Back to front: a socially-stratified ultrasound tongue imaging study of Scottish English /u/

James M. Scobbie, Jane Stuart-Smith & Eleanor Lawson

We explore the vowel space, with a particular focus on the phonetic location (and phonological interpretation) of the vowel /u/ (GOOSE, FOOT) in Scottish accented English, using a socially-stratified articulatory and acoustic corpus of fifteen teenage speakers of both sexes (ECB08). The articulatory data consists of midsagittal tongue contours extracted from ultrasound tongue images, and the acoustic vowel space is modelled with F1 and F2 (in Bark). We explore the methodological issue of how to quickly measure a given vowel's lingual location relative to others in a space of such tongue curves, given a very small sample from each speaker: we measure the linear distance between the highest point of different tongue curves. This lets us compare the somewhat metaphorical but widely accepted equivalence of frontness in acoustic space to high F2, to a more literal but still indirect measure of frontness, namely the relative closeness of the high points of /u/ and /i/. /e/ and /o/ are also measured, for comparison. We investigate two quite different rotations of the space, which reflect different hypotheses about what is an appropriate orientation for the horizontal axis. Both rotations give similar results, supporting our qualitative analysis of /u/, but we recommend the use of the occlusal plane to define horizontality (preferably measured directly on a speaker-by-speaker basis).

In Scottish English, the /u/ vowel has previously been described as approximately central between cardinal vowel 1 and cardinal vowel 8, and high. Our qualitative analysis, supported by our acoustic and articulatory measurement, revises this finding: in these speakers /u/ is indeed front of centre but, however, it is not high – it is in fact a frontish, mid-high (rounded) vowel. Phonologically, a number of interpretations would be available, all of which alter the shape of the Scottish English system by accepting that "/u/" is not high and back. Moreover, /o/ is far closer to cardinal 8 in both acoustic and articulatory senses, and its location in phonetic space strongly suggests that it is (now) the high back vowel of the phonological system. Alternatively, there is no such vowel. We also find articulatory support for the very strong sociophonetic difference in the location of the /ɪ/ and /ɛ/ vowels between more working class and more middle class speakers, and discuss whether this clear phonetic difference should be modelled phonologically.

Rivista di Linguistica 24.1 (2012), pp. 103-148 *(received February 2012)*

We set this discussion in the context of a critique of strongly modular approaches to the phonetics/phonology interface. We ask whether phonological labels are in any way relevant to explaining phonological change, or, equivalently, whether the phonetic measurement of a vowel category provides straightforward evidence of what the phonological label actually is. We argue that it would only be possible to find unequivocal evidence for phonological change after the fact, and that labels are likely to be assigned non-deterministically and related to phonetics in an abstract way. We therefore conclude that the use of phonetically realistic labels for phonological categories does not have any straightforward explanatory purpose, unlike the number of categories and their location in phonetic space, which, however, does not require modularity.*

1. Introduction

1.1. What is front about a front vowel?

In descriptive phonetics, and hence in phonology and in sociolinguistics, there is a longstanding tradition of classifying the quality of vowels in terms of a range of dimensions such as the height, rounding and frontness of each vowel. These dimensions are represented in the vertical and horizontal dimensions of the cardinal vowel space of Daniel Jones's primary and secondary vowel set, as well as being reflected in many other systems of classification and transcription. They permeate thinking about the vowel space, not just informally, or in transcription using International Phonetic Alphabet vowel symbols and diacritics, but also when the vowel space is quantified in formal acoustic dimensions based on the measurement and estimation of its first two acoustic resonances; the first and second formants. The relation between the use and representation of the first and second formants as key correlates of vowel quality and the cardinal vowel space[1] has been frequent topic of study since Joos (1948) (e.g., Iivonen 1994).

There is an articulatory basis to the height, rounding and frontness dimensions – or at the very least there is an articulatory *metaphor*. There is obviously also an articulatory foundation to vowels, in speech planning and production, based in part in the location of the tongue in the vocal tract. So how do these various levels and systems relate, in fact and in metaphor? After all, the tongue is a supremely flexible articulator capable of producing a complex vocal tract geometry. Is anything to be gained by examining vowel targets in articulatory space as opposed to just acoustic space, and how do the two perspectives relate? It is clear that it is not as simple as an F1 = height

and F2 = frontness metaphor implies, and not just because the formant frequencies are modified by other articulatory components, most notably the lips.

For example, a so-called high vowel is generally expected to have a lingual constriction near the hard or soft palate, with a relatively closely approximated mandible and maxilla. A front vowel will be generally expected to carry, in very general terms, the body of the tongue closer to the lower incisors (if low) or the post-alveolar surface (if high) than a corresponding back vowel. Pharyngeal articulations are also present, of course, but the traditional labels very much refer to the anterior portion of the vocal tract, with the tongue root position being an (optional) extra and usually unspecified in traditional vowel description. But the acoustic characteristics of vowels' resonances – specifically the frequency of the formants – are based on the entire articulatory space, including the contribution of the lips and tongue root, not just the anterior lingual location.

Daniel Jones apparently believed that there was an articulatory, as well as an auditory (and acoustic) basis to the whole vowel space, and three corner cardinal vowels, [i], [ɑ], and [u] have an articulatory definition, with the others being auditorily equidistant between them (Jones 1918, 1972). For example, cardinal vowel [u] is produced with the tongue as far back and as high in the mouth as possible, and with tight lip rounding. A diagrammatic representation of the vowel space in two dimensions (height and frontness) has always been used, and the locations of the vowels were thought to be isomorphic to the location of the highest point of the tongue in physical space, as shown in the X-ray photographs of Jones by H. Trevelyan George (Jones 1917a).[2] The outer limit of the vowel space was quickly regularised into a quadrilateral or triangular shape for pedagogical reasons, but originally the extreme high points / vowel locations along the outside of the possible space for vowel production were thought to more accurately form an ovoid (Jones 1918). The orientation of front and high was, however, only informally fixed. There was, it would seem, only a general "whole body" definition of horizontal and vertical in these X-ray photographs, parallel to, we assume, the edges of the photographic images as reproduced within their rectangular frames. The occlusal plane (as defined by the fitting of a flat plate against the upper teeth) is, nowadays, used as an intra-oral definition of horizontality in speech research (Scobbie et al. 2011). We estimate (from skeletal features) that the occlusal plane is in these x-rays is sloping downwards (towards the anterior) by approximately 6°.

The schematic vocal tract vowel diagrams in Jones (1972: 32) are based on the X-ray photographs of Jones' production of [i], [a], [ɑ], [u], augmented with four other estimated primary cardinal vowel curves.[3] These diagrammatic tongue curves are each labelled with a single point, the highest point on the curve of the tongue in the images. These high points are then related to the relative locations of each cardinal vowel in the vowel space.

While current articulatory research focuses on three-dimensional tube geometries, kinematics, and other sophisticated analytic techniques far beyond anything possible 100 years ago, nevertheless the concepts of the front-back and high-low tongue positions pervade our thinking about vowels.

1.2. Phonological features

The relationship between the phonetic quality of a vowel and its phonological location in the system inventory can be complicated, even if an attempt is made to use congruent analytic dimensions, leading on occasion to the rejection of any such connection between the linguistic levels (e.g., Hale & Reiss 2000). Phonologically, theoretical drivers such as markedness, diachronic and cross-dialectal homogeneity (or faithfulness), systematic symmetry and phonetic grounding support the use of certain phonetically-inspired labels for phonemes (or any segment-like phonological category) in what appears to be a phonetic featural space. Actually any phonological category exists in a rather abstract transferral of phonetic space into the phonological level of description. Features are *not* phonetic. In consequence, a set of "similar" phonemic inventory member categories (e.g., across dialects) with relatively "similar" phonetic qualities will be labelled with the *same* IPA symbol or the same set of distinctive features, despite clear observable phonetic differences (Docherty 1992, Pierrehumbert et al. 2000) and despite, on occasion, quite dramatic discrepancy between any straightforward interpretation of the phonological label and the phonetic facts.

For example, despite phonemic differences in inventories due to diachronic changes, the vowel in the GOOSE lexical set in English is generally labelled /u/ and defined abstractly as some theoretical description cognate to [HIGH, BACK, ROUND]. But what can /u/ actually sound like? Or not? There is an almost complete lack of theoretical understanding of the circumstances in which the inevitable tension between phonetics and phonology is too untenable to be defensible. So, is the vowel of the English GOOSE lexical set, for example, really a high, back and round vowel in phonology, call it /u/;

or is it a central /ʉ/; or is it something else? How can we tell? And, in the context of diachronic change, what theoretical predictions are there about when the threshold of phonological reanalysis will be crossed? How far in phonetic space from [u] must /u/ be, for /u/ to be unsupportable as an analytic label?[4] In Ohala's (1981) conception of change, when does the listener-acquirer change category?

This paper addresses English /u/ as a specific instance of this phonetic-phonological tension. The English high back vowel (or, more commonly, vowels, since in many varieties of English there is a tense and lax pair, namely GOOSE /u/ and FOOT /ʊ/) is an important sociolinguistic variable in a number of accents of English, which has received attention in recent literature as being "fronted" phonetically, relative both to its high back phonological categorisation, and to historic phonetic realisations. Fronting can also affect other back rounded vowels, which might be moving in a chain, but we will focus here on /u/.[5]

1.3. Phonetic investigations of /u/ *in English*

A variety of methodological approaches can be used to study phonetic vowel spaces. Ferragne & Pellegrino (2010) present a broad cross-dialectal empirical acoustic survey of various British English vowel systems, including both Scottish and Anglo-English systems. Wells (1982) also provides a comprehensive description of all significant English varieties, relying more heavily on transcription. Lass (1989) is a useful phonological discussion of front rounded vowels in English, including /u/. Mesthrie (2010) adopts a socially-stratified Labovian methodology to explore /u/ fronting in South African English. To focus just on acoustic analyses, the place of /u/ in a number of the national English dialects, in addition to South African English, has been examined, for example: American and Canadian (Labov et al. 2006, Fridland 2008, Boberg 2011); Australian (Cox 1999, Cox & Palethorpe 2001); New Zealand (Gordon et al. 2004). Southern Standard British English has been addressed by Henton (1983), Hawkins & Midgley (2005), Fabricius (2007), Harrington et al. (2008) and McDougall & Nolan (2007) among others; and of course Harrington (2007) explores the relationship between British English and the Queen's own accent. Indeed, GOOSE-fronting is so ubiquitous that it would be unsurprising if it were examined in every recent variationist study undertaken in the UK, as well being a very important theoretical driver for research in many other geographical locations, including many of the acoustic studies mentioned above.

While acoustic measurement is now standard in linguistic studies of accent variation, direct measurement of articulation is still relatively uncommon in phonetics. Harrington et al. (2011b) includes a detailed articulatory study of /u/ in SSBE, and we will return to it below. Articulatory analysis in sociophonetics is, moreover, positively rare. Our particular focus here is a dialect in which the /u/ vowel has been phonetically central, rather than back, for quite some time, namely Scottish English, since we think it makes an interesting counterpoint to other varieties of English in which diachronic change seems to be more active, currently. Specifically, we will look at Eastern Central Belt Scottish Standard English using a socially-stratified sample, and we will employ both acoustic and articulatory analyses.

1.4. /u/ *in Scottish English*

Scottish English[6] (as well as more vernacular accents, not addressed here) has long been said to have a fronted (i.e., central or front-central) /u/ (McAllister 1938, Speitel & Johnston 1983, Macaulay 1977, Johnston 1997, Schüztler 2011). Stuart-Smith's 1997 Glasgow data (Stuart-Smith 1999) analysed acoustically by Scobbie et al. (1999b) shows quantitatively that /u/ is central. Preliminary results from acoustic analysis of Macaulay's 1972/3 data also confirm /u /is a non-back vowel (Stuart-Smith et al. 2012). We thus will assume that there is an even stronger basis in Scottish English for "back" /u/ to be re-analysed by linguists and/or speakers as central or front than in accents where variation in greater and back realisations might be regarded prescriptively as more acceptable.

Ferragne & Pellegrino (2010) recently confirmed that Glasgow and Elgin /u/ are high and central, using data from the ABI corpus, and they confirm our own anecdotal observations that it is the /o/ vowel that is high and back in the acoustic and impressionistic vowel spaces, "to the extent that *hoed* in [Elgin and Glasgow] probably has the vowel that comes closest to cardinal [u] in the whole corpus" (ibid: 24). Brato (2012: 114) notes that in the Aberdeen area, local, distinct forms of FOOT and GOOSE "range from a fully back [u] to an extremely fronted [y]" and in general the locality shows fronting and merging, so that they both "have been levelled towards a more SSE-like [ʉ]" (ibid.).

There are two metaphors at play here – the term "front" and the symbol "u" which is used for the phoneme in the GOOSE set of words in English generally, and additionally for FOOT in Scottish English, which lacks the relevant tense/lax contrast.[7] The phonemic symbol in /u/ is inevitably associated with the phonetic cardinal vowel #8, [u],

which in IPA terms, recall, is the highest, backest, roundest vowel. Acoustically, cardinal #8 can be expected to have a very low F1 and F2. Since the symbol for Anglo-English GOOSE has traditionally been [u], which is associated with a high, back tongue position and rounded lips, Anglo-English GOOSE is today said to be fronted or centralised, because, impressionistically, it has been observed to have a fronter quality, and/or it has been found to have a relatively higher F2 in acoustic analysis than cardinal [u]. Since within a speech community, different groups of speakers show different F2 values, the extent of /u/ fronting has been found to carry social meaning (e.g., Baranowski 2008, Cheshire et al. 2011). Finally, previous generations of Anglo-English speakers had an /u/ vowel which was backer: diachronic fronting is supported by the real-time study of Harrington (2007) and apparent time studies such as Cheshire et al. (2011), as well as comparisons between modern Southern Standard British English speakers (Hawkins & Midgley 2005, Ferragne & Pellegrino 2010) and descriptions of and primers for traditional RP (the erstwhile standard) based on the seminal works and revisions of Jones (1917, 1918) and Gimson (1962), and the more contemporary and comparative Wells (1982).

Discussion of a diachronic movement forwards in the vowel space is based on acoustic measures showing higher F2 values in more recent times or (preferably normalised) in younger speakers, as well as by impressionistic transcriptions in the IPA cardinal vowel space. Presumably, therefore, at some point phonologists will also propose that these varieties' GOOSE and FOOT vowels have undergone phonological change and are no longer analysable as being phonologically BACK.

1.5. Articulation, acoustics and front /u/

There has been some recent work quantifying how impressionistic conceptions of frontness and height are grounded in speech production. Such work can help us address basic questions concerning the relationship between phonological and phonetic conceptions of linguistic categories. What might "fronting" of /u/ mean at an articulatory level, specifically in terms of lingual articulation? Does "front" mean the same thing across languages or dialects? How do various aspects of articulation interact with each other to create audiovisual cues to perceived frontness, such as a loss of lip rounding which results in an increase in F2, but is visible to a perceiver?

Harrington et al. (2011b, 2008), for example, provided articulatory answers to these questions for Southern Standard British English

(SSBE) by examining both lip rounding and tongue position. They found evidence in favour of their hypothesis that diachronic fronting was initiated when /u/ followed coronal consonants, due to coarticulatory pressures on /u/ from the front articulation required for the coronal consonant. The degree of lingual articulatory fronting in SSBE has, it seems, been underestimated, due in part to suggestions that unrounding might have been partly responsible for raising of F2.

Harrington et al.'s (2011b) articulatory study of five speakers aged in their early 20s concluded, from Electromagnetic Articulometry (EMA) data, that SSBE /u/ is indeed much more front than central, along an articulatory horizontal dimension defined on the occlusal plane of the speaker, which, as noted above, is a standard reference collected as part of EMA experiments. As is common in EMA studies, data on the position and movement of three coils (hence three flesh points) on the upper anterior surface of the tongue was analysed. The coils were TT (about 1cm behind the tongue tip), TM (tongue mid) and TB (tongue back, roughly opposite the velum). These showed the tongue target configuration is similar in the three phonemes /i/, /u/ and /ɪ/, i.e., that /u/ is fronted, which was supported by a principl components analysis in which /u/ and /ɪ/ patterned together, close to but distinct from /i/. Based on the highest flesh point shown (the more posterior "TB-tongue body" coil for 4 speakers and TM-tongue mid for one) (Harrington et al. 2011b: Fig. 8), we can estimate that /u/ is on average about 2mm retracted from /i/, and 2.5mm lowered. From video data used in a cross-linguistic perceptual (mis)-classification experiment as well as from EMA data, SSBE /u/ was shown to be lip rounded. The quantitative EMA data alone, being based just on a lower lip EMA coil, rather than, say, cross-sectional area of lip aperture, is not, however, a particularly convincing demonstration that /u/ has not unrounded to some extent, but the audiovisual perceptual and acoustic analyses all combine to support their conclusion that rounding is a key component of /u/.

One of the strong advantages of EMA is the wide range of established techniques for comparing and discriminating articulations, as well as its excellent spatio-temporal sampling abilities and its suitability for long data collection sessions. Statistical analysis techniques are also highly developed. However, the tongue coils only provide information on three or four anterior points on the tongue spanning about 5cm, and do not give a convenient qualitative overview. In particular, information on tongue root position is lacking. Moreover, EMA is technically complex, and it is hard to obtain data from large numbers of speakers as are often required for sociophonetic investigations,

particularly if the goal is to obtain qualitative data or rather broad quantitative data. So we will turn now in the next section to a much more accessible technology for qualitative analysis, which in the next decade is likely to see an enormous growth of use in phonetic analysis. As we will see, however, there is a pressing need for the development of new tools for detailed quantitative analysis, in order to emulate what has been achieved with EMA, and, more relevant for this paper, there is even a lot of work to do to agree on qualitative methods of interpretation.

1.6. Ultrasound tongue imaging

Most ultrasound-based phonetics research is undertaken using the video output of standard medical ultrasound machines with a standard audio signal from an independent microphone. For an overview of the general principles and some general articulatory methodological considerations, see Davidson (2012). For a specific discussion of how articulatory methods, especially ultrasound, have and can be applied in sociolinguistic research, see Lawson et al. (2008, submitted) and Stuart-Smith et al. (submitted).

There is a huge variety of scanners that can sample roughly fan-shaped, two-dimensional areas of the vocal tract via probes that are held against the underside of the chin (the submental surface). They are able to create images of vocal tract tissues lying within their field of view, so long as bone or a body of air does not lie between the probe and the object of interest. A small probe pressed against the submental surface emits a series of echo-pulse beams in the mid-sagittal plane in a fan (with an angle of around 120°), detecting the echoes. Since a major source of echoes is the tongue-air boundary, ultrasound scanners create an image of a sagittal slice of the tongue interior[8] and surface from the tip, or near the tip when there is a sublingual airspace, to near the root. As will be discussed, the orientation and location of the tongue within this image varies from speaker to speaker, depending on physiology and the placement and angling of the probe.

Ultrasound scanners perform many scans per second, but for linguistic analysis, the images they create have to be compiled, sequenced, and then synchronised with the audio speech signal as part of digitisation on a computer. The standard output from ultrasound scanners is a video signal at a rate of approximately 30 frames (images) per second (fps), corresponding more-or-less to the standard NTSC television standard rate of 29.97fps. In practice, the actual frame rate should be measured electronically, because it may easily vary by ±1fps in dif-

ferent scanners, and may vary as the depth, power, and scan rate settings are changed. It is unlikely in video output that each video frame corresponds to exactly one complete scan from the probe: usually the image in each frame is a composite created from a number of such sweeps, and data from different parts of a single scan might appear in different images, introducing problems with temporal synchronisation, ambiguities and artefacts (Wrench & Scobbie 2006). (More problematic is a scanning rate lower than about 30Hz, which fails to provide new images for each video frame, but even so, slow-moving articulations can be analysed). A fast scan rate is the first part of a solution, to make sure each sweep is completed within a short window. The second is to de-interface each image, if, as seems to typically be the case, the video frames are themselves interlaced during their creation by the scanner. If the images are interlaced, buffered scan data is fed to the odd and then the even horizontal pixel lines in sequential blocks. If the initial scan data is greater than twice the frame rate, e.g., approximately 100fps, de-interlacing doubles the effective frame rate to 60fps, albeit with a consequent halving of image resolution vertically. This reduces, but does not eliminate temporal overlap, the doubling of data between frames, and temporal smearing, but it is a big improvement if temporal re-alignment is carried out correctly.

Ultimately, the solution will be cineloop or high-speed UTI (Hueber et al. 2008, Wrench & Scobbie 2008, 2011, Miller & Finch 2011), so long as synchrony between ultrasound scans and the acoustic channel can be ensured. At faster rates, temporal resolution, accuracy and synchronisation can be better, but at very high rates, the density of echopulses or size of the area being imaged is reduced. If there is a greater distance between echo pulse lines, more extensive interpolation between echo pulse lines is required, giving a more smeared quality to the ultrasound images, more-so further from the probe. The systems and processes needed for high-speed UTI are currently more specialised, expensive, complex and/or laboratory-based than video UTI, and for the foreseeable future, sociolinguistic research is more likely to use the video-based systems.

Stabilisation of the probe to the head of the speaker is a further important consideration. Some approaches rely on immobilising the speaker's head, e.g., against a chair or similar supporting device (Stone & Davis 2005, Gick et al. 2005, Davidson 2012). If the probe is secured in the mid-sagittal plane and its movements within that plane relative to the head can be tracked, then correction of those movements can be made (Whalen et al. 2005, Mielke et al. 2005). However, if synchronisation is not perfect, or if the frame rates of

the images or correcting data are low, then the sort of fast-moving jaw-lowering movements that displace the probe are hard to correct accurately (Scobbie et al. 2008b). Moreover, the extra equipment and added complexity of head-correction suggest that it is less likely to be used in sociolinguistic research or fieldwork. In the other main method, the head is left more naturally mobile, and a stabilising headset is used to secure the probe relative to the head. For example, Articulate Instruments' headset is intended to reduce lateral and rotational probe movement, whilst also allowing relatively unconstrained movement of the head and upper body (Articulate Instruments 2008, Scobbie et al. 2008b), and it has been used in sociolinguistic and linguistic fieldwork (Lawson et al. 2008, Scobbie et al. 2008b, Miller & Finch 2011). The main drawback with the headset is that discomfort can foreshorten the duration of data capture to as little as 20 minutes.

Ultrasound scanners are often noisy, so the speaker should be shielded from this acoustic interference. Ideally the scanner, with the control and synchronisation computers for recording (and perhaps also the researchers) should be in a different room to the speaker(s), but this may not be possible, in which case steps should be taken to minimise the noise. In any case, it may be desirable for the speaker not to feel that their speech is being listened to or judged through the presence of a researcher. However, ultrasound itself can be characterised as a technique that "measures the tongue", an ethically-defensible experimental misdirection which reduces the Observer's Paradox (Labov 1972) and helps to ensure that articulatory data is at least as ecologically valid and vernacular as acoustic data (Lawson et al. 2008). Participants readily accept that they need to speak aloud as part of this "tongue measuring" process, whether it's natural or structured discourse, or wordlists, since tongues change shape while people are talking. There is something about the idea that a physical organ is being measured that can hide the fact that the purpose of the experimental instrumentation is, in fact, to record speech itself; such misdirection is harder to achieve with a microphone alone. It is also hard to achieve when participants are themselves university students of linguistics or related disciplines.

This paper will compare acoustic (Bark-transformed F1 and F2) and UTI data, with a focus on exploring the frontness of Scottish English /u/. We will see whether, like SSBE, /u/ is high and front, near /i/, as in Harrington et al.'s (2011b) study of Anglo-English /u/. The analysis will be based on the absolutely minimum amount of data from a socially stratified corpus of 15 speakers collected for another purpose, namely one token of each vowel per speaker, because that is all that is available. This study lets us also consider what can be

achieved quickly with ultrasound as a sociolinguistic research tool. The other main purpose of the paper is to explore the metaphorical basis of [frontness] in an articulatory sense, in two orientations of the xy mid-sagittal space. Together, these acoustic and articulatory analyses help us question the phonological assumption that the Scottish GOOSE vowel is a phonologically high and back vowel, namely /u/, while being phonetically central and high.

2. Method and qualitative appraisal of data

2.1. Ultrasound tongue images and general design

Data was collected as part of a broader project to investigate the efficacy of ultrasound tongue imaging as a joint sociolinguistic and phonetic tool (Scobbie et al. 2008a). The project involved the collection of the corpus ECB08 (Eastern Central Belt 2008), along with an earlier corpus WL07 (West Lothian 2007, collected from ECB08's working class location with some of the same participants). Participants were aged 12-13 and were recorded in same-sex friendship pairs undertaking spontaneous speech tasks and unstructured spontaneous conversation, and then reading wordlists. Analysis of vernacular speech variables showed that the use of ultrasound tongue imaging instrumentation did not alter the levels of vernacular speech in the participants any more than an audio recording (Lawson et al. 2008), and indeed their spontaneous speech appeared highly naturalistic and unmonitored. Younger teenagers were used as younger speakers tend to obtain clearer ultrasound images than older speakers, and older males in particular.

In ECB08 the speakers comprised a socially-stratified group of 15 teenage speakers of both sexes (Tab. 1), and it is these speakers who are analysed here. The two social backgrounds are shorthand: the WC (working class) school was selected from an area of social deprivation in Livingston as indexed by local government, and the MC (middle class) school was a fee-paying school in Edinburgh. Both the participants in each friendship pair wore articulatory instrumentation during a recording session where they a) read aloud individual lexical prompts from a monitor, b) took part in a structured discourse task (map task followed by an unstructured conversation with a friend). As one MC male missed the recording due to illness, one of the other participants doubled up as his conversational partner. In neither condition were the researchers in the room. For full details, see Scobbie et al. (2008a).

Table 1. Participant details. WC speakers are prefixed L (Livingston), MC speakers E (Edinburgh). They were aged 12-13 and recorded in same sex friendship pairs.

	WC Working Class (West Lothian)	MC Middle Class (Edinburgh)
Male	n=4 (LM15-18)	n=3 (EM3-5)
Female	n=4 (LF1-4)	n=4 (EF3-6)

The main phonetic purpose of this corpus was to investigate variation in the lingual articulation of coda /r/, an important sociolinguistic variable in Scotland (Stuart-Smith 2003, Stuart-Smith et al. submitted). Collection of wordlist data lasted approximately 20-30 minutes, and materials were therefore mainly focussed on /r/ words or control words. Only a small amount of time was available to collect other data, e.g., on plain vowels. For more on vowels before /r/, including a merger and the development of /ɚ/ by MC speakers, see Lawson et al. (submitted).

The materials analysed here (Tab. 2) were an annex to the main dataset investigating /r/. The use of just a single prompt for each vowel meant that we were able to ask whether a useful overall picture of the vowel space can be obtained in less than a minute of data collection. Some other variables that it would have been good to examine in detail with ultrasound include vocalisation of coda /l/ and glottal replacement of /t/. For the very practical reasons outlined, the vowel annex was as small as possible: in fact only a single token of each of the nine monophthongal vowels in Scottish English was collected. The materials were presented on screen as an orthographic prompt at the beginning of the 20-30 minute period of data collection. While the main dataset was randomised, the vowel materials appeared as a list, one word on screen at a time, and each speaker produced the words in the same order.

Table 2. Monophthongal Scottish English vowels: conventional phoneme symbols and the lexical items used to elicit them. Lexical items contain only non-lingual consonants, to avoid the effects of coarticulation.

BEAM	FAME	HIP	HEM	MAP	HUM	AWE	HOPE	BOOM
i	e	ɪ	ɛ	a	ʌ	ɔ	o	u

Clearly, it would be preferable to undertake a more extensive study, but this is only possible as resources allow. A methodological research question is, therefore, to ask how useful such an extremely limited, indeed minimal, articulatory dataset can be. The broader linguistic and sociophonetic research questions relate to the overall relationships within this vowel space. In particular, what is the articulatory

nature of the lingual articulation of the vowel /u/ which is phonologically high and back in most theoretical work, but in phonetic realisation is widely accepted to be relatively central or even front in the vowel space.

A standard medical portable ultrasound scanner (Concept M6 with a 120° Field of View microconvex probe) was used for data capture, in a speech laboratory setting. The depth (magnification) of the image and the scanner frequency varied depending on the size of the head of the speaker. The probe, with gel, was in direct contact with the submental surface, with the Articulate Instruments headset for stabilisation.

2.2. Acoustic analysis

Because the speakers were adolescents of both sexes with a wide variety of orofacial sizes especially among the boys, the vowel formants were measured by hand using the Articulate Assistant Advanced™, or AAA software (Articulate Instruments 2012) by the first author, and independently by the third author using PRAAT (Boersma & Weenink 2012). The acoustic measurements reported here were made without reference to any articulatory information, but with knowledge of the lexical target.[9]

After one year, the first author checked his own measurements of a randomly selected 20% of all the tokens (at which point, of course, the general results were known). A Pearson's correlation showed an adequate match for F1 ($r = 0.89$), and in the remeasure, F1 values trended on average higher by 43Hz (s.d. 92Hz). F2 trended on average higher by 20Hz (s.d. 180Hz, Pearson's $r = 0.96$). (These remeasured values replaced the original values, because some of the original set contained errors.)

Given their importance here, all tokens of /i/, /e/, /o/ and /u/ which differed by more than 100Hz were re-examined jointly, and we agreed qualitatively on the formant centre frequency location, and the two measurers then recorded a new value from AAA or Praat spectrographic displays, as appropriate. Collaborative correction of large disparities in the other five vowels was also undertaken. After this process, on average, the two measurers' F1 differed by 3Hz and their F2 by 18Hz (with standard deviations of 63Hz and 67Hz respectively). The F2 difference (in a paired t-test) remains significant ($t(123) = 1.98$, $p < 0.005$): the first author's F2 measurements are on average 18Hz higher. Given the wide range of formants in a vowel set, unsurprisingly the correlation between the measurers is very high: Pearson's r values are 0.94 for F1 and 0.99 for F2.

In summary, each quantitative value analysed here is the average of two manual measurements, one from each measurer, which have been checked and are broadly in agreement. All measures have been transformed into the psychoacoustic Bark scale (Traunmüller 1990) for display, statistical analysis and reporting in order to more closely approximate formant relationships as they are perceived by the hearer.

While a normalisation process was used to gauge the location of /u/ (see below), mean vowel formant values reported were not normalised, given the small dataset, and are indicative only. The main focus is a WC vs. MC comparison, so boys and girls are pooled within each social group. The male/female differences are smaller in these adolescent speakers than they would be in adults, but interspeaker differences are large. The vowel ellipses plotted below (Fig. 4) therefore reflect the fact that they are based on only a few tokens each from physically varied speakers, but are comparable across social groups.

In the quantitative analysis of /u/, F1 and F2 will be presented relative to each speaker's own /i/ (and /o/), which appear to be stable and widely spaced in the horizontal dimension of the vowel space, so /u/ will be normalised in that regard. In fact, non-normalised measurements and Euclidian distances give similar results.

There were at most seven tokens for each MC vowel and eight for each WC vowel. Due to an error in data collection, there is no token for /e/ from one WC speaker (LM15), with a knock-on effect affecting the number of tokens in the frontness measures (both acoustic and articulatory). There is also a missing token of /ɪ/ from LF2. Thus 133 tokens were measured ((9×7)+(9×8)-2).

2.3. Articulatory Analysis

2.3.1. Tongue curve drawing

It is important to examine the vowel spaces of each individual speaker, because different speakers have different vocal tract sizes, data capture settings, and different probe-cranium orientations. It would be preferable to make measurements on averaged data from each speaker, but as explained, we have only one token of each vowel per speaker. Averaging within a speaker is extremely fast and simple within AAA software (Articulate Instruments 2012) and can be input to a fast statistical comparison using multiple t-tests, but averaging across speakers without explicit normalisation procedures can only be viewed as illustrative (see below). Analytic tongue curves were semi-automatically fitted (using point-to-point smoothing) to the images

using AAA software (versions 2.13 & 2.14) with correction of errors by hand.

Figure 1 and Figure 2 below provide typical examples of an individual's vowels and vowel space, based on single tokens of each vowel. The figures show fairly clear separation of vowels, despite the increased analytic noise that results from this absolutely minimal data sample. The use of individual tokens is evident in the wiggly appearance of many of the tongue curves in Figure 3 or Figure 1.

Greater degrees of tongue-curve smoothing of each single token can be achieved if a spline that is fitted to the tongue using a smaller number of control "knots" or anchor points, even though the underlying image is no different in quality. Smoothing creates individual tongue curves that contain natural-looking curves, but care must always be taken to also respect the underlying data. The analysis presented here was undertaken both with AAA's curve smoothing and, in an earlier draft, without it. The overall results with respect to the location of /u/ are unaffected.

One potential concern is that the lowness of phonologically low vowels /a/ and /ʌ/ may be under-represented, because we have not corrected for any downward movement of the probe under the influence of jaw opening. This was not technically possible at the time of data collection in our laboratory. Also, it's possible that the stabilised probe's presence may discourage a speaker from opening as much as they otherwise might. If either of these problems arose, the effect would be that the vowel space would be bigger in the vertical dimension of the image than it appears in our plots, for the low vowels. However, we are primarily interested in high and higher-mid vowels, and we are confident that there is no effect of probe movement or stability with these, so our analysis of /u/ in relation to the higher vowels will not be affected.

Measurements in AAA are made on a fan-shaped grid, the origin of which lies within the area of the image corresponding to the probe. It has 42 measurement radii whose orientation is based on radial echopulse-based measures which ultimately provide the data on which the interpolated 2D image of the tongue is based. The actual number of echopulse beams is not relevant – the same grid is superimposed on all images, re-angled to fit probe's view, whether the probe has a 90°,120° or other size field of view.

2.3.2. Examples of individual articulatory vowel spaces

Figure 1 demonstrates a vowel space for one speaker based on smoothed tongue curves of single tokens, where the frame selected

was one that was judged to holistically capture the vowels target. It shows how powerful ultrasound is as a field tool – this is what can be quickly extracted with minimal intervention from single observations. Each speaker's vowel space took around a minute to elicit, not much longer than an audio recording would be. If the focus of research were on dynamic changes in articulation or the relation of the articulatory change to acoustic events, we would recommend using high-speed UTI to ensure accuracy in synchronisation, and here we focus on analysis of a single target shape. The overall shape seen here is echoed by all the other subjects, except for features that we mention below. Later in the paper,we will present rotated vowel spaces, but the figures in this section show the raw orientation.

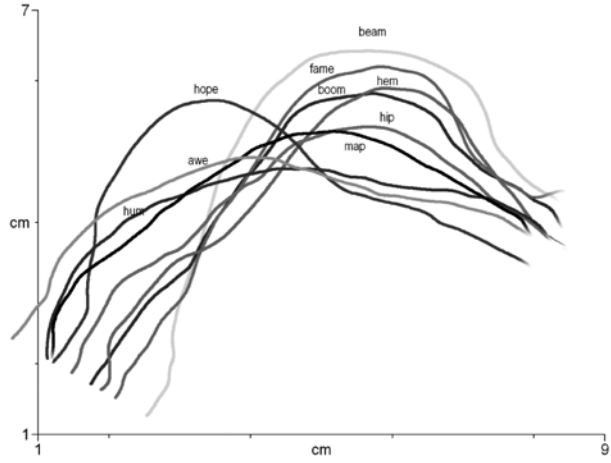

Figure 1. Example of a single speaker's vowel space as constructed from tongue images of single tokens of each of the 9 Scottish English monophthongal phonemes (labelled with prompt, see Tab. 2 for phoneme labels). Horizontal and vertical axes have an arbitrary orientation determined by the probe angle. Anterior to right. Speaker EF4, middle-class female.

The vowels are well-separated. They show a large range of tongue surface locations in the root area (lower left) with an overall distance of about 2cm of tongue root fronting and backing (from *beam* to *awe*). The tube width in the anterior area also varies greatly, with the tongue front in *beam* being about 2cm higher than the floor of the mouth (revealed by retraction) in *hope*, *awe* and *hum*. The root ends of each vowel are abruptly cut off at the edge of the fan-shaped sampling area before the tongue could be seen curving inwards again to the very root at the bottom of the epiglottic valecula, as MRI would

reveal (Cleland et al. 2011), but given our 120°, we see more of this part of the tongue than some other ultrasound imaging systems. Approximately 9cm of the tongue's surface is imaged, depending on the vowel and speaker. For example, EF4's traces (Fig. 1) vary from about 10cm for /i/ down to 8.5cm for /a/, and LM17's from 10.5cm for /i/ to 8.8cm for /a/ and 7.8cm for /ɪ/ (Fig. 1). These examples are comparable to the adult mean of 8.5cm for Scottish /a/ in the images used by Zharkova et al. (2011). As noted above, for vowels like /o/ and /ɔ/, the floor of the mouth in front of the retracted tongue tip, rather than tongue muscle, is what corresponds to the most anterior part of this curve. Though we need always to be cautious about drawing firm conclusions about the relative locations, sizes and shapes of adjacent vowels from just single tokens in individual speakers, the collection of fifteen single speaker/single token diagrams are nevertheless useful, given their consistency and comparability. On that basis, we conclude that the vowel space in Figure 1 is valid in its general shape, because in general terms it is typical of all the corpus' speakers.

Consider now the relative positions of /i/ and /o/. Informally, they appear to have the highest front (right) and highest back (left) constrictions. We can see that the position of /u/ in relation to them is very front.[10] However, it is not high, like /i/, as its phonological label (and all previous literature that we are aware of) would lead us to expect. In Figure 1, /u/ occupies a mid (indeed open-mid) position not dissimilar to /ɛ/ — and such a non-high /u/ is entirely typical (Tab. 3). See the typical WC speaker in Figure 2.[11] Every WC /u/ is noticeably lower than /ɛ/. This could be a fact about /u/, or because WC /ɛ/ is higher, and closer to /e/, than MC /ɛ/ is. We will return to the height of /u/ below.

Table 3. Number of speakers with the token of /u/ categorised for overall height relative to their /e/ and /ɛ/ tokens.

	MC	WC
/u/ below /e/	7/7	8/8
/u/ as low as /ɛ/, or lower	5/7	8/8

Consider also the /ɪ/ vowel, which is expected to be lower for WC speakers than MC ones (Stuart-Smith 1999, Eremeeva & Stuart-Smith 2003). EF4 (Fig. 1) is a MC speaker, and her /ɪ/ is apparently rather lower than both /e/ and /ɛ/, not between them, as the literature, based on F1 and impressionistic quality, leads us to expect. However, as expected, generally in MC speakers, /ɪ/ is closer to and higher than /ɛ/, while in WC speakers, such as LM17 (Fig. 2) this distance is greater. (See also Fig. 10 below for indicative averages).

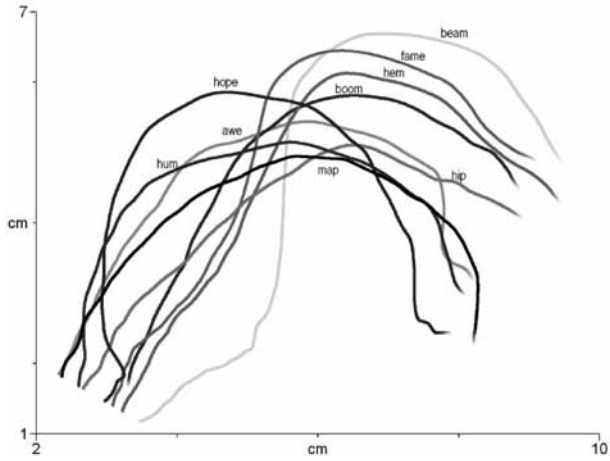

Figure 2. Example of a single speaker's vowel space as constructed from tongue images of each of the 9 monophthongs (labelled). Horizontal and vertical axes are in cm, with an arbitrary orientation determined by the probe angle. Anterior to right. Speaker LM17, working-class male.

This consistent relativistic positioning of /u/ leads us to conclude that, impressionistically, Scottish /u/ in these young speakers is front and non-high, and moreover that it tends to be in a location no higher than the unrounded Scottish /e/; a pattern repeated in the hyperarticulated speech of a young adult in Scobbie et al. (2012). We can also see that /ɪ/ tends to be lower in WC speakers than in MC speakers.

Rather than presenting 15 diagrams in the results section, it would be better to try to quantify tongue body position relationships in the Scottish monophthongal vowel space in more objective terms, and thereby enable more cross-speaker comparisons. To this end, we will now explore some relevant issues within a quantitative analysis of /u/. The shape and location of /u/ and of comparator vowels /e/, /o/ and /i/ were measured, the latter providing crucial contextual information, given that we cannot relate /u/ in this dataset accurately to cranial features, which do not appear consistently on ultrasound recordings as they do in, for example MRI recordings, apart from parts of the hard palate, which are revealed when the speaker presses their tongue against the palate, or swallows liquid. Quantitative articulatory analysis will thus be based on 59 tokens (4×15 minus one missing token of /e/).

2.3.3. Defining horizontal

When quantifying the fronting of the curved tongue surface of /u/ in articulatory space, it is necessary to define both horizonality in the vocal tract and which part or aspect of /u/ will be measured relative to it, i.e., whether we apply measures based on the conception of the highest point of the tongue, or something more sophisticated. The tract itself curves through roughly 90°, making the dimensionality both arbitrary and misleading unless the definitions are closely considered (and open to discussion once they are so defined). The tongue is a curving body that may define a narrow constriction in this tube, but with ultrasound the entire 2D cross-sectional area is not measurable. EMA research also typically uses the occlusal plane as its base for horizontality, and may compute locations and velocities in this single dimension (or vertically), as well as in 2D (which manages to avoid the issue). However, EMA only examines the anterior portion of the vocal tract as the electromagnetic coils used to track articulator movement cannot be comfortably fitted further back in the vocal tract than the back of the tongue.

Even at an informal level, the concept of horizontality is complex and vague. It seems to be based on both the physiological structure of the anterior parts of the vocal tract and their orientation in a habitual upright human stance. A person, standing upright, looking at an object at "eye-level", enables a lay definition of horizontal in the vocal tract as being parallel to this eye-level (which will also typically be parallel to the floor). Our speakers were sitting, and the probe was not necessarily vertical in orientation to the floor, nor were the prompts necessarily at eye level. What remains unknown is whether the internal articulatory view of the vocal tract in this orientation, would, in fact, correspond to phoneticians' informal and varying definitions of "horizontal" as used in textbook diagrams.

As is common in ultrasound research, our speakers were sitting on adjustable chairs, were of different heights, but were reading from a screen of a fixed height. More importantly, the ultrasound probe is generally fitted to the external surfaces of the head to provide as good an image as possible of the tongue. The probe was angled in such a way as to balance the visibility of the surface near the blade and tip within the angle of view. The probe was not orientated either to the room's conventional horizontal nor to a speaker-internal definition (e.g., the occlusal plane). As noted, speakers were wearing the Articulate Instruments headset, so were free to move their heads naturally while talking and reading, while the probe was kept in the same location relative to the head. Thus we need to establish a definition of "horizontal", on which we can quantify one dimensional front-

ing based on each speaker, not on the external orientation of a speaker's head held steady within the laboratory space, but from aspects of the vocal tract itself. This could have the advantage of being applied to other instrumental data, e.g., MRI or EMA. Even so, using ultrasound images alone for this process will limit the range of possibilities, given the lack of passive articulator information contained in these images.

In the rest of this section we will sketch two possible approaches to the definition of horizontal using ultrasound data, before moving on to presenting the results. In both cases, we also have to deal with locating the constant curving shape of the tongue surface as a point on this dimension. We will assign a location in the vowel space by picking out the point on the curve nearest to the horizontal axis, i.e., the high point. An alternative would to pick a centre of gravity or other point or area to try to characterise tongue location, though the fact that not all of the tongue surface has been imaged argues against this approach. It would be more suitable for the analysis of MRI images, for example. More options are possible, especially if passive articulator locations can be estimated or measured, such as the upper teeth, back of the hard palate, a straight line estimation of the rear wall of the pharynx. Such approaches are not pursued further here.

The most straightforward approach would be to use measures parallel to the x-axis and y-axis of the scanner's images, but this offers no prospect of normalisation across sessions, let alone speakers.

2.3.4. Using the vowel space itself to define the horizontal axis

There are a number of ways in which the vowel space, i.e., the range of locations of the tongue itself as it forms vowels or consonants, without direct reference to passive articulators, could be used to define axes, and any would lend themselves to use with ultrasound data. The first, very simple, approach explored here is to propose that this arbitrary coordinate space is defined on two consistent and relevant vowels, namely the highest front (/i/) and the highest back vowel in the system, which in this case appears to be /o/. To characterise the relative location of /u/, a common tangent was drawn to link these two phonetically peripheral vowels (Fig. 3). This unique /i/-/o/ reference line is then treated as the horizontal orientation for the speaker in question on which to define /u/-fronting. Perpendicular lines can be dropped (or raised) from this plane to the unique closest point on the tongue curve for /u/. This line, the /i/-/o/ plane, therefore lets us define both "frontness and "lowering". Here, frontness will be defined as the distance back from /i/, which was assumed to be the most stable vowel (Gendrot & Adda-Decker 2007). Both frontness and lowering (the per-

pendicular distance dropped from this plane) are therefore calculated to a single point on another vowel's tongue surface, namely the point on its curve closest to the /i/-/o/ plane.

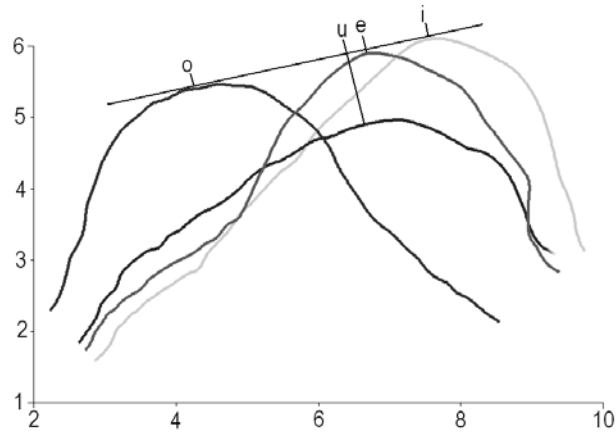

Figure 3. Example of measurement using a common tangent (the /i/-/o/ plane) to define the locations of other vowels. The x and y axes are provided by the edges of the rectangular image generated by the ultrasound scanner.

Picking a single point on a convex curve has the advantage of providing a replicable point of measure, even though it is not a flesh point. Smooth tongue curves provide an easier basis for measurement reducing any possibility of ambiguity, and there was in fact little or no ambiguity in the application of this method here.

The angle of the /i/-/o/ plane relative to the scanner's x/y axes (as shown in the image) varies from speaker to speaker but is approximately 30° anticlockwise from the scanner's x axis, i.e., the anterior (right) side of this plane is raised about 30° relative to the scanner's horizontal. Impressionistically, this orientation seems to be too much of a rotation compared to expectations, but a "high back" vowel constriction should be expected to be "turning the corner" of the vocal tract, so perhaps the orientation in Figure 3 is a suitable orientation and we do not think it is appropriate to reject it on these a priori grounds. However, defining this as horizontal effectively rotates the anterior portion of the images downwards and the root up and back, creating unfamiliar-looking diagrams that don't fit phonetic intuition well and will give one dimensional measures of fronting that in EMA studies would be seen as incorporating an element of raising.

2.3.5. Using an estimate of the occlusal plane as "horizontal"

The second approach we consider is to use passive articulators to create a speaker-specific, cranially-based set of dimensions. In particular, a well-accepted approach in articulatory research is to capture each speaker's unique occlusal plane; a common landmark defined using the upper teeth as the passive structures. An alternative passive structure that could be used, and it is particularly useful for alignment of different sessions from the same speaker is the hard palate and alveolar ridge (Wrench et al. 2011), but the occlusal plane offers a direct and consistent definition of the horizontal dimension.

If the occlusal plane had been recorded for each speaker in ECB08, then it would have been simple to rotate the articulatory space correctly in a speaker-specific way, but we had not, at that point, developed a method of obtaining traces of the speaker's occlusal plane (Scobbie et al. 2011). In the absence of this information, an estimation of the appropriate occlusal rotation must be used. In future, we would aim to capture occlusal information for all speakers, if practicable. To deal with the pre-existing data from ECB08 (or data collected in other laboratories, or when biteplane/occlusal orientation is not possible), we opt for a +20° (clockwise, with anterior to right) downward sloping occlusal plane, based on analysis of other data (Scobbie et al. 2011). Defining this as horizontal effectively rotates the image, raising up and retracting the anterior portion of the images (the tongue front) and lowering and advancing the posterior part of the images (the tongue root). It therefore increases the apparent gradient of the /i/-/o/ plane, making it appear far less suitable as a horizontal measure than it was in Figure 3. The /i/-/o/ plane viewed from the point of view of an occlusal horizontal is tilted by about +50°, making the /i/-/o/ plane almost half-way between horizontal and vertical.

Comparison of /u/ frontness using the "common tangent" and the "occlusal plane" methods is a useful check on what is, after all, the rather arbitrary extraction of a single point to quantify the location of each vowel in a curving space. Each method will inevitably result in a different point on the surface curve being used to calculate the location of the vowels relative to each other. Comparison of these two systems of measurement also lets us examine a very anterior-focussed definition of vowel quality against one that combines vowel constrictions in both anterior and posterior parts of the upper vocal tract.

2.3.6. Methodological conclusions

The acoustic space of vowels is complex, with very many correlates of phonological identity in multiple dimensions being available, from which the most statistically reliable cues to contrast and identity in a given language are drawn. There is, however, a widespread informal assumption that F2 alone is a suitable measure of the frontness of vowels (and F1 of height), particularly where there are no complicating factors. Frontness of vowels and consonants is also often mentioned in articulatory studies using techniques such as Electromagnetic Articulography, where the focus is on the anterior part of the vocal tract, and where horizontal and vertical axes in 2D plots have to be defined. Horizontal measures are sometimes used in quantitative, one-dimensional analyses of frontness, though articulatory analysts are also well-aware that this is a simplified short-cut. While it is entirely acceptable to discuss the frontness of vowels informally, or in introductory text books, or in impressionistic, transcription-based IPA analysis, or in terms of phonological features, it must be remembered in such contexts the term "front" is only tenuously and informally related to the vocal tract space, from both acoustic and articulatory perspectives.

The articulatory data provided by ultrasound tongue imaging, like X-ray or MRI images, both covers a large part of the vocal tract, as the tube bends through 90° or so, and has horizontal and vertical axes somewhat randomly, based on the orientation of the scanner's probe. Given that *some* orientation of the 2D space is essential, it would be useful to have a consistent definition of "horizontal", and even better to have a meaningful one, such as one based on a real or an estimated occlusal plane. It would then be possible to visualise and make initial quantitative measurements of frontness and height from tongue curves in the anterior portions of that space, and to do so for both UTI and EMA.

We now turn to our comparative analysis of contemporary Scottish English /u/ using some of these simple articulatory/acoustic techniques to express the vowel's phonetic frontness and height, and to discuss these findings in terms of its previously-assumed identity as a high, back or central vowel.

3. Quantitative Results

3.1. Pooled acoustic results

Based on acoustic evidence, the vowel /u/, as expected, is more central than back for both MC and WC speakers in the ECB08 corpus, perhaps even front of centre in MC speakers, and not at all near the high back position (Fig. 4). It is also non-high; being around a height comparable to /e/ and, to a lesser extent, /ɔ/. In terms of F1 values, /i/ and /o/ are very similar. This matches Scobbie et al.'s (1999b) analysis of /i u o ɔ/ in Stuart-Smith's 1997 sociolinguistic corpus of 32 Glasgow speakers. The biggest social difference in the MC and WC monophthongal vowel systems that can be seen in the data presented in Figure 4 is that WC /ɪ/ is lower and backer than WC /ɛ/, occupying a central position under /u/, whereas the MC speakers' /ɪ/ is located between their /e/ and /ɛ/. That is to say, the relative positions of /ɛ/ and /ɪ/ are reversed in MC and WC speech.

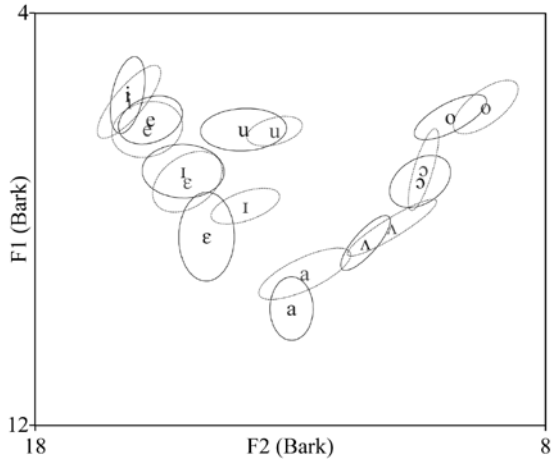

Figure 4. MC (solid) and WC (dashed) labelled ellipses (±1σ) for each of the 9 Scottish Standard English monophthongal vowels. Anterior is to the left. A colour version, with MC blue and WC red is online at: http://linguistica.sns.it/RdL/2012.htm

The narrow ellipses in vowel variation are probably the result of mixing of small numbers of male and female speakers. They are, moreover, adolescents whose vocal tracts vary in size. However, examination of individual and pooled data from the four sex-class subsets show that the overall vowel spaces shown in Figure 4 are indeed representative, and in fact even the means based on small amounts of data in the sex-class cells (mostly n = 4 but MCM males n = 3) are as expected (Tab. 4).

Table 4. Sex-social group means for the four high & upper-mid vowels (Bark) organised by decreasing F2 and increasing F1.

		POOLED	WCM	WCF	MCM	MCF
F2	/i/	16.2	15.7	16.6	16.2	16.2
	/e/	15.8	15.3	16.2	15.8	15.8
	/u/	13.6	13.1	13.6	14.0	13.9
	/o/	9.5	8.8	9.6	9.9	9.8
F1	/i/	5.7	5.2	6.2	5.2	5.9
	/o/	5.9	5.4	6.2	6.0	6.0
	/e/	6.2	5.9	6.5	6.1	6.1
	/u/	6.3	6.1	6.5	6.5	6.1

The phonologically front vowels /e/ and /i/, despite the fact that pooled data is far from homogeneous, can be shown to be significantly different in raw F2 (Bark) in a paired t-test, $t(13) = 4.76$, $p < 0.0005$, though the average difference between them is just 0.4 (s.d. 0.3). In all speakers, /i/ has a higher F2 than /e/. In summary, /i/ and /e/ are front, /o/ is back, and /u/ is front-central, perhaps trending even more front in MC speakers (Fig. 5).

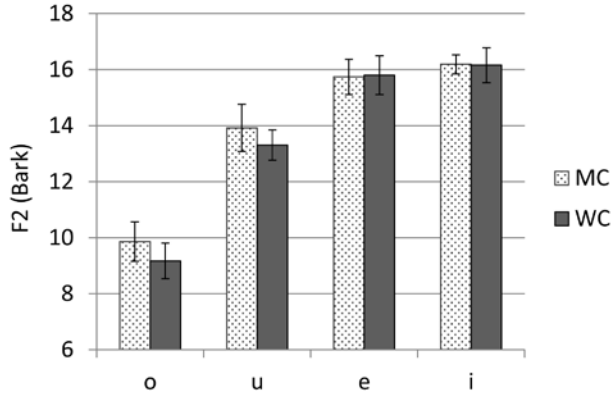

Figure 5. Mean F2 for /u/ and reference vowels /i/, /e/, /o/, pooled by social class.

Given the inter-speaker differences, we will quantify the acoustic frontness of /u/ more carefully, relative to /i/ (and /e/ and /o/), and will present these results below in conjunction with the articulatory results. In both acoustics and articulation, by comparing vowels within speaker to their own /i/, a degree of normalisation is achievable. In both domains, the F2 difference between /i/ and /u/ will be compared

pairwise to the F2 difference between /i/ and /e/, rather than comparing raw F2 of /u/, /i/ and /e/, and the same relative measurement will be carried out for horizontal articulatory distance. On the basis of Figure 4, Figure 5 and Tab. 4, we expect the difference /i/-/e/ to be significantly smaller than /i/-/u/ in both the acoustic and articulatory domains.

3.2. Relative frontness of /u/

3.2.1. Acoustic frontness, F2

We present first the F2 (Bark) differences, reporting planned 2-tailed paired sample *t*-tests. First, we confirmed that the difference between pooled /u/ F2 vs. pooled /e/ F2 (e.g., Tab. 4) is significant when approached from a normalised perspective, i.e., /i/-/e/ vs. /i/-/u/ (Fig. 6a, below). Relative to /i/, the pairwise results are significant for F2, $t(13) = 12.8$, $p < 0.0001$,[12] with mean differences of 0.4 (/i/-/e/ s.d. 0.3) and 2.6 (/i/-/u/, s.d. 0.6). The mean of each speaker's /u/-/o/ difference is 4.1 (s.d. 0.7), so /i/-/u/ is on average smaller than /u/-/o/, $t(14) = 4.99$, $p < 0.0005$, leading to the conclusion that /u/ is closer to /i/ than to /o/.

Each speaker has a normalised percentage measure of the frontness of /u/ (and other vowels), relative to /i/ and /o/, where 100% would mean that /u/ had the same F2 as /i/ and 0% would mean that /u/ had the same F2 as /o/. The mean value for /e/ is 94% (s.d. 4%), which is indeed very front. The mean frontness of /u/ is 61% (s.d. 9%).[13] In ECB08, only two speakers have /u/ further back than the halfway point between /i/ and /o/ in F2 (LF1 and EF3), both by just 0.7Bark. EF4 shows the largest imbalance (70% fronted) since for her /i/ is 16.3 and /u/ 14.2, while /o/ is 9.3.

These % values are the most meaningful way of comparing the acoustic and articulatory results, to which we now turn.

3.2.2. Articulatory fronting in the common tangent plane and in the assumed occlusal plane

In articulatory space where we take the plane with a tangent drawn between the /i/-/o/ tongue curves as horizontal, a pairwise comparison of the distance from /u/ to /i/ and /e/ to /i/ shows that /u/ is significantly less front than /e/, $t(13) = 3.64$, $p < 0.005$, with a mean distance from /i/ of 2.3mm (s.d. 3.7mm) for /e/ and 5.8mm (s.d. 3.5mm) for /u/ (Fig. 6b). (The acoustic distance data from Section 3.2.1. is shown in Fig. 6a.) Compared to the acoustic results, the articulatory data based on normalised frontness along the common tangent of /i/ and /o/ makes /u/ appear even more front. In % terms, /e/ is 91% front (s.d. 16%), and /u/ is 74% (s.d. 15%). For comparison, and in raw values, /o/ is 22mm back (s.d. 4.9mm) from /i/.

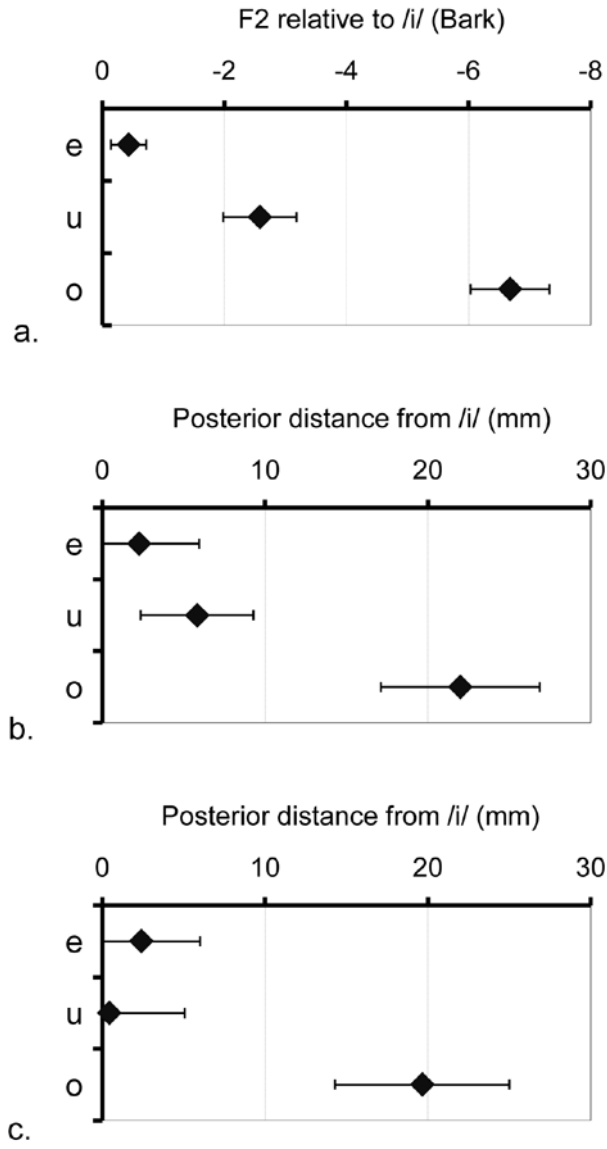

Figure 6. (a.) Mean decrease in F2(Bark) indicating backness of vowels relative to each speaker's /i/'s F2. (b.) Mean distance back along the /i/-/o/ plane from each speaker's /i/. (c.) Mean distance back from each speaker's /i/ along the estimated occlusal plane. One s.d. marked in each case.

Now consider the quantification of fronting when the data is rotated to an assumed occlusal plane (Fig. 6c). Here /u/ is even more highly fronted: there is, in fact, no statistical difference in the frontness of /u/ and /e/ on a pairwise *t*-test. On average, /u/ is just 0.4mm backer than /i/ (s.d. 4.6mm), while /e/ is 2.4mm further back (s.d. 3.6mm). The equal frontness of /e/, /u/ and, by assumption, /i/ is supported by categorical trends in the articulatory data: in 9/14 tokens in the pairwise comparison, /u/ is fronter than /e/, and for 6/15 speakers, the token of /u/ is even fronter than the token of /i/. Relatively, compared to the /i/-/o/ distance, /u/ is just 0.4% back from /i/, i.e., 99.6% fronted (s.d. 26%) is while /e/ is 90% front (s.d. 17%). /o/, for comparison is 20mm back (s.d. 5.4mm) from /i/.

Is one of these articulatory results a more accurate measure of frontness? Well, recall that "horizontal" is just a construct, and our measure of a single point is just one convenient way to locate the whole tongue within that space, one which echoes Daniel Jones's original proposals. It happens to work well for comparing two vowels that are close in space, capturing this clear qualitative impression that the UTI images provide. What matters is that, on either approach, though the horizontal axis differs by around 45°, tokens of /u/ from the ECB08 corpus are truly "front" in the articulatory space, rather than central, as they were in the acoustic analysis of F2.

3.3. /u/ *height*

In the articulatory metaphor, "lower" vowels have higher F1. In raw F1, /e/ is lower than /i/, i.e., has a higher F1 in Bark, $t(13) = 2.37$, $p<0.05$, as is /u/, $t(14) = 3.59$, $p < 0.005$. We found that /e/ and /u/ were pairwise equally "lower" compared to /i/, with no significant difference. On average the F1 of /e/ is 0.5 Bark higher than /i/ (s.d. 0.7) and /u/ is 0.6 Bark higher (s.d. 0.7). Thus in Figure 7a, with its reversed y-axis, neither vowel appears as high as the baseline of /i/, and both are lowered about equally (recall Fig. 4). A categorical perspective, supports this finding, because there was a great deal of inter-speaker variability in the relative F1 values of /e/ and /u/. In 5/14 speakers, the token of /u/ had a higher F1 than that of /e/.

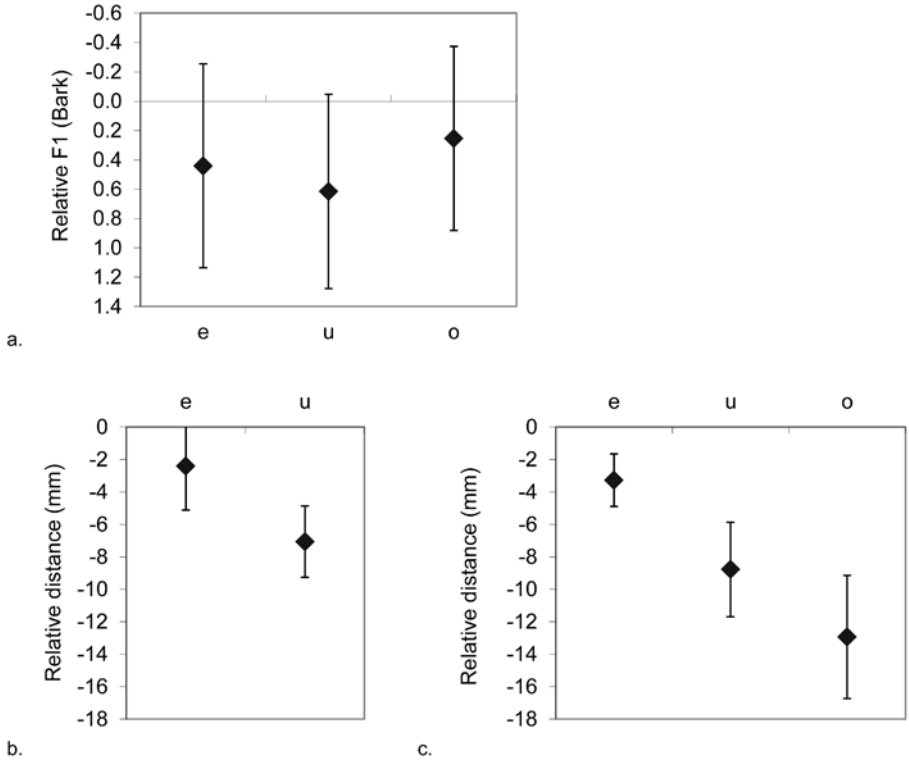

Figure 7. (a.) Mean increase in F1 indicating lowering (one s.d. marked), relative to the average of /o/ and /i/'s mean F1. (b.) Mean (negative) distance perpendicularly below the /i/-/o/ plane of the nearest point of the tongue surface for the vowels /e/ and /u/. (c.) Mean (negative) distance perpendicularly below the level of /i/ on an estimated occlusal plane of the nearest point on the tongue's surface for the vowels /e/, /o/ and /u/.

In an articulatory analysis in which the /i/-/o/ plane is defined as horizontal (Fig. 7b), a paired sample *t*-test showed there is greater articulatory lowering relative to this plane of /u/ compared to /e/, $t(13) = 7.1$, $p < 0.00001$. On average, within speaker, /u/ is 7.2mm lower than the plane (s.d. 2.2mm) while /e/ is just 2.4mm lower (s.d. 2.7mm). From a categorical perspective, /u/ is lower in 13/14 pairs. On average the difference is 4.8mm (s.d. 2.5mm). The height of /o/ on the /i/-/o/ plane is 0mm by definition. On this articulatory measure which assumes /i/ and /o/ are equally high (and in which /e/ is 2.3mm more front than /u/, see above), /u/ is clearly lowered, and /e/ is probably lowered, and is nearly 5mm higher than /u/.

Relative to the assumed occlusal plane (Fig. 7c), a paired samples t-test again shows that there is greater lowering of /u/ than /e/, $t(13) = 7.9$, $p < 0.00001$. On average /e/ is just 3.3mm lower than /i/ (s.d. 1.6mm) and /u/ is 8.9mm lower (s.d. 3.0mm). Categorically, /u/ is lower than /e/ for 14/14 speakers. The average extent is 5.6mm (s.d. 2.7mm). On the occlusal plane, however, the putatively high /o/ is far lower than /i/, and in 15/15 speakers. The difference is 12.9mm (s.d. 3.8mm). This is because it is a long way back from /i/, in a curving vocal tract. /u/, /e/ and /i/ are only comparable for occusal height because all are similarly front.

Analysis of scatter plots of individual speakers showed no significant correlation of articulatory and acoustic differences within vowel categories, though of course there are strong correlations across different vowels. More data would be needed to explore the articulatory-acoustic relations of these measures in detail.

To sum up, on either articulatory orientation, and acoustically, /e/ in the ECB08 corpus is not a phonologically high vowel. /u/ is articulatorily lower more than /e/ by about the same amount as /e/ is lowered from /i/, though acoustically in F1 they are the same. We can easily conclude therefore that /u/ in the ECB08 corpus is not high, and could conclude also that it is articulatorily lower than /e/, but only trends lower in the acoustic domain for this F1 measure of height.

4. A qualitative view of the average vowel space

It would be useful, at least to aid qualitative analysis, to create an average set of mid-sagittal tongue-shapes for each vowel across speakers with similar sized heads and probe locations. For quantitative analysis, other steps would be required, since speakers have in fact rather different sized vocal tracts and tongues. A procedure to normalise the articulatory space similar to that used in vowel-space normalisation would be desirable. In addition, articulatory normalisation ought to deal with the fact that the coordinate space for measurements differs due to different location and orientation of the probe, caused in part by different external and internal physiologies and the goal to get the best possible tongue image.

For an idea of what might be possible, and as the first step in visualising the overall /u/ fronting and lowering in Scottish English quantified above, see Figure 8, where /u/, /o/ and /i/ from four different speakers (all MC females) have been overlaid following translation (as indicated by the different axes). In this case, /o/ was used as an

informal reference and speakers' /o/ tongue contours were used as an anchor point to allow comparison of inter-speaker tongue-body location for other vowels. For three of the speakers, /i/ curves fit nicely too, but one speaker's data is on a different scale (EF5), and there are three different combinations of data collection settings.[14] Even so, a broad qualitative picture emerges. It is not clear whether the differences in /u/ are meaningful or not and further data is required. Such images do convey, however, the need for cross-speaker normalisation and averaging, on the one hand, and the value of looking at whole-tongue images, even of single tokens, on the other.

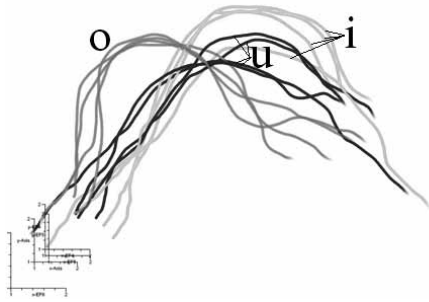

Figure 8. Example of overlaid single vowel tongue curves from four MC Scottish females for /o/, /i/ and /u/. Anterior to right. The images have been translated to line up on /o/ but not rotated or resized. Scales may vary from speaker to speaker. Curves have been gently smoothed. A colour version, with /o/ in orange, /i/ in green and /u/ in blue is online at: http://linguistica.sns.it/RdL/2012.htm

As a rough check on the results calculated on a speaker-by-speaker basis from individual tokens, and to test a different style of analysis, a composite or ensemble average was constructed of /o/, /i/, /e/ and /u/, and (given that there is no scale for this), the relative fronting of mean /e/ and mean /u/ were calculated as a percentage of the distance from mean /i/ to mean /o/ using both measures outlined above. The ensemble average (Fig. 9) was created in the AAA workspace averaging the distance (in cm) of each vowel tokens' surface curve at its crossing point along each of the 42 fan-line on which it had been traced. Given the different placement of the probe relative to the vowel space in each individual, the different angles between fan radii, and the different locations, these measures should be approached with caution. Note that at the ends of the average tongue shapes, particularly at the anterior end, artefacts appear as the number of tongue curves being averaged drops from 15 to zero.

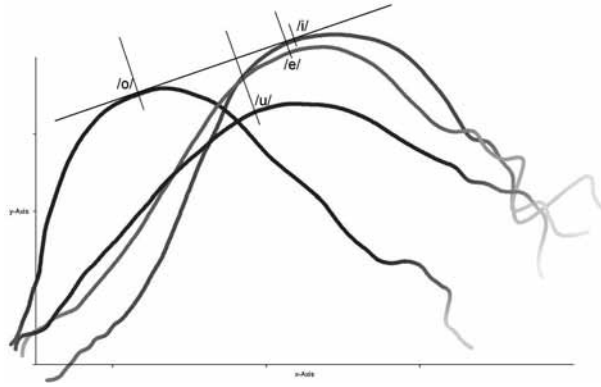

Figure 9. Illustrative ensemble average of 15 speaker's vowel shapes in unitless space, showing common tangent and perpendicular drop lines used to calculate relative fronting on the /i/-/o/ dimension.

Table 5. Comparison of relative fronting and lowering calculated on speaker-specific basis to the relative values calculated from the ensemble average tongue shapes.

fronting %	/e/	/u/
Mean of individual % on /i/-/o/ plane	91	74
Ensemble average, /i/-/o/ plane	95	62
Mean of individual % on estimated occlusal plane	90	100
Ensemble average, on estimated occlusal plane	80	100

Figure 9 provides an attractive smoothed schematic image which superficially presents a similar layout and proportional relationships as seen for individual speakers. Interestingly, the quantitative location of /u/ is broadly similar to the averages reported above (Tab. 5), both on the /i/-/o/ plane, in which space /u/ is fronter than central, but not fully front, and after rotation by 20°, in which space both the individual measures and measurement of the ensemble image suggest that /u/ is fully front.

So, based on both visual examination of this ensemble composite against individual systems and the relative fronting values, even such a rough approximation is useful. As noted above, true normalisation is required, based on multiple tokens from each speaker, and employing translation, rotation and normalisation. But even so, it is not clear how much more useful such ensemble tongue surface images would be than the rough average above, unless UTI data can also incorporate other articulators like the lips, to scale, and normalise them.

Normalised 3D cross-sectional shape data from along each speaker's differently shaped and sized vocal tract is needed, if we aim to build a more useful and acoustically predictive tube model.

If, then, the shapes in Figure 9 can be taken as a characterisation of /u/ in Scottish English in relation to other vowels, we should conclude by presenting fuller vowel spaces for the WC and the MC speakers for comparison (Fig. 10). Seven vowels are presented from the 9 available, partly to keep the diagrams simple and focussed on providing a context for /u/, and partly because vertical probe movement has not been corrected for in the production of the two low vowels /a/ and /ʌ/. In addition, the far left and right ends of the tongue curves have been manually removed at the point where artefacts of averaging were obviously affecting their shape and location. Unsurprisingly, the curves for the vowels are smoother than the individual tokens shown in earlier diagrams, but it is surprising how well the vowels appear impressionistically to be good representations in terms of shape and location, being based on only 7 or 8 tokens each.

The main aspects to note are the similarity of the location of /u/ in both WC and MC speakers. Having quantified through a more detailed articulatory analysis that /u/ is lower than /i/ and about as front as /e/ and /i/, Figure 10 conveys this location accurately, and adds in the qualitative indication that /u/ is even lower than /ɛ/. Secondly, a well known social difference in /ɪ/ and /ɛ/ which was clearly shown in the acoustic plots above (Fig. 4) is also revealed here: in WC speakers, the mean /ɪ/ tongue body position is lower than that of /ɛ/ and for MC speakers it is the reverse, although the MC mean /ɪ/ and /ɛ/ tongue body positions are very similar.

Figure 10. Cross-speaker averaged ECB08 vowel spaces for 7 non-low vowels, rotated 20° after the ensemble averaging was undertaken. WC left, MC right. A colour version is online at: http://linguistica.sns.it/RdL/2012.htm

5. Conclusions

For articulatory analysis in sociolinguistics, perhaps only audio-synched video data (e.g., of the face) is simpler and easier than ultrasound tongue images to collect. UTI is relatively easy, cheap, accessible, and, as we hope we have shown here, it can reveal articulatory information of theoretical importance. There are a number of methodological issues that must be addressed. We explored one of particular relevance to the topic of vowel fronting, and found that the precise definition of horizontal / vertical did not in fact affect these results, in part because we used a normalised measure of frontness and measured a vowel very close to [i], and in part because we applied Jones's concept of the high-point of the tongue, which is relatively informal and convenient. For comparability across studies and articulatory methods, we suggest that UTI data should be rotated to the occlusal plane of the speaker, ideally using their own biteplane (Scobbie et al. 2011), or, as here, by an estimated correction (e.g., here we rotated 20°). While not appropriate for detailed kinematic analysis of tongue movement, location, shape, or constriction, our quick quantitative study both supported our qualitative conclusions, and showed how much can be revealed from what was a comparatively small component of the ECB08 corpus, based on just one token of each vowel per speaker. For many applications, the ease of use of ultrasound tongue imaging is extremely important, and we have found it a very valuable tool for sociophonetic research.

Our data showed conclusively that the Scottish English vowel formally known as /u/ is in fact both neither back nor high, phonetically. Depending on how we quantify frontness, and how we balance articulatory and acoustic (Bark) evidence, "/u/" is front-central. Impressionistically it sounds lax, rounded and front, more like [ʊ̈] or [ʉ̟] or [ɵ̟]. It certainly does not sound like [y]. Phonologically, a single front-central rounded vowel in a linguistic description may be a priori likely to be labelled as /y/, but in our ECB08 data this is only a little less artificial than /u/. Impressionistically, we can see lip rounding and protrusion on /u/, but quantification of labialisation as well as information on audio-visual perception and contextual variation awaits future research (cf. Scobbie et al. 2012).

This discussion of category labels uses IPA symbols in an attempt to remain atheoretic. Whatever the formal representation of /u/ and /o/, if the formalism takes into account the phonetic realisation at all, it is important this is done accurately, otherwise cross-linguistic discussions of markedness and the structure of the vowel

system based on the labels lack any phonetic credibility. This is not to assume, however, that articulatory and acoustic vowel spaces are necessarily congruent – indeed we can see that the midsagittal lingual and Bark spaces we have compared are not. So at the most basic level, it is not clear to us whether acoustic or articulatory phonetic data should take precedence in the choice of formal phonological labels, both here and more generally.

If a phonology is rather abstract, /u/ may as well be said to be high and back. When nothing hinges on phonetic accuracy, phonological labels will tend to be cross-dialectally and historically conservative. In such a model, it is hard to see how phonology will ever manage to predict phonetic change. If a phonology is, on the other hand, more transparent, it would probably classify /u/ as a non-high front or central rounded vowel. This would mean Scottish English has rather a cross-linguistically marked system. This markedness would, moreover, have a different source for its non-high front rounded or central vowel from some other varieties of English whose non-high front rounded vowel is for NURSE (Wells 1982, Lass 1989). For an atheoretical label, central high /ʉ/ seems a suitable compromise, and it is indeed often used for the Scottish English vowel in GOOSE and FOOT (as discussed in the introduction).

As for /o/, /it appears to actually be a high back vowel, in both the articulatory and acoustic spaces examined, and we think it could easily be rephonologised as the system's corner vowel. However, further research is needed to see how phonetically similar Scottish /o/ is to truly high and back /u/ in other languages, and we should not uncritically assume that the highest backest vowel in a phonological sense has to be the nearest to cardinal vowel [u] without a more sophisticated understanding of what the vowel space actually is, and how production and perception combine.[15] In the meantime, we would refrain from advising that the vowel of GOAT should be relabelled as /u/ in descriptive works, though formal phonological analyses of this suggested theoretical change would be welcome.

The dialect-internal and cross-dialectal ramifications of merger-free vowel shifts are usually approached phonetically and functionally. If phonology has any theoretical predictive power, then we should be able to make predictions about future changes on the basis of whether a rephonologisation (rather than just from the movement in continuous phonetic space) has occurred. The main prediction we would make from the reclassification of /u/ to /ʉ/ and /o/ to /u/ would be that that if diphthong /au/ (in MOUTH) keeps its label – which it might if there a tendency for that diphthong to terminate in a corner

of the vowel space, it might flip in one diachronic generation from its current fronted phonetic quality currently rather like [əʉ], back to something more like [au]. In a context of gentle continuous diachronic phonetic change, any rapid and large shift is likely to have been caused by category change elsewhere. It's also possible that the Scottish Vowel Length Rule would be disrupted by phonological lowering of /u/, since it affects just the high vowels /i/ and /u/ (Scobbie et al. 1999a,b).

We have shown that teenage speakers from eastern Central Scotland from two different social backgrounds have lingual and F2 placement for /u/ and for /o/ in vowel space locations that belie the conventional phonological labels. There appear to be small sociophonetic differences which invite closer study, but it is the broad consistency across social class that may be the factor that persuades those phonologists who would describe sociolinguistic variation as "merely" phonetic that *phonological* featural and phonemic category labels like /u/ and /o/ are not well-supported, phonetically. Clear social differences were evident in the articulatory and acoustic relationships of the vowels /ɛ/ and /ɪ/, as expected, but even so, it would probably be seen as controversial to propose that these categories should differ phonologically as well as phonetically between WC and MC systems, at least in modular approaches to the phonetics/phonology interface.

Thus we can see that, despite being quite clear about the articulatory location of /u/, phonetic analyses alone do not provide any easy answers for the thorny problem of phonological interpretation, let alone what phonological label or feature to assign a phonological category. Nor is it clear from a diachronic perspective when such labels should change (and for a useful recent discussion, see Fruehwald 2010). It seems logically and empirically clear that systematic phonetic change precedes phonological change, and that in any speech community, systematic phonetic changes will arise gradually in particular social groups in response to prior smaller and less systematic changes in the speech of others. What then is the role of phonological formalism in predicting change, if any (cf. Janda 2008)? What is the role of phonological formalism in even providing phonetically accurate labels for the phonological categories?

Our view (Scobbie 2006, 2007, Scobbie & Stuart-Smith 2008), following Docherty (1992) among others, is that phonological features (and hence their labels) are a type of emergent abstraction from an interplay of phonological and non-phonological factors, arising from but not imprisoned within phonetic substance. If such (fuzzy) categories recur cross-linguistically, this should be explained through

theories of phonetics, psycholinguistics, sociolinguistics or acquisition: a phonological feature theory based on universal labels has no explanatory role. The label that attaches to the category of /u/ should be just an accurate, quantitative, gradient, variable phonetic label, not a set of features, because in the latter case the label has to suddenly change if any phonological change is to be posited, while phonetic substance can smoothly transition from one state to another. Unambiguous category change might seem feasible from a historical distance of some centuries, after change and any resulting mergers or splits is complete, but in the midst of non-merging variation and change, it seems unlikely, to say the least, that there is an unequivocal point around which continuous phonetic change maps on to two clear categories. In the case of /u/, how unlike [u] does it have to become, in the absence of merger or neutralisation, before a change of phonological category label is deemed essential?

Only finer-grained phonetic data can resolve some of these issues, by allowing us to focus on the competitive interplay of various underlying causes in a realistic manner, rather than forcing us to operate only at the abstract phonological level, in broad categories. Harrington et al. (2011b: 153) are able to hypothesise more convincingly than we can, that in SSBE /u/, "the lingual position… is now so front that lip-protrusion is the principal feature for its differentiation from /i/". But nor can we extrapolate from these findings about SSBE. In our ECB08 Scottish English corpus, we hypothesise that the combination of F1/F2 is sufficient to cue the category, but other acoustic or visual cues might be important, and we do not know yet what combinations of lingual and labial articulations are used or how they are socially structured. We cannot even assume the same synchronic direction of change: the real-time Stuart-Smith et al. (2012) suggests change in Scottish may even be backing and lowering. What seems clear (from small pilot studies and student projects) is that some speakers do articulate /u/ as a fairly high vowel, between /e/ and /i/, and that /u/ may be as rounded as /o/ and /ɔ/ (Scobbie et al. 2012). None of these Scottish /u/ sound, impressionistically, like SSBE /u/ or German or French /y/, being generally laxer, or slightly more retracted.[16] It remains to be seen how Scottish /u/ is cued, and how it is distinguished from the other vowels. In addition to rounding, duration is almost certain to play a part, since /u/ is in general so short before stops (Scobbie et al. 1999a,b).

It would be interesting to see, moreover, whether /u/ in Scottish English is fronted variably in a similar way to SSBE after coronals vs. non-coronals (cf. Zharkova 2007, Brato 2012). To explore such fine

differences in the specification of vowel targets requires production, perception and acoustic data from a wide range of standard and vernacular speakers, in greater numbers, undertaking a range of tasks, and with high-speed ultrasound able to resolve fine-grained spatio-temporal processes.

However, it is clear from the small study presented here that ultrasound tongue imaging has a huge potential to provide new and different data to help us understand vowels and vowels systems for a wide range of theoretical questions. It is relatively straightforward to collect the data, and the technique has been shown to be powerful. Even with a minimal amount of data, the technique can reveal linguistic patterns that were unexpected and challenge contemporary understanding, as well as shedding light on established problems.

We have also explored the articulatory metaphors of frontness and height, both in relation to the first two formants, and in the articulatory data itself, taking two rather different approaches to the orientation of midsagittal vocal tract images. Qualitatively, it seems clear that there is a good match between F1/F2 space and the articulatory vowel space. A broad quantitative analysis of the place of /u/ in the ultrasound images was also attempted, which, given the partial tongue images which ultrasound provides, was an adaptation of Daniel Jones's conception that the high-point of the tongue provides a convenient basis for quantification. This century-old method proved, perhaps surprisingly, appropriate for the task and the data.

It is quite remarkable that Scottish /u/, which has long been known to be central from an impressionistic point of view, is *as* articulatorily front, and *as* articulatorily low, as we have seen in our socially stratified study of young speakers from the Eastern Central Belt. The acoustic consequences are that /u/'s F2 is front of central, and its F1 is higher than /e/. The mid height of /u/, observed with UTI, is surprising, but may explain some of our anecdotal impressions of mid central /u/ in speakers from the Western Central Belt, e.g., in the media and from personal contact. Further research is needed, and it is interesting that these findings contrast with Brato's (2012) recent detailed examination of Aberdeen speakers, in which he notes that "Sociolinguistic variation patterns in (BOOT) are usually only attested for the front-back dimension… and this study is no exception. There is *some small variation in the height dimension*…" (ibid, 99, emphasis added). In fact, given our findings, we now predict large variation in the height of /u/ and expect that it may well be found to be functioning socially in some locations.

The convenience of treating the Scottish English /u/ as a phonological high back vowel is well and truly over. Phonologically, our ultrasound study suggests that the Scottish vowel system either has a single front rounded vowel, which is non-high, or perhaps the vowel is central. More information is clearly needed, particularly with respect to kinematic behaviour, perceptual boundaries, speaker intuition, and patterns of variation and real-time diachronic change. All these are relevant sources of evidence for the phonological status of /u/, especially in the absence of merger or changes to categorical behaviour. Even with a full empirical set of diagnostics, we doubt there can ever be clarity as to whether and when a vowel changes its phonological label for a given speaker. Yet, in a modular view of phonetics and phonology, such clarity should be expected and must be sought. We suspect that the phonological facts only are unarguably clear, if ever, long after a change has been completed, for an entire community, from a distant diachronic perspective, especially one that focuses only on category identity. This gives the appearance of modularity. In the meantime, we are always in the midst of change, despite the phonological fiction that we are in a period of stability, between changes. If we wish to understand the issues, research into contemporary vowel systems and their properties should focus on articulatory, acoustic, perceptual and of course interactional perspectives within a socially-structured methodology, in order to explain how vowel systems arise, function, and change.

Addresses of the Authors

James M. Scobbie and Eleanor Lawson, CASL (Clinical Audiology, Speech and Language) Research Centre, Queen Margaret University, Edinburgh, EH21 6UU, Scotland, UK <{jscobbie,elawson}@qmu.ac.uk>

Jane Stuart-Smith, School of Critical Studies, University of Glasgow, 4 University Gardens, Glasgow, G12 8QQ, Scotland, UK <Jane.Stuart-Smith@glasgow.ac.uk>

Notes

* The corpus was funded by ESRC (RES-000-22-2032). Initial vowel measures were made by Rochelle Cantwell. Technical and moral support from Steve Cowen and Alan Wrench are very gratefully acknowledged. The paper has benefitted from review by the editors and reviewers Jonathan Harrington and Alexei Kochetov, to whom we are very grateful, as we are to our colleagues and audiences for dis-

cussion. In particular we'd like to thank John Coleman, Joanne Cleland, Olga Gordeeva and Natasha Zharkova, and audiences and colleagues in Grenoble, Lecce and Bolzano.

[1] The cardinal vowel space had been systematised by Daniel Jones in print (and committed to a gramophone recording) in 1917 and popularised in his groundbreaking *An Outline of English Phonetics* (Jones 1918, 1972).

[2] They are viewable online at http://en.wikipedia.org/wiki/File:Cardinal_vowels-Jones_x-ray.jpg.

[3] The four empirical tongue curves are based on the location of a chain of lead plates worn on the midsagittal tongue during the X-ray procedure.

[4] In previous work, we have used both /u/ and /ʉ/ for this vowel in Scottish English. In this paper, since we will present quantitative phonetic data, we opt for /u/ in the main body of the paper, since it is the more abstract and neutral of the two. We return to the phonological identity of the vowel and the choice of label in the conclusion.

[5] As we will see, Scottish English has no /ʊ/.

[6] More accurately, Scottish Standard English, or SSE.

[7] These lexical sets have been grouped together as BOOT by Johnson (1997) for Scottish English, while GOOSE was split into BRUISE and BRUTE by Scobbie et al.(2007) to reflect the quasi-phonemic contrast of the Scottish Vowel Length Rule. BRUTE and FOOT are expected to be homophonous in Scottish English, while BRUISE has a durationally longer vowel in long SVLR contexts (Scobbie et al. 1999a,b). The single word analysed here is SVLR short *boom*, from BRUTE. We can now note in passing that Ferragne & Pellegrino (2010) found that Elgin *hood* i.e. BRUTE was fronter than *who'd* i.e. BRUISE, without the expected duration difference, which confirms our own anecdotal observations on Highland English, and suggests a rather different SLVR system.

[8] Though the internal structure of the tongue is also imaged, it is rarely used quantitatively, but is useful for qualitative analysis.

[9] Scottish English is characterised by vowels that are phonological monophthongs, which in phonetic production in most contexts show relatively flat and unchanging formants in the centre of the vowel. Formant values were extracted at the vowel midpoint (or F2 extremum if that was different) by the first author and at the vowel intensity peak by the third author.

[10] We regret that it will be confusing for many readers that UTI images by tradition are orientated with anterior to the right.

[11] We think this is the first mention in print of the height of Scottish English /u/ being lower-mid (let alone higher-mid) rather than high, as expected. If /u/ is indeed lower in WC speakers than MC ones, this would be compatible with there being a vernacular change from below of lowering in progress. A preliminary real-time comparison of middle-aged and adolescent Glaswegian WC speakers suggests slight lowering and stronger *retraction* of /u/ (with respect to /i/ and /a/) between the 70s and 00s (Stuart-Smith et al. 2012).

[12] Recall one value for /e/ is missing.

[13] This front-of-central value is similar to one that can be calculated for Stuart-Smith's 1997 Glasgow corpus. Averaging the four age-sex groups reported in Scobbie et al. (1999b), after converting to Bark, gives a mean relative frontness of 63% for /u/.

[14] The axes are included to help indicate of how much movement was required to line up the curves, because they were all in the same place before the images were translated..

[15] Thanks to Jonathan Harrington (p.c.) for this point. It is true that it is just an assumption on our part based on traditional descriptions that [ʉ] for /u/ is less com-

mon and natural than [u]. As for Scottish /o/, we note that it may never be possible to show unequivocally that it is a corner vowel, even if we combine sociophonetic data with the sort of detailed perception-production study of German /u/ and /o/ undertaken by Harrington et al. (2011a) – even phonetically natural variation in phonetic realisation can lead to variable, fuzzy or non-deterministic phonologisation. [16] Harrington et al. (2011b) suggest present-day SSBE /ʉ/ does not sound like French /y/ because English /i/ is not as peripheral as French /i/ (and by inference neither is contemporary SSBE /ʉ/ as peripheral as French /y/, on the assumption that the tongue differences between French /i̯y/ and SSBE /i-ʉ/ are broadly similar. Scottish /i/ however does seem quite peripheral, and /e/ (which corresponds to a rather open diphthong in SSBE) likewise is rather high. Scottish /u/, while fronted, is clearly not high and is unlikely to be a lip-rounded version of either of these vowels.

Bibliographical References

Articulate Instruments Ltd. 2008. *Ultrasound Stabilisation Headset Users Manual: Revision 1.13*. Edinburgh: Articulate Instruments Ltd.

Articulate Instruments 2012. *Articulate Assistant Advanced Ultrasound Module User Manual, Revision 2.14*. Edinburgh: Articulate Instruments Ltd. [Associated computer program version 2.13 to 2.14, http://www.articulateinstruments.com]

Baranowski Marcie 2008. The fronting of the back upgliding vowels in Charleston, South Carolina. *Language Variation and Change* 20. 527-551.

Boberg Charles 2011. Reshaping the vowel system: an index of phonetic innovation in Canadian English. *University of Pennsylvania Working Papers in Linguistics* 17, 2 (Article 4).

Boersma Paul & David Weenink 2012. *Praat: Doing phonetics by computer* [Computer program, version 5, http://www.praat.org/]

Brato Thorsten 2012. *A Sociophonetic Study of Aberdeen English: Innovation and Conservatism*. Gießen: Justus-Liebig-Universität Gießen. PhD Dissertation.

Cheshire Jenny, Paul Kerswill, Felicity Fox & Eivind Nessa Torgersen 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15. 151-196.

Cleland Joanne, Alan A. Wrench, James M. Scobbie & Scott Semple 2011. Comparing articulatory images: An MRI / Ultrasound Tongue Image database. In *Proceedings of ISSP'11. IXth International Seminar on Speech Production* (June 20-23, 2011). Montreal, Canada. 163-170.

Docherty Gerard J. 1992. *The timing of voicing in British English obstruents*. Berlin: Foris Publications.

Eremeeva Viktoria & Jane Stuart-Smith 2003. A sociophonetic investigation of the vowels OUT and BIT in Glaswegian. In Solé Maria J., Daniel Recasens & Joaquim Romero (eds.). *Proceedings of the XVth International Congress of Phonetic Sciences* (August 3-9, 2003). Barcelona, Spain. Vol. 3. 1205-1208.

Ferragne Emmanuel & François Pellegrino 2010. Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association* 40. 1-34.

Foulkes Gerard & Gerard Docherty (eds.) 1999. *Urban Voices: Accent Studies in the British Isles*. London: Arnold.

Gendrot Cedric & Martine Adda-Decker 2007. Impact of duration and vowel inventory size on formant values of oral vowels: an automated formant analysis from eight languages. In Trouvain Jürgen & William J. Barry (eds.). *Proceedings of the XVIth International Congress of Phonetic Sciences* (August 6-10, 2007). Saarbrücken. Germany. Vol. 3. 1417-1420.

Gick Brian, Sonya Bird & Ian Wilson 2005. Techniques for field application of lingual ultrasound imaging. *Clinical Linguistics and Phonetics* 19. 503-514.

Gimson Alfred C. 1962. *An Introduction to the Pronunciation of English*. [Revised and updated as Cruttenden Alan (ed.) 2001. *Gimson's Introduction to the Pronunciation of English* (6th ed.). London: Hodder Arnold.]

Hale Mark & Charles Reiss 2000. Substance abuse and dysfunctionalism: current trends in phonology. *Linguistic Inquiry* 31. 157-169.

Harrington Jonathan 2007. Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In Cole Jennifer & José Ignacio Hualde (eds.). *Papers in Laboratory Phonology 9*. Berlin: Mouton de Gruyter. 125-143.

Harrington Jonathan, Felicitas Kleber & Ulrich Reubold 2008. Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America* 123. 2825-2835.

Harrington Jonathan, Phil Hoole, Felicitas Kleber & Ulrich Reubold 2011a. The physiological, acoustic, and perceptual basis of high back vowel fronting: evidence from German tense and lax vowels. *Journal of Phonetics* 39. 121-131.

Harrington Jonathan, Felicitas Kleber & Ulrich Reubold 2011b. The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English. *Journal of the International Phonetic Association* 41. 137-156.

Hueber Thomas, Gérard Chollet, Bruce Denby & Maureen Stone 2008. Acquisition of ultrasound, video and acoustic speech data for a silent-speech interface application. In Sock Rudolph, Susanne Fuchs & Yves Laprie (eds.). *Proceedings of ISSP'08. VIIIth International Seminar on Speech Production* (December 8-12, 2008). Strasbourg, France. 365-368.

Iivonen Antti 1994. A psychoacoustical explanation for the number of major IPA vowels. *Journal of the International Phonetic Association* 24. 73-90.

Janda Richard D. 2003. "Phonologization" as the start of dephoneticization — Or, on sound change and its aftermath: Of extension, generalization, lexicalization, and morphologization. In Joseph Brian D. & Richard D. Janda (eds.). *The Handbook of Historical Linguistics*. Oxford: Blackwell. 401-422.

Johnston Paul 1997. Regional variation. In Jones Charles (ed.). *The Edinburgh History of the Scots Language*. Edinburgh: Edinburgh University Press. 433-513.

JONES Daniel 1917a. *An English pronouncing dictionary*. London: Dent. [Reprinted in COLLINS Beverley & Inger M. MEES (eds.). *Daniel Jones: Collected Works*. Vol. 1-8. London: Routledge.]

JONES Daniel 1917b. Experimental phonetics and its utility to the linguist. *Nature* 100. 96-98.

JONES Daniel 1918. *An Outline of English Phonetics*. Leipzig: Teubner. [Reprinted in COLLINS Beverley & Inger M. MEES (eds.). *Daniel Jones: Collected Works*. Vol. 1-8. London: Routledge.]

JONES Daniel 1972. *An outline of English phonetics* (9th ed.). Cambridge: W. Heffer & Sons Ltd.

JOOS Martin 1948. Acoustic phonetics. *Language* 24. 1-136.

LABOV William 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.

LASS Roger 1989. System shape and the eternal return: front rounded vowels in English. *Folia Linguistica Historica* 10. 163-198.

LAWSON Eleanor, Jane STUART-SMITH & James M. SCOBBIE 2008. Articulatory insights into language variation and change: Preliminary findings from an Ultrasound study of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics* 14. 102-110.

LAWSON Eleanor, James M. SCOBBIE & Jane STUART-SMITH (submitted). The impact of covert articulatory /ɪ/ variants on prerhotic vowels: an ultrasound tongue imaging analysis.

MACAULAY Roger 1977. *Language, social class and education: a Glasgow study*. Edinburgh: University of Edinburgh Press.

MCALLISTER Anne H. 1938. *A Year's Course in Speech Training*. London: University of London Press Ltd.

MESTHRIE Rajend 2010. Socio-phonetics and social change: Deracialisation of the GOOSE vowel in South African English. *Journal of Sociolinguistics* 14. 3-33.

MIELKE Jeff, Adam BAKER, Diana ARCHANGELI & Sumayya S. RACY 2005. Palatron: A Technique for Aligning Ultrasound Images of the Tongue and Palate. In SIDDIQI Daniel & Benjamin V. TUCKER (eds.). *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona*. 14. 97-108.

MILLER Amanda L. & Kenneth B. FINCH 2011. Corrected High Frame Rate Anchored Ultrasound with Software Alignment. *Journal of Speech, Language, and Hearing Research.* 54. 471-486.

OHALA John J. 1981. The listener as a source of sound change. In MASEK Carrie S., Roberta A. HENDRICK & Mary F. MILLER (eds.). *Papers from the Parasession on Language and Behavior.* Chicago: Chicago Linguistic Society. 178-203.

PIERREHUMBERT Janet, Mary BECKMAN & Robert D. LADD 2000. Conceptual foundations of phonology as a laboratory science. In BURTON-ROBERTS Noel, Philip CARR & Gerard J. DOCHERTY (eds.). *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press. 273-303.

SCHÜTZLER Ole 2011. Charting vowel spaces in Edinburgh middle-class speech. *English World-Wide* 32. 24-45.

SCOBBIE James M. 2006. Flexibility in the face of incompatible English VOT systems. In GOLDSTEIN Louis M., Douglas H. WHALEN, Catherine T. BEST

(eds.). *Papers in Laboratory Phonology 8: Varieties of Phonological Competence*. Berlin: Mouton de Gruyter. 367-392.

Scobbie James M. 2007. Interface and overlap in phonetics and phonology. In Ramchand Gillian & Charles Reiss (eds.). *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press. 17-52.

Scobbie James M. 2009. Scottish English sound systems: levels, segments, and transitions. Oral paper presented at *ICLCE3. IIIrd International Conference on the Linguistics of Contemporary English* (July 14-17, 2009). London, UK.

Scobbie James M., Nigel Hewlett & Alice E. Turk 1999a. Standard English in Edinburgh and Glasgow: the Scottish vowel length rule revealed. In Foulkes & Docherty 1999. 230-245.

Scobbie James M., Alice E. Turk & Nigel Hewlett 1999b. Morphemes, phonetics and lexical items: the case of the Scottish Vowel Length Rule. In Ohala John J., Yoko Haswgawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey (eds.). *Proceedings of the XIVth International Congress of Phonetic Sciences* (August 1-7, 1999). San Francisco, California. Vol. 2. 1617-1620.

Scobbie James M. & Jane Stuart-Smith 2008. Quasi-phonemic contrast and the fuzzy inventory: examples from Scottish English. In Avery Peter, B. Elan Dresher & Keren Rice (eds.). *Contrast in Phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter. 87-113.

Scobbie James M., Jane Stuart-Smith & Eleanor Lawson 2008a. Looking variation and change in the mouth: developing the sociolinguistic potential of Ultrasound Tongue Imaging. Final Report for ESRC Project RES-000-22-2032.

Scobbie James M., A. Alan Wrench & Marietta van der Linden 2008b. *Head-Probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement*. In Sock Rudolph, Susanne Fuchs and Yves Laprie (eds.). *Proceedings of ISSP'08. VIIIth International Seminar on Speech Production* (December 8-12, 2008). Strasbourg. France. 373-376.

Scobbie James M., Eleanor Lawson, Steven Cowen, Joanne Cleland & Alan A. Wrench 2011. A common co-ordinate system for mid-sagittal articulatory measurement. CASL Working Papers, Queen Margaret University, WP-20.

Scobbie James M. & Jane Stuart-Smith 2012. Socially stratified sampling in laboratory-based phonological experimentation. In Cohn Abigail C., Cécile Fougeron & Marie K. Huffman (eds.). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. 607-621.

Scobbie James M., Joan Ma & Joanna White 2012. Expanding the vowel space: a pilot study of Lombard speech with ultrasound tongue imaging and video lip imaging. CASL Working Papers, Queen Margaret University, WP-21.

Speitel Hans-Henning & Paul Johnston 1983. *A Sociolinguistic Investigation of Edinburgh Speech*. ESRC End of Grant Report.

Stone Maureen L. & Edward P. Davis 2005. A head and transducer support system for making ultrasound images of tongue/jaw movement. *Journal of the Acoustical Society of America* 98. 3107-3112.

STUART-SMITH Jane 1999. Glasgow: accent and voice quality. In FOULKES & DOCHERTY 1999. 203-223.

STUART-SMITH Jane 2003. The phonology of Modern Urban Scots. In CORBETT John, Derrick MCCLURE & Jane STUART-SMITH (eds.). *The Edinburgh Companion to Scots*. Edinburgh: Edinburgh University Press. 110-137.

STUART-SMITH Jane, Eleanor LAWSON & James M. SCOBBIE (submitted). Derhoticisation in Scottish English: lessons we can learn from sociophonetic data.

STUART-SMITH Jane, Tamara RATHCKE & Claire TIMMINS 2012. Best BOOT forward? A real-time study of Scottish English /u/ over forty years. Poster presented at *IInd Workshop on Sound Change* (May 2-4, 2012). Kloster Seeon, Germany.

TRAUNMÜLLER Hartmut 1990. Analytical expressions for the tonotopic sensory scale *Journal of the Acoustical Society of America* 88. 97-100.

WELLS John Christopher 1982. *Accents of English*. Vol. 1-3. Cambridge: Cambridge University Press.

WHALEN Doug H., Khalil ISKAROUS, Mark K. TIEDE, David J. OSTRY, Heike LEHNERT-LEHOUILLIER, Eric VATIKIOTIS-BATESON & Donald S. HAILEY 2005. The Haskins Optically Corrected Ultrasound System (HOCUS). *Journal of Speech Language and Hearing Research* 48. 543-553.

WRENCH Alan A. & James M. SCOBBIE 2006. Spatio-temporal inaccuracies of video-based ultrasound images of the tongue. In YEHIA Hani Camille, Didier DEMOLIN & Rafael LABOISSIÈRE (eds.). *Proceedings of ISSP'06. VIth International Seminar on Speech Production* (December 13-15, 2006). Ubatuba, Brazil. 451-458.

WRENCH Alan A. & James M. SCOBBIE 2008. High-speed cineloop ultrasound vs. video ultrasound tongue imaging: comparison of front and back lingual gesture location and relative timing. In SOCK Rudolph, Susanne FUCHS and Yves LAPRIE (eds.). *Proceedings of ISSP'08. VIIIth International Seminar on Speech Production* (December 8-12, 2008). Strasbourg, France. 57-60.

WRENCH Alan A. & James M. SCOBBIE 2011. Very high frame rate ultrasound tongue imaging. In *Proceedings of ISSP'11. IXth International Seminar on Speech Production* (June 20-23, 2011). Montreal, Canada. 155-162.

WRENCH Alan A., Joanne CLELAND & James M. SCOBBIE 2011. An ultrasound protocol for comparing tongue contours: upright vs. supine. In LEE Wai-Sum & Eric ZEE (eds.). *Proceedings of the XVIIth International Congress of Phonetic Sciences* (August 17-21, 2011). Hong Kong, China. 2161-2164.

ZHARKOVA Natalia 2007. *An investigation of coarticulation resistance in speech production using ultrasound*. Edinburgh: Queen Margaret University. PhD dissertation.

ZHARKOVA Natalia, Nigel HEWLETT & William J. HARDCASTLE 2011. Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control* 15. 118-140.