Fronting of Southern British English high-back vowels in articulation and acoustics

DOI:
10.1121/1.4991010

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
The Journal of the Acoustical Society of America

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Download date: 20. Sep. 2020
Fronting of Southern British English high-back vowels
in articulation and acoustics

Patrycja Strycharczuk\textsuperscript{1} and James M. Scobbie\textsuperscript{2}

\textsuperscript{1} Linguistics and English Language, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
e-mail: pstrycharczuk@gmail.com

\textsuperscript{2} Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, Musselburgh, EH21 6UU, United Kingdom

Acknowledgments: We wish to thank the speakers for participating in our study, Steve Cowen for assistance with the recordings, and Alan Wrench for help with the ultrasound system. We thank the editor, Cynthia Clopper, and two anonymous reviewers for their comments and suggestions. Any remaining errors are our own. The research reported in this paper was supported by a British Academy Postdoctoral Fellowship PDF/pf130029 to the first author.
Abstract

The fronting of the two high-back vowels /u:/ and /u/ in Southern British English is very well documented, but mainly in the acoustic domain. This paper presents articulatory (ultrasound) data, comparing the relative tongue position of these vowels in fronting and non-fronting consonantal contexts, i.e. preceding a coronal consonant (food, foot) and preceding a coda /l/ (fool, full). Particular attention is paid to the comparison between articulatory results and corresponding acoustic measurements of F2 in both vowels. Results show that the average differences between food and foot and their dynamic profiles are similar in articulation and acoustics. In /u:l/ sequences (fool), tongue position is more advanced than could be inferred from its low F2. In addition, even though the tongues position in fool and full is clearly distinct, there is no comparable corresponding difference in F2. This suggests that the common articulatory metaphor that characterises F2 increase as fronting must be used cautiously. In the case of English high-back vowel fronting, special attention must be paid to the flanking consonants when estimating vowel distances. This paper also provides specific recommendations for recording and analysing ultrasound data in research on vowel variation and change.
I. Introduction

Fronting processes affecting high-back vowels /u:/ and /u/ (lexical sets of GOOSE and FOOT; Wells 1982) are among some of the most striking sound changes observed for English in recent years. Both vowels have shifted in numerous English dialects, and the change has been especially dramatic in Southern British English (SBE; see Ferragne & Pellegrino 2010 and Lawson et al. 2015 for cross-dialectal comparisons). Furthermore, fronting is subject to numerous segmental, structural and sociolinguistic constraints, giving us the opportunity to study the emergence of complex conditioning effects in sound change. Previous research shows that /u:/- and /u/-fronting is blocked, or at least seriously limited, before a coda /l/ (Kleber et al., 2011). Similar limitations in the degree of fronting may apply as a result of interaction between segmental factors (the presence of a following /l/) and higher-level linguistic structure, including morphological information. For instance the /u:/ vowels in morphologically-complex words like fool-ing are characterised by increased degree of lingual retraction in comparison to monomorphemic words like hula, although the tongue may be not as far back as in fool (Strycharczuk & Scobbie, 2015). In contrast, the /u/ vowels in words like full-ish (morphologically-complex) and bully (monomorphemic) undergo fronting to a similar extent. Such different gradient patterns of allophony highlight the need for analytic methods that allow us to quantify the degree of vowel fronting across speakers and across
speaker groups.

In this paper, we compare articulatory and acoustic methods in quantifying vowel fronting. We propose a replicable articulatory method which compares the degree of tongue displacement, imaged using ultrasound, along the occlusal plane. The occlusal plane is a consistent cross-speaker (and cross-session) reference vector commonly used in EMA research. Scobbie et al. (2011) proposed its utility for ultrasound data and developed a procedure for imaging it. We relate our measure of articulatory frontness to F2, an equally replicable measure which is known to be an acoustic correlate of tongue position. Given the dynamic nature of the coarticulatory blocking of fronting being investigated, both measures are analysed dynamically as they vary over the time-course of the segment rather than just as segmental targets. The results confirm that there is a robust relationship between horizontal tongue displacement and F2 increase, which validates our articulatory measure. The relationship is, as expected, not linear. Such non-linearity reinforces the advantages provided by acquiring direct articulatory evidence in experimental studies of ongoing sound change.

A. Acoustic and perceptual studies of /u:/- and /ʊ/-fronting

The diachronic change in the acoustic quality of /u:/ and /ʊ/ vowels in SBE is apparent from the comparison of different studies carried out over the last 5 decades. As summarised in Table 1, consecutive studies generally report increasingly high F2 values for the /u:/ and /ʊ/ vowels, whereas the F2 values for /i:/ remain more stable. (The relatively low F2 value
for /u/ produced by male speakers in Williams & Escudero 2014 is an exception to this trend.)

Similar changes in F2 values for high-back vowels are reported by a number of apparent time studies. Hawkins & Midgley (2005) compared formant frequencies in 20 male speakers representing four different age groups: 20–25, 35–40, 50–55, and 65–73 (speaker ages as of 2001). The F2 results for /u:/ /u/ and /i:/ are summarised in Table 2. F2 for both /u:/ and /u/ increased in apparent time, albeit at a different pace. The rise in F2 for /u:/ was steady across all age groups, whereas for /u/, there was only a notable F2 increase for youngest speakers compared to the other age groups.

The findings by Hawkins & Midgley (2005) concerning /u/-fronting are also corroborated by acoustic and perceptual results reported in Harrington et al. (2008). Harrington et al. 2008 find that the F2 values for the /u:/ vowel in younger speakers (18–20) overlapped partially with the F2 values for /i:/ category, whereas for older speakers (50+), the two vowel categories were clearly separate along the F2 dimension. The same study shows a perceptual shift in discrimination of the /u:/ and /i:/ vowels, where the younger speakers’ perceptual /u:/ category became closer to /i:/.

Kleber et al. (2011) look in more detail at the /u/ vowel by analysing the degree of coarticulation affecting it, using a combination of production and perception studies. They report a shift in production, where /o/ becomes increasingly front in most environments,
but not in the context of back consonants (e.g. in wool\textsuperscript{1}). As a result, the F2 distance between vowels in words like *hood* and *wool* increases in apparent time. However, there was no systematic evidence that this shift in production was accompanied by a shift in perception that would involve younger listeners compensating less for coarticulatory effects compared to older listeners.

**B. Articulation of /u:/ and /u/**

Observations concerning acoustic changes in the qualities of the /u:/ and /u/ vowels have largely been interpreted either in terms of acoustics (increase of F2), or in terms of articulation where the F2 rise results from a primary change in the relative locations of constrictions in the vocal tract involved in the production of these vowels. Generally the tongue is seen as the primary factor, as is implicit in the use of the term ‘fronting’ when referring to ‘F2 increase’, with coronal lingual consonants being seen as coarticulatory actuators of the sound change. This articulatory perspective that the fronting is instigated by coarticulatory influences from flanking coronal consonants is explicitly elaborated by the proposals concerning the causes of high-back vowel fronting by Harrington et al. (2008) and Kleber et al. (2011). A central factor involved in the gradient fronting of the vowel target is the misalignment between production and perception, in which coarticulation pressures cease to be perceptually ‘undone’ by listeners.\textsuperscript{2} As a result, listeners cease to attribute the front variants to the presence of a coarticulation trigger, and extend the use of such variants also to the contexts
where fronting is not phonetically motivated. This reinterpretation hypothesis, which builds on the work by Ohala (Ohala, 1981, 1990), has at its heart the assumption that high-back vowel fronting is lingual: it is triggered by lingual coarticulation, and so it is mainly realised in the domain of tongue position.

This lingual hypothesis is supported by the articulatory data in Harrington et al. (2011). Electromagnetic articulography (EMA) data from 5 young speakers of SBE show that the average tongue position for /u:/ is very front, patterning closely with the kit vowel. The average tongue position for /ʊ/ is also relatively front, clearly more front than that for /ʌ/ or /ɒ/, although not as front as that for /u:/.

Harrington et al. (2011) also analyse lip movement data obtained in the same experiment, showing that there is similar lip protrusion in /u:/ and /ʊ/ as in the baseline /ɔ:/ (thought), whereas the /i:/ vowel is distinct in showing considerably less lip protrusion. The argument that the degree of lip protrusion in /u:/ has not been recently reduced is also supported by the results of an audiovisual perception experiment, where young speakers’ /u:/ vowel was judged to be produced with rounded lips. The findings confirming the presence of lip rounding in fronted variants of /u:/ and /ʊ/ are crucial, since the acoustic effect of F2 increase can be achieved through lingual fronting, lip unrounding, or a combination of the two.

In another study of lingual fronting in SBE, Lawson et al. (2015) ask whether increased F2 in the production of the goose vowel is indeed achieved by tongue fronting across different
accents of English, or whether lip unrounding is a contributing factor. Based on ultrasound data from 20 speakers from a variety of different locations in the British Isles, Lawson et al. show that there are dialectal differences concerning the relative contribution of tongue and lips to the production of acoustically front *goose*. Speakers from England tend to produce this vowel by means of lingual fronting. This is different from Scottish speakers, who achieve increase in F2 through lip spreading. Furthermore, the English speakers in the study seem to form a continuum of lingual fronting, with more front vowels produced by the speakers of southern English dialects (e.g. Kent, Southampton) in comparison to speakers from the north of England (e.g. Newcastle, Yorkshire, Manchester). However, since most dialects in the study were represented by one speaker, more systematic research will be needed to see if there are genuinely dialectal patterns rather than individual differences.

While the articulatory studies by Harrington et al. (2011) and Lawson et al. (2015) are mainly concerned with illuminating the articulatory mechanisms involved in /u:/- and /ʊ/-fronting, we may ask what methodological advances they made in studying these processes, and what further advances are needed. As ongoing fronting seems to correlate with changes in tongue position, is it the case that acoustic data can tell us all there is to know about high-back vowel fronting? A positive answer could have far-reaching practical implications. Audio recordings are cheap and easy to acquire, and they also include pre-existing data, e.g. from historical sources, allowing us to study sound change in real time. Furthermore, there
is a rich research tradition of analysing and normalising acoustic data, which provides the foundations for comparing results from different studies. However, in spite of what we know so far about lingual fronting in /u:/ and /ʊ/, we cannot safely conclude that acoustic data provide a sufficiently good window into the articulatory behaviour underlying the production of different /u:/ and /ʊ/ variants.

The complexity of acoustics-to-articulation mapping is a central theme of for a body of experimental and modelling work (e.g. Richmond et al. 2003; Parrell 2010; Ananthakrishnan 2011). For the specific case of /u:/ and /ʊ/, the main complication concerns interpreting acoustic distances in articulatory terms. Even if we eliminate lip rounding as a major factor contributing to F2 fluctuations in the case of diachronic changes affecting /u:/ and /ʊ/, the relationship between tongue position and F2 is not a linear one. Small differences in tongue position may have a different effect on F2, depending on the location of the constriction. Such non-linearities have long been recognised, and they have indeed formed the foundation for the Quantal Theory of speech (Stevens et al., 1986; Stevens & Keyser, 1989, 2010; Keyser & Stevens, 2006), where areas of stability in articulation-acoustics mappings are selected as loci for contrastive linguistic units. A related issue, and one which is not easily modelled from the theoretical point of view, concerns individual differences between speakers. Individuals vary in their vocal tract morphology, concerning both tongue size and palate shape, and therefore they may use different articulatory strategies to produce a similar acoustic effect.
Previous work on compensatory articulation shows that adjustments in different articulators may offset each other, such that articulatory variability is reduced in the corresponding acoustic dimension (Maeda, 1990; Guenther et al., 1999). In the context of dialect variation, the complex nature of articulation-acoustics mappings prompts us to consider the need for a wider use of direct articulatory evidence in dialectal studies. This, however, brings on the question of how to normalise articulatory information from different speakers. We consider some of the challenges involved in this task in Section C below.

**C. Articulatory measures and normalisation**

Anatomical differences introduce a difficulty in comparing speech data across speakers. This issue has long been recognised in studies of vowel formants, which are known to be systematically affected by the length of the vocal tract. A large body of research has been dedicated to addressing this issue, and various normalisation methods have been proposed to make formant-based vowel system comparisons more reliable (see Kohn & Farrington 2012 for a recent overview). In the articulatory domain, anatomical differences are potentially an even stronger influence, but since articulatory variation is a relatively young field of study, previous work on normalisation is somewhat scarce. The data we present in this study come from ultrasound tongue imaging. A popular technique in analysing such data involves statistical comparison of average smoothed tongue contours (Davidson, 2006). However, such comparisons are only possible within speaker, as tongue shapes and tongue lengths vary between
different individuals. Tongue musculature and palate shape vary too, and all combine to affect both single time point images of the tongue surface and dynamic patterns of change.

Existing attempts to normalise single-time point tongue surface curves extracted from individual ultrasound images often rely on reducing the dimensions in the image to a specified representative point which acts as a proxy for quantifying displacement in a region of interest. Such an approach is taken by Scobbie et al. (2012), who quantify articulatory distances based on a replicable reference plane defined by a common tangent linking /i/ and /o/ vowels. The articulatory position of other vowels is then measured relative to this plane, based on the location of the closest approximation of their tongue surface curve to the reference plane (see Figure 1 for illustration). The motivation for using the /i/-/o/ plane as a reference is that these two vowels represent the upper corners of the vowel space: /i/ is the highest front and /o/ is the highest back vowel in Scottish English, the variety studied by Scobbie et al. The horizontal degree of fronting is then normalised to the distance between /i/ and /o/. The authors note, however, that the occlusal plane might have been a more suitable basis for a replicable definition of horizontality, had it been available.

Building on this study, Lawson et al. (2015) study the position and height of the /u:/ vowel across a number of English dialects, using ultrasound recordings. Lawson et al. use a /i/-/w/ plane as reference, since most accents of English lack a monophthongal /o/, and /w/ is assumed to be a suitably stable high-back reference point. They again quantify tongue
position and tongue height for the /u:/ vowel via the horizontal and vertical distances of the closest approximation of the tongue to the anchor vowel /i:/ in the reference plane. The highest point on the tongue as a proxy for vowel place of articulation is also used by Bennett et al. (2011) and Noiray et al. (2014).

The highest point of the tongue’s curving midsagittal surface during the articulation of a vowel, relative to a definition of a horizontal reference plane and to other vowels, as used in the above studies, is intended to represent the location of a vocalic constriction. Another approach to capturing constriction is based on movement, and it involves using the intersection of the tongue with a vector where the greatest degree of articulatory movement is observed, or expected (Gick et al., 2006; Rastadmehr et al., 2008; Strycharczuk & Scobbie, 2015).

In this paper, we propose an articulatory measure inspired by the studies cited above, but modified to suit our research questions, which concern vowel targets and the contextual effect of flanking consonants. We are primarily interested in the dynamically changing degree of articulatory fronting/retraction though the rime. We are also interested in the location of the lower dorsal to root areas of the tongue surface of the vowels; those areas are relevant to the darkness of /l/. Measures of root retraction are not usually presented, partly because they can be expected to correlate closely with anterior constrictions in vowels, but more importantly, they cannot be detected with EMA. However, since /l/ does not have a canonical
single lingual gesture, unlike /u:/ and /u/, and since root retraction is likely to be crucial to the effect of /l/ on these adjacent vowels, we need to measure it.

We track the upper tongue root in the upper pharynx or uvular-velar area. Visual inspection of the articulatory movements involved suggests that the occlusal plane is a suitable reference vector also for this area of the tongue. It lets us capture the backwards or forwards movement of the tongue root in a comparable way across speakers, by recording the anteriority on the vector of the location where the tongue root crosses it. We also use this plane to express relative distances in the position of the vowels. In order to make such distances comparable across speakers, we normalise them to the reference vowel, /i:/, following Scobie et al. (2012) and Lawson et al. (2015).

D. Research aims and research questions

This paper sets out to quantify the degree of /u/- and /u/- fronting in SBE in two domains: articulatory and acoustic. Our main concern is effect sizes: we wish to see whether the same relative degree of fronting can be seen for specific vowels in both domains. We look at two consonantal contexts: pre-coronal stop and pre-/l/. One of the reasons for studying these two contexts is that they allow us to compare the extremes of any influence on fronting: it is expected to be enhanced before coronals, and limited before tautomorphemic, tautosyllabic /l/ (see Section A above). We consider this segmental interaction to address a more general question of how well F2 correlates with articulatory position, depending on the vowel and
depending on the presence or absence of a following /l/.

II. Method

A. Test items

The test items were 8 monosyllabic words containing the /u:/ or /o/ vowel preceded by a non-lingual consonant (labial or glottal). This choice of preceding consonant was made in order to minimise any progressive coarticulatory effects on the vowel, and so to isolate the effect of the following consonant. The final consonant was either /t/ or /d/, representing a context for fronting, or /l/, where fronting is expected to be blocked/limited. A summary of the test items is in Table 3. As a baseline representing the /i:/ vowel, we used the test item heap.

B. Speakers

The speakers were 10 females in two age groups who had been born and grown up in the south of England or English Midlands. They all self-identified as speakers of Southern British English, though they were not necessarily speakers of the standard variety (Received Pronunciation). What they all had in common was a Southern English vowel system, which has some salient differences from Northern Englishes, such as the /a:/ vowel for the BATH lexical set, and a FOOT-STRUT split (see Wells 1982 for a description of some typical southern features, and Williams & Escudero 2014 for recent acoustic results comparing vowels
in Southern and Northern English). The younger speakers were between 20 and 25 years of age (mean age=22.6), and the older speakers were between 45 and 62 years of age (mean age=56). These speakers were a subset of 20 recorded for a larger study. They were selected for analysis in the present study due to high quality of their ultrasound data (clear image), which was crucial to the dynamic articulatory analysis.

C. Procedure

In the experiment, we recorded the ultrasound image of the speakers’ tongue and time-aligned audio signal. The synchronisation was controlled by the Articulate Assistant Advanced software (Articulate Instruments Ltd, 2014). The ultrasound data were acquired using a high-speed Sonix RP system (Frame Rate = 121.5 fps, Scanlines = 63, Pixels per Scanline = 412, Field of Vision = 134.9°, Pixel offset = 51, Depth = 80 mm). The audio data were captured using a lavalier Audio-Technica AT803 condenser microphone. The audio data were sampled at 22 kHz. At the beginning of the experiment, a stabilisation head set was fitted on the participant’s head, in order to minimise the movement of the ultrasound probe throughout the recording (Articulate Instruments Ltd, 2008). The probe was positioned under the participant’s chin, aligned to the midsagittal plane. Following that, the occlusal plane was imaged, by asking the participant to bite on a flat piece of plastic (a bite plate), and press her tongue up against it (Scobbie et al., 2011). Following that, the participant was asked to swallow, so that we could obtain an image of the hard palate (Epstein & Stone, 2005).
During the recording, the speakers sat in front of a computer screen, and read the experimental prompts which appeared one by one. The prompts contained the test items embedded in a standard carrier phrase *Say X five times*. In addition to the test items described in Section A, we also included a range of words containing the test vowels, /uː/ or /uː/, in a variety of morphosyntactic contexts (see Strycharczuk & Scobbie (2015) for an analysis of this part of the corpus). The participants read four repetitions of the experimental material. Altogether, each participant read 98 items. Of these, we analysed 36 items per participant: 4 repetitions of the 8 test items and 4 repetition of one baseline item. Each recording lasted between 20 and 30 minutes, which is a suitable limit for this kind of study due to potential discomfort from the headset for some participants.

D. Segmentation

The audio data were exported as wav files from AAA, and automatically segmented using the University of Pennsylvania Forced Aligner (FAVE, Rosenfelder et al. 2011). The boundaries were then checked by the first author, and corrected as necessary; only two are used here for each item. The left edge of the measurement domain was defined as the onset of the vowel (based on the increase in amplitude, the onset of formant structure, and the onset of voicing). The offset of voicing was taken as the right edge of the measurement domain. In the tokens containing final /l/, the lateral was included in the domain along with the vowel. This non-segmental approach was motivated by the fact that coarticularatory dynamics blend
a vowel and a coda /l/ in ways that make any abrupt segmentation exceedingly arbitrary both from the articulatory and the acoustic point of view. This is particularly true in our case, which involves high-back vowels, /u:/ and /u/, followed by dark /l/. These vowels are acoustically so similar to the final /l/ that we were unable to define a set of consistently identifiable segmentation criteria that would allow us to identify the left edge of the /l/.

E. Articulatory measure

The articulatory measure that we developed for these data captures the displacement of the tongue in the occlusal plane. We analysed such displacement over time, throughout the measurement domain (the vocalic portion of the rime, see above). As the first step in the analysis, we traced the tongue contour throughout the vocalic portion, using a semi-automatic tracker implemented in AAA version 2.16.13. We then exported up to 42 tongue contour coordinates (at an equal angled increment) at 10% time intervals of the vocalic portion. As part of this export, the tongue contour data were rotated in the occlusal plane, in order to standardise the rotation of the data across the participants. Figure 2 shows the rotation of an example token for speaker YF8. All the subsequent analysis of tongue contour data was carried out in R (R Development Core Team, 2005).

Figure 3 shows average rotated tongue contours for an example speaker YF1 at a single selected time point: in this case, the acoustic onset of the vowel. Rotating the X-axis to the occlusal plane means the figure directly reflects the conventional EMA approach to
horizontality (consistent across speakers), and displays our quantified measures of tongue position in terms of the X coordinates of the crossing point of any part of the upper root of the tongue on the occlusal plane. According to this measure, for speaker YF1 at the acoustic onset of the vowel, the tongue retraction increases from the most front *food*, via *foot* and *fool*, which are retracted to a similar degree, to the most retracted *full*.

The origin of the coordinate frame, as shown in Figure 3 corresponds to the origin of the rotation. As this varies across the participants, the raw coordinate values are not comparable across the participants. Therefore, we relocated and rescaled the retraction measurements within speaker, also using the reference values for the vowel /i/. Using /i/ as a reference has previously been proposed by Scobbie et al. (2012) and Lawson et al. (2015), and it draws on the idea that the /i/ constriction represents a corner for the speaker’s vowel space in this variety of English (and indeed, most languages). For each production of the /i/, we extracted the X-coordinates of the /i/ constriction point (the highest point of the tongue). An illustration of the relevant measurements for /u:/ and /i/ are in Figure 4. We then pooled those values with the measurements of retraction for our test items (dynamic contours for /u:/ and /u:/), and we scaled those measurements within each speaker using z-score normalisation. It would have been possible to directly measure the distance between the retraction point for /u:/ and the reference point for /i/. However, we suspect that such distances are sensitive to anatomical differences between speakers, where distances are over-
all greater for speakers with larger oral cavities, whereas scaling normalises for both probe placement and anatomical differences between individuals.

**F. Acoustic measure**

We used raw F2 (in Hz) as our main acoustic correlate of acoustic fronting/retraction, in line with previous acoustic studies on /u:/- and /u/- fronting (see Section A). F2 was measured dynamically, at 10% intervals throughout the vocalic portion, to correspond with our dynamic articulatory measurements. We used a modified script by Remijsen (2004) for the measurements. Formant tracking was monitored in real time throughout the measurements, and the formant settings were optimised for each speaker to ensure accuracy.

The F2 measurements were normalised analogically to the articulatory normalisation (see Section E above), i.e. through scaling F2 measurements for /u:/, /u/ and /i:/ within speaker.

**G. Statistical analysis**

The statistical analysis was carried out using a series of General Additive Mixed Models (GAMMs, Wood 2006; van Rij et al. 2015). This method was chosen for its capacity to accommodate non-linear effects, of the kind expected to occur in our data, without pre-specifying the shape of the curve. The specification for key models is discussed in Section III below.
III. Results

The articulatory and acoustic fronting measures were analysed using generalised additive mixed models. For articulatory fronting, the dependent variable was the $z$-score of the fronting/retraction measure, whereas for acoustic fronting, the dependent variable was normalised F2, measured in Hz. Both variables were analysed as a function of normalised time within context (food, foot, fool or full). The random effect structure included random smooths for temporal change in fronting by context within speaker, and random intercepts for trial and item. The random smooths were included to accommodate the possibility that the progress of fronting/retraction develops differently, depending on the speaker. For both models, including such individual variation significantly improved the model fit at $p<.001$ level, which confirms there are indeed individual differences with respect to degree of fronting, and in how fronting develops over time.

The average main results for both models are presented in Figure 5. The left panel shows the average smooths for the degree of articulatory fronting in normalised time. As expected, we find more fronting in food compared to fool, and more in foot compared to full. Also, the food~fool difference within the GOOSE vowel is noticeably larger than the foot~full difference in FOOT, which is consistent with the generalisation that /ʊ/-fronting is a more recent, less advanced sound change than /uː/-fronting (Hawkins & Midgley, 2005; Kleber et al., 2011).
To an extent, similar generalisations emerge from the model of acoustic fronting, as illustrated in the right panel of Figure 5. What is particularly interesting, however, is the comparison of fronting degrees in the two related phonetic domains: articulation and acoustics. Whereas the relative positions of vowels in *food, foot* and *full* are comparable, the vowel in *fool* is more front in articulatory space than it is in acoustic space. The F2 trajectories for *fool* and *full* are very similar indeed, but the two vowels are clearly distinct with respect to articulatory retraction, with *fool* being more advanced. What is also striking is the dynamic difference between articulatory fronting contours and F2 contours for *fool* and *full*. The F2 contours show relatively more movement: the F2 decreases rapidly to begin with, reaching a dip ca. halfway through the vowel + /l/ sequence, where it begins to rise again.

In order to capture the observations about the articulation-acoustics relationship more systematically, we fitted another model predicting the degree of articulatory fronting/retraction. We kept the same random and fixed predictors as in the previous model: a fixed effect of normalised time by context, a random smooth for normalised time within speaker, and random intercepts for trial and item. In addition, we included another fixed effect, that of normalised F2 by context, in order to explore the relationship between acoustic and articulatory fronting, and how it varied depending on the vowel and the consonantal context. Results illustrating this relationship are plotted in Figure 6.
The results are in line with some of the observations made above, based on comparing the articulatory and the acoustic model. To an extent, we see a correlation between lingual fronting and F2 raising, insofar as the most retraced vowels (fool, full) also show the relatively lowest F2. However, there is a difference in lingual retraction between fool and full, which is not reflected in the F2. Furthermore, the strength of the correlation between F2 and lingual fronting is different within each vowel context. Three contexts, fool, full and foot, seem to show some variation in the F2 dimension that corresponds only weakly to the variation in lingual position among these more retracted vowels.

IV. Discussion

One of the goals of this paper was to provide an articulatory perspective on the contextual /u:/- and /u/- fronting in English, in particular the SBE accent. We were especially interested in the relative tongue position for both vowels in the two contexts studied. Our results largely confirm the hypotheses concerning the effect of following consonant, where a following /l/ blocks /u:/- and /u/- fronting. We also find that the /u:/- vowel is more advanced than the /u/- vowel within the same consonantal context. What is striking, is that the back allophone of /u:/ (in fool) has a comparable articulatory position to the front allophone of /u/ (in foot), whereas in acoustics, fool shows considerable F2 lowering, and the F2 values for fool are similar to the F2 values for the articulatorily more retracted full.

In general, we can expect F2 lowering in the presence of a following /l/, especially
when /l/ is dark. A relationship between /l/-darkening an F2 lowering has been observed by a number of studies looking into /l/ in English (Ladefoged & Maddieson, 1996; Carter, 2002, 2003; Carter & Local, 2007; Hawkins & Nguyen, 2004). But why would this affect in fool in particular? If the effect of a following /l/ was simply additive, we would expect a similar degree of F2 lowering in full, and we would expect the articulatory difference between fool and full to be preserved in the acoustics. That is not the case: fool and full are almost merged along the F2 dimension, with confidence intervals between average smooths overlapping considerably (see right panel of Figure 5).

The fact that the vowel in fool has a more advanced tongue position compared to full could be reconciled with the acoustic similarities in F2 between the two vowels if the lip rounding of the two vowels differed. Since both lingual retraction and lip rounding may lower the F2, we may expect degree of articulatory compensation between these two articulators. Specifically, if there were relatively more lip rounding in /u:/ compared to /u/, we would expect the F2 lowering we find in food and fool, compared to their articulatory position. Note, however, that the data in Harrington et al. (2011) show that lip protrusion is similar in /u:/ and /u/ in SBE. It is possible that our speakers have a consistent difference, unlike those of Harrington et al., or that there is a rounding difference specific to a pre-/l/ environment. In the absence of specific data, however, these possibilities remain speculative.

While some differences in lingual retraction are underestimated by the F2, we also
see the opposite, as some variation along the F2 dimension is not due to tongue position. Comparing across different contexts, we can generalise that the more retracted vowels also have the lowest F2. Within each context, however, there is a degree of F2 variation that is not correlated with the lingual retraction (Figure 6). We can presume that this variation is caused by other articulatory adjustments.

Taken together, our results compel us to recommend caution when offering articulatory interpretations of acoustic differences affecting high-back vowels followed by /l/, and of specific F2 values, particularly when it comes to small effects or null results. Null acoustic results especially can be somewhat uninformative on their own, as they may conceal a fairly robust covert articulatory difference (such is the case with fool and full in our data, which have similar F2, but differ considerably in their articulatory position). Such non-linearities between articulation and acoustics also problematise potential considerations of categoricity and gradience: categorical shifts in one of those domains may translate into gradient shifts in the other (see also Section I.B and the work on Quantal Theory cited there).

In a broader perspective, our results reinforce the need to supplement acoustic studies on vocalic sound changes with articulatory analysis. Related to this, more methodological work is necessary, as further development of variationist articulatory work crucially depends on development of relevant experimental protocols and successful normalisation techniques for articulatory data. With respect to experimental procedure, two steps are recommended
for ensuring that data are comparable: probe stabilisation and orientation. An easy way to achieve probe stabilisation, also outside of a laboratory, is through the use of a probe holder (Zharkova et al., 2015; Derrick et al., 2015). Standardisation regarding the rotation of the data can be attained by imaging the occlusal plane during the recording, by asking the speaker to bite on a piece of flat plastic (Scobbie et al. 2011, see also Section II.C).

Assuming that experimental procedures have been followed to ensure data comparability, further issues to consider are choosing the relevant measure, and standardising such a measure across speakers. The former has been extensively addressed by recent work reviewed in Section I.C. Our extension to the existing methods is the proposal to use the intersection between the occlusal plane and the tongue root surface to characterise the position of specific vowels and consonants. This method is relatively straightforward, and inherently dynamic, as it allows us to capture not only tongue position, but also tongue displacement.

An area that requires further validation is standardising articulatory measurements. Raw distance measures between vowels are sensitive to anatomical differences between speakers: such distances are expected to be greater for speakers with relatively larger vowel spaces. As a way of standardising distance measures, we used measurements of a reference vowel, /i/, and z-score normalisation. This approach is inspired by existing techniques in acoustic analysis, such as Lobanov normalisation (Lobanov, 1971). However, since we did not record a full set of vowels, we are unable to verify how well the normalisation works across speak-
ers. We believe that the normalisation was sufficient for our purposes, since we focused on within-speaker comparisons. However, further validation work in this area is recommended before comparisons between speaker groups (e.g. dialect, sex, or age-based comparisons) are made.

Our final comment concerns vowel dynamics. In this work, we pursued a dynamic approach to vowel measurements, as previously done by Harrington et al. (2008), Williams & Escudero (2014) and Docherty et al. (2015), *inter alia*. There is a growing consensus in the field that dynamic measurements of vowels provide a more complete view of vowel characteristics, and they avoid a necessarily arbitrary choice of selecting a specific time point where the measurements are taken. Our results also confirm that dynamics matters – both tongue position and F2 in our data undergo non-linear changes in time (see Figure 5). Consequently, conducting measurements at a selected time point would considerably affect the measured distances between specific vowels, and in some cases, the time point selection even bears on the presence or absence of contrast between two vowels/vowel contexts. Thus, dynamic effects deserve to be considered both in articulatory methodology and in vowel description. This is further reinforced by recent findings that dynamic characteristics of the acoustic signal are predictive of sociophonetic factors, such as speaker age and social class (Haddican et al., 2013; Hughes & Foulkes, 2015), as well as dialect (Williams & Escudero, 2014).
Dynamic analysis of articulatory variables extracted from ultrasound is a novel approach, and hence it is more challenging than dynamic analysis of vowel formants. However, further methodological work in this domain is worthwhile, since high-speed ultrasound provides a relatively cheap and accessible tool for the study of phonetics and phonology. Recently, there have been major improvements in the efficiency of using this method, mainly thanks to the development of reliable automated tracking. Particular advantages of ultrasound for sociolinguistic research are for the study of the tongue root and for multiple-speaker research. We therefore look forward to the increased uptake, development and improvement of analytic tools by the phonetics and variationist research communities, in the laboratory and in fieldwork, where these unique characteristics can be fully exploited.

Notes

1Kleber et al. (2011) as well as Harrington et al. (2008) mainly talk about the preceding consonant as conditioning fronting/retraction, but they acknowledge that a following /l/ may block fronting independently. In this study, we focus on the effect of the following consonant exclusively.

2Sonderegger & Yu (2010) model this phenomenon not as perceptual error, but rather as part of continuum, where degree of perceptual compensation for coarticulation is shaped by
listener’s experience and properties of the ambient language.

REFERENCES


Harrington, J., F. Kleber & U. Reubold (2011). The contributions of the lips and the


co-ordinate system for mid-sagittal articulatory measurement, URL


Table 1: Summary of F2 frequencies for /uː/, /ʊ/ and /iː/ vowels reported in selected acoustic studies on Southern British English.

<table>
<thead>
<tr>
<th>Source</th>
<th>Speakers</th>
<th>Speaker age (at time of recording)</th>
<th>Average F2 (at time of recording)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wells (1962)</td>
<td>25 males</td>
<td>18+</td>
<td>939 950 2373</td>
</tr>
<tr>
<td>Deterding (1997)</td>
<td>5 females</td>
<td>Unspecified</td>
<td>1437 1340 2654</td>
</tr>
<tr>
<td></td>
<td>5 males</td>
<td>Unspecified</td>
<td>1191 1550 2249</td>
</tr>
<tr>
<td>Ferragne &amp; Pellegrino (2010)</td>
<td>6 males</td>
<td>Unspecified</td>
<td>1672 1550 2289</td>
</tr>
<tr>
<td>Williams &amp; Escudero (2014)</td>
<td>10 females</td>
<td>18–30</td>
<td>2202 1705 2760</td>
</tr>
<tr>
<td></td>
<td>7 males</td>
<td></td>
<td>1683 1320 2289</td>
</tr>
</tbody>
</table>
Table 2: Summary of F2 frequencies for /u:/, /u/ and /i:/ across four different age groups recorded in 2001. Source: Hawkins & Midgley (2005)

<table>
<thead>
<tr>
<th>Speakers</th>
<th>Speaker age (in 2001)</th>
<th>Average F2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/u:/</td>
<td>/u/</td>
</tr>
<tr>
<td>5 males</td>
<td>65–73</td>
<td>994</td>
</tr>
<tr>
<td>5 males</td>
<td>50–55</td>
<td>1112</td>
</tr>
<tr>
<td>5 males</td>
<td>35–40</td>
<td>1336</td>
</tr>
<tr>
<td>5 males</td>
<td>20–25</td>
<td>1626</td>
</tr>
</tbody>
</table>

Table 3: Experimental stimuli. The words in bold are treated as keywords throughout the paper, representing the category.

<table>
<thead>
<tr>
<th>Fronting context</th>
<th>Blocking context</th>
</tr>
</thead>
<tbody>
<tr>
<td>/u:/</td>
<td>/u:/</td>
</tr>
<tr>
<td>food</td>
<td>fool</td>
</tr>
<tr>
<td>boot</td>
<td>pool</td>
</tr>
<tr>
<td>/u/</td>
<td>/u/</td>
</tr>
<tr>
<td>foot</td>
<td>full</td>
</tr>
<tr>
<td>put</td>
<td>pull</td>
</tr>
</tbody>
</table>
Figure 1: (Colour online) Tongue position measurements relative to the /i/-/o/ plane by Scobbie et al. (2012). A common tangent linking the mean high front and high back vowels (line FB) defines the horizontal top limit of the vowel space. The frontness of a target vowel token is measured in absolute or relative terms via a perpendicular line starting at $\alpha$ dropped to the vowel’s nearest approximation to line FB, using the length of F-$\alpha$ on FB as the absolute difference in frontness. The absolute difference in height is $\delta$. The X- and Y- axes align to the unrotated edges of the rectangular image generated by the ultrasound scanner. Two other potential references, the hard palate and occlusal plane, are shown for comparison.
Figure 2: Example of unrotated (left), and rotated (right) tongue tracing. Tongue tip is on the right.
Figure 3: (Colour online) Averaged tongue contours (loess smoothed) for Speaker YF1 at the acoustic onset of the vowel. Tongue tip is on the right.
Figure 4: (Colour online) Measurements of /i:/ and /u:/ position along the occlusal plane. Tongue tip is on the right.
Figure 5: (Colour online) Dynamic effects of vowel and consonant context on the degree of fronting in the articulatory (left) and acoustic (right) domain.
Figure 6: (Colour online) The interactive effect of normalised F2 and context on articulatory fronting