptive material," in *Urban Voices: Accent Studies in the British Isles*, edited by P. Foulkes, and G. Docherty (Arnold, London), Chap. 12, pp. 223–230.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. (**1952**). "Some Experiments on the Perception of Synthetic Speech Sounds," J. Acoust. Soc. Am. **24**(6), 597–606.

Corbett, J., and Stuart-Smith, J. (**2012**). "Standard English in Scotland," in *Standards of English: Codified Varieties around the World*, edited by R. Hickey (Cambridge University Press, Cambridge, UK), Chap. 4, pp. 72–95.

Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J. (**1952**). "An experimental study of the acoustic determinants of vowel color; observations on one- and two- formant vowels synthesized from spectrographic patterns," Word **8**, 195–210.

Fabricius, A. (**2007**). "Vowel formants and angle measurements in diachronic sociophonetic studies: FOOT-fronting in RP," in *Proceedings of the 16th International Congress of Phonetic Sciences*, 6–10 August, Saarbrucken, Germany, pp. 1477–1480.

Fabricius, A., and Watt, D. (**2002**). "Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the $F1$–$F2$ plane," Leeds Work. Pap. Linguist. Phonetics **9**(9), 159–173.

Fant, G. (**1992**). "Vocal tract area functions of Swedish vowels and a new three-parameter model," in *Proceedings of the 2nd International Conference on Spoken Language Processing*, 13–16 October, Banff, Alberta, Canada, pp. 807–810.

Ferragne, E., and Pellegrino, F. (**2010**). "Formant frequencies of vowels in 13 accents of British English," J. Int. Phon. Assoc. **40**(1), 1–34.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., and Trudgill, P. (**2004**). *New Zealand English: Its Origins and Evolution* (Cambridge University Press, Cambridge, UK).

Harrington, J., Kleber, F., and Reubold, U. (**2008**). "Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study," J. Acoust. Soc. Am. **123**(5), 2825–2835.

Harrington, J., Kleber, F., and and Reubold, U. (**2011**). "The contributions of the lips and tongue to the diachronic fronting of high back vowels in standard southern British English," J. Int. Phon. Assoc. **41**, 137–156.

Hawkins, S., and Midgley, J. (**2005**). "Formant frequencies of RP monophthongs in four age groups of speakers," J. Int. Phon. Assoc. **35**(2), 183–199.

Hickey, R. (ed.) (**2016**). "English in Ireland: Development and varieties," in *Sociolinguistics in Ireland* (Palgrage MacMillan, Basingstoke), pp. 3–41.

Johnston, P. A. (**1997**). "Older Scots phonology and its regional variation," in *The Edinburgh History of the Scots Language*, edited by C. Jones (Edinburgh University Press, Edinburgh, UK), pp. 47–111.

Jones, D. (**1909**). *The Pronunciation of English* (CUP, Cambridge, UK).

Jones, D. (**1917**). *An English Pronouncing Dictionary*. (Dent, London, UK).

Jones, D. (**1947**). *An Outline of English Phonetics* (Heffer, Cambridge, UK).

Joos, M. (**1948**). "Acoustic phonetics," Language **24**, 1–136.

Kendall, T., and Thomas, E. R. (**2018**). "Vowels: Vowel manipulation, normalization, and plotting," in *R Package (version 1.2-2) [computer software]* (R Foundation for Statistical Computing, Vienna, Austria).

Kerswill, P., and Williams, A. (**2005**). "New towns and Koineization: Linguistic and social correlates," Linguistics **43**(5), 1023–1048.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2017**). "lmerTest Package: Tests in linear mixed effects models," J. Stat. Softw. **82**(13), 1–26.

Labov, W., Rosenfelder, I., and Fruehwald, J. (**2013**). "One hundred years of sound change in Philadelphia: Linear incrementation, reversal and reanalysis," Language **89**(1), 30–65.

Ladefoged, P. (**1964**). *A Phonetic Study of West African Languages* (Cambridge University Press, Cambridge).

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (**2015**). "The role of anterior lingual gesture delay in coda /r/ lenition: An ultrasound tongue imaging study," in *Proceedings of the 18th International Congress of Phonetic Sciences*, 10–14 August 2015, Glasgow, UK.

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (**2014**). "A socio-articulatory study of Scottish rhoticity," in *Sociolinguistics in Scotland*, edited by R. Lawson (Palgrave Macmillan, London, UK), pp. 53–78.

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (**2011**). "The social stratification of tongue shape for postvocalic /r/ in Scottish English," J. Socioling. **15**(2), 256–268.

Lawson, E., Stuart-Smith, J., and Scobbie, J. M. (**2018**). "The role of gesture delay in coda /r/ weakening: An articulatory, auditory and acoustic study," J. Acoust. Soc. Am. **143**, 1646–1657.

Lee, J., Shaiman, S., and Weismer, G. (**2016**). "Relationship between tongue positions and formant frequencies in female speakers," J. Acoust. Soc. Am. **139**(1), 426–440.

Lenth, R. (**2016**). "Least-squares means: The R package lsmeans," J. Stat. Softw. **69**(1), 1–33.

Lindau, M. (**1978**). "Vowel features," Language **54**(3), 541–563.

Lindblom, B., and Sundberg, J. (**1971**). "Acoustic consequences of articulatory movement," J. Acoust. Soc. Am. **50**(4), 1166–1179.

Lobanov, B. M. (**1971**). "Classification of Russian vowels spoken by different speakers," J. Acoust. Soc. Am. **49**(2), 606–608.

Mather, J. Y., and Speitel, H. H. (**1986**). *The Linguistic Atlas of Scotland. Vol III. Phonology* (Croom Helm, London).

Mayr, R. (**2010**). "What exactly is a front rounded vowel? An acoustic and articulatory investigation of the nurse vowel in South Wales English," J. Int. Phon. Assoc. **40**(1), 93–112.

McAllister, A. (**1938**). *A Year's Course in Speech Training* (University of London Press Ltd., London, UK).

McGurk, H., and MacDonald, J. (**1976**). "Hearing lips and seeing voices," Nature **264**, 746–748.

Mesthrie, R. (**2010**). "Socio-phonetics and social change: Deracialisation of the GOOSE vowel in South African English," J. Socioling. **14**(1), 3–33.

Mielke, J., Baker, A., Archangeli, D., and Racy, S. (**2005**). "Palatron: A technique for aligning ultrasound images of the tongue and palate," in *Coyote Papers: Working Papers in Linguistics*, edited by S. Jackson and D. Siddiqi (University of Arizona, Tucson, AZ).

Narayanan, S., Bresch, E., Ghosh, P., Goldstein, L., Katsamanis, A., Kim, Y., Lammert, A., Proctor, M., Ramanrayanan, V., and Zhu, Y. (**2011**). "A multimodal real-time MRI articulatory corpus for speech research," in *Proceedings* *of the 12th Conference of the International Speech Communication Association*, 27–31 August, 2011, Florence, Italy, pp. 837–840.

Orton, H., Sanderson, S., and Widdowson, J. (**1978**). *The Linguistic Atlas of England* (Croom Helm, London, UK).

Perkell, J. S., Matthies, M. L., Svirsky, M. A., and Jordan, M. I. (**1993**). "Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot 'motor equivalence' study," J. Acoust. Soc. Am. **93**(5), 2948–2961.

Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**(2), 175–184.

Przedlacka, J. (**2001**). "Estuary English and RP: Some recent findings," Studia Anglica Posnaniensia **36**, 35–50.

R Core Team. (**2018**). "*R*: A language and environment for statistical computing (version 3.3.2) [computer software]" (R Foundation for Statistical Computing, Vienna, Austria).

Savariaux, C., Perrier, P., and Orliaguet, J. P. (**1995**). "Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube: A study of the control space in speech production," J. Acoust. Soc. Am. **98**(5), 2428–2442.

Scobbie, J. M., and Pouplier, M. (**2010**). "The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/," J. Phonetics **38**(2), 240–259.

Scobbie, J. M., Stuart-Smith, J., and Lawson, E. (**2012**). "Back to front: A socially-stratified ultrasound tongue imaging study of Scottish English /u/," Riv. Linguist./Ital. J. Linguist., Special Issue: Articulatory Techniques for Sociophonetic Research **1**, 103–148.

Scobbie, J. M., and Wrench, A. A. (**2003**). "An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 3–9 August 2003, pp. 1871–1874.

Scobbie, J. M., Wrench, A. A., and van der Linden, M. (**2008**). "Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, pp. 373–376.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

Stevens, K. N., and House, A. S. (**1955**). "Development of a quantitative description of vowel articulation," J. Acoust. Soc. Am. **27**(3), 484–493.

Stone, M. (**2005**). "A guide to analysing tongue motion from ultrasound images," Clin. Linguist. Phonet. **19**(6-7), 455–501.

Stuart-Smith, J. (**1999**). "Glasgow: Accent and voice quality," in *Urban Voices*, edited by P. Foulkes and G. Docherty (Arnold, London, UK), pp. 203–222.

Stuart-Smith, J., Jose, B., Rathcke, T., Macdonald, R., and Lawson, E. (**2017**). "Changing sounds in a changing city: An acoustic phonetic investigation of real-time change over a century of Glaswegian," in *Language and a Sense of Place: Studies in Language and Region*, edited by C. Montgomery and E. Moore (Cambridge University Press, Cambridge, UK), pp. 38–64.

Syrdal, A. K., and Gopal, H. S. (**1986**). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," J. Acoust. Soc. Am. **79**(4), 1086–1100.

Venables, W. N., and Ripley, B. D. (**2002**). *Modern Applied Statistics with S*, 4th ed. (Springer, New York).

Watt, D., and Milroy, L. (**1999**). "Patterns of variation and change in three Newcastle vowels: Is this dialect levelling?," in *Urban Voices*, edited by P. Foulkes and G. Docherty (Arnold, London, UK), Chap. 2, pp. 25–46.

Wells, J. C. (**1982a**). *Accents of English 1: Introduction* (Cambridge University Press, Cambridge, UK).

Wells, J. C. (**1982b**). *Accents of English 2: The British Isles* (Cambridge University Press, Cambridge, UK).

Wrench, A. (**2012**). *Articulate Assistant Advanced User Guide* (Articulate Instruments Ltd., Edinburgh, UK).

Wrench, A., and Hardcastle, W. (**2000**). "A multichannel articulatory speech database and its application for automatic speech recognition," in *Proceedings of the 5th Seminar on Speech Production*, May 1–4, Kloster Seeon, Germany, pp. 305–308.

Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Holland, C., and Choe, A. (**2008**). "A magnetic resonance imaging-based articulatory and acoustic study of 'retroflex' and 'bunched' American English /r/," J. Acoust. Soc. Am. **123**(6), 4466–4481.