



# Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English

Hannah King, Emmanuel Ferragne

## ► To cite this version:

Hannah King, Emmanuel Ferragne. Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English. *Journal of Phonetics*, 2020, 80, pp.100978. 10.1016/j.wocn.2020.100978 . hal-02549386

**HAL Id: hal-02549386**

**<https://hal.science/hal-02549386>**

Submitted on 21 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English

Hannah King<sup>a,\*</sup>, Emmanuel Ferragne<sup>b</sup>

<sup>a</sup>*CLILLAC-ARP, EA 3967, Université de Paris - Paris Diderot, 8 Place Paul Ricoeur, 75013 Paris, France*

<sup>b</sup>*Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Université Sorbonne Nouvelle - Paris 3, 19 rue des Bernardins, 75005 Paris, France*

---

## Abstract

This paper presents acoustic and articulatory data from prevocalic /r/ in the non-rhotic variety of English spoken in England, Anglo-English. Although traditional descriptions suggest that Anglo-English /r/ is produced using a tip-up tongue configuration, ultrasound data from 24 speakers show similar patterns of lingual variation to those reported in rhotic varieties, with a continuum of possible tongue shapes from bunched to retroflex. However, the number of Anglo-English speakers using exclusively tip-up variants is higher than that reported in American English across all phonetic contexts. It is generally agreed that English /r/ may be labialised, but the exact contribution of the lips has yet to be explored. Lip camera data reveal significantly more lip protrusion in bunched tongue configurations than retroflex ones. These results indicate that the differing degrees of lip protrusion may contribute to maintaining a stable acoustic output across the different tongue shapes. An articulatory-acoustic trading relation between the sublingual space and the degree of lip protrusion is proposed. Finally, we suggest that Anglo-English /r/ has a specific lip posture which differs from that of /w/. We relate the development of such a posture to Anglo-English speakers' exposure to labiodental variants and to the pressure to maintain a perceptual contrast between /r/ and /w/.

*Keywords:* Anglo-English, Rhotics, Articulation, Ultrasound, Lips, Labiodentalisation

---

## 1. Introduction

### 1.1. Tongue shape diversity

It is well-documented in rhotic varieties of English that approximant realisations of the phoneme /r/ (i.e. [ɹ]) may be produced with a number of different tongue shapes, which are categorised on a continuum between two extreme configurations: retroflex, with a raised and curled-up tongue tip and a lowered tongue body; and bunched, with a lowered tongue tip and a raised tongue body (Delattre & Freeman, 1968; Tiede et al., 2004; Zawadzki & Kuehn, 1980). Some speakers produce one configuration exclusively, while others present individual-level or contextually-conditioned variation (see Mielke et al., 2016, and references therein). For example, it has been observed in American English that prevocalic /r/ is produced with higher

---

\*Corresponding author

Email address: [hannah.king@univ-paris-diderot.fr](mailto:hannah.king@univ-paris-diderot.fr) (H. King)

degrees of retroflexion than postvocalic /r/ (Delattre & Freeman, 1968; Hagiwara, 1995; Mielke et al., 2016) and that retroflexion is favoured by back and perhaps also by open vowels (Mielke et al., 2016; Ong & Stone, 1998; Tiede et al., 2010). Furthermore, Westbury et al. (1998) observed that speakers with extreme bunched tongue shapes in the word *row* show less extreme bunching in the word *street*, which suggests that neighbouring vowels have a co-articulatory influence on bunched realisations too.

Despite the extensive literature on lingual variation in rhotic English varieties (e.g., North America: Dediu & Moisik, 2019; Magloughlin, 2016; Mielke et al., 2016 and Scotland: Lawson et al., 2011, 2018; Scobbie et al., 2015), the articulation of /r/ in non-rhotic varieties, particularly in the English spoken in England, henceforth Anglo-English, remains largely unexplored. It is not yet known to what extent prevocalic /r/ in non-rhotic Anglo-English differs from rhotic variants, although tip-up /r/ is generally more associated with Anglo-English than rhotic Englishes. Descriptions as early as Sweet (1877) refer to tip-up articulations as opposed to tip down ones. Jones (1972) describes the sound of /r/ as ‘the equivalent to a weakly pronounced retroflexed ə’ (p. 206). The three Anglo-English speakers presented in Delattre & Freeman (1968) used an ‘extreme’ tip-up shape prevocally, which differed from American English shapes. Similarly, Ladefoged & Disner (2012) explain that many ‘BBC English speakers’ use tongue tip raising towards the alveolar ridge, while many American English speakers bunch the body of the tongue up (p. 121). Interestingly, bunched /r/ is rarely, if ever, mentioned as an alternative strategy in pronunciation manuals for second language learners of English, particularly in those based on Standard Southern British English<sup>1</sup>. These manuals strongly focus on retroflexion, encouraging learners to curl the tongue tip back and often provide stylised midsagittal drawings indicating retroflexion (e.g., Ashton & Shepherd, 2012; Hancock, 2003; Marks, 2007; Roach, 1983; Underhill, 1994). Drawing on their experiences as voice and dialect coaches of British English, Ashton & Shepherd (2012) go as far as to suggest that the ‘correct position’ to produce the /r/ sound in English is with the tongue tip curled back and upwards towards the roof of the mouth (p. 48). Despite the abundance of tip-up and retroflex descriptions in the literature on Standard Southern British English, similar articulatory patterns to those found in rhotic English /r/ have recently been observed in New Zealand English (Heyne et al., 2018) and in a small-scale study of Anglo-English (Lindley & Lawson, 2016).

### 1.2. *Accompanying labial gesture*

Although the vast majority of articulatory work on /r/ focuses on its lingual gesture (Docherty & Foulkes, 2001), it is generally agreed that /r/ may be labialised but the exact phonetic implementation of labialisation is unknown. It has been observed that lip rounding is likely to occur in prevocalic and pre-stress syllable positions in both American English (Delattre & Freeman, 1968; Mielke et al., 2016; Proctor et al., 2019; Uldall, 1958; Zawadzki & Kuehn, 1980) and Anglo-English (Abercrombie, 1967; Jones, 1972; Scobbie, 2006), regardless of the shape of the tongue. On the other hand, Gimson (1980) suggests that lip rounding in Anglo-English /r/ is largely conditioned by the quality of the following vowel, with /r/ preceding rounded vowels exhibiting more rounding than /r/ preceding non-rounded vowels. However, it has been observed that English speakers do not always round their lips for so-called rounded vowels (Brown, 1981), and that they

---

<sup>1</sup>We found one mention of bunching in a teachers’ manual on American English pronunciation (Ehrlich & Avery, 2013). The authors indicate that although there is a ‘disagreement’ regarding the characterisation of /r/ as either retroflex or bunched which may be due to ‘dialectal differences’, they stress that retroflexion is the most useful characterisation for pedagogical purposes.

use less rounding than speakers of other languages with phonologically equivalent rounded vowels, such as French (Badin et al., 2014; Wilson, 2006). Ladefoged & Disner (2012) note that modern productions of the vowel /u:/ have relatively spread lips in comparison to productions of the recent past, although articulatory studies have indicated that while /u:/ remains rounded, it is no longer a back vowel (e.g., Harrington et al., 2011; King & Ferragne, 2018; Lawson et al., 2015). Brown (1981) even goes as far as to suggest that the main origin of lip rounding in English derives not from rounded vowels, but rounded consonants, and that the most marked lip movement can be found in the consonants /ʃ, tʃ, ʒ, dʒ/ and /r/, although this idea does not seem to have been developed further. English pronunciation manuals vary with their treatment of the labial gesture. O'Connor (1967) recommends learners approach [ɹ] from [w], and then curl the tip of the tongue back until it is pointing at the hard palate, which presumably supposes that the lip postures for [ɹ] from [w] are identical. Others warn learners not to exaggerate rounding for /r/ because it would have the effect of producing the percept of a [w] (e.g., Lilly & Viel, 1977; Roach, 1983). While Ehrlich & Avery (2013) indicate that lip rounding is a possibility, Ashton & Shepherd (2012) inform learners that using their lips to help them form the /r/ sound is 'wrong' and recommend learners use their fingers to hold their lips still in order to practise using just their tongue (p. 49).

The phonetic implementation of labialisation in consonants is surprisingly rarely addressed. Indeed, Laver (1980) explains that the label 'labialisation' has been used so extensively that the only appropriate articulatory action to which the various usages refer is likely a horizontal constriction of the interlabial space. Horizontal constriction occurs when the lip corners are compressed making the space between the lips smaller, using the orbicularis muscle (Laver, 1980). The opposite lip posture, horizontal expansion, results in an articulation resembling a 'fixed, slight grin' (Laver, 1980, p.36), i.e., lip spreading. According to Laver (1980), horizontal constriction is the articulatory property that all rounded vowels and consonants have in common, implying that any labial configuration without horizontal constriction would not be considered rounded or labialised. He remarks that labial protrusion is almost always accompanied by a certain degree of horizontal constriction of the space between the lips, although substantial lip protrusion without horizontal constriction is physiologically possible (Laver, 1980). Horizontal constriction of the lip corners towards the centre has been described as 'pouting' by Catford (Catford, 1977, 1988). Although he bases his observations predominantly on the articulation of vowels, Catford distinguishes two types of rounding: endolabial and exolabial, which parallel Sweet (1877)'s classification of inner and outer rounding in vowels. In endolabial or inner rounding, which is typical of back vowels, the lip corners are brought in towards the centre (i.e., 'pouted'), pushing the lips forwards to form a channel between the inner surfaces of the lips. In exolabial or outer rounding, which is typical of front vowels, the corners of the mouth are vertically compressed without 'pouting', leaving a slit-like elliptical shape between the lips, rather than actually round.

With regards to English /r/, the terms lip protrusion and lip rounding seem to be used interchangeably, perhaps because, as Laver (1980) indicates, protrusion without lip rounding is rare in the world's languages. However, also inspired by Sweet (1877)'s articulatory account of rounding in vowels, Brown (1981) explicitly differentiates the two: rounding restricts lip aperture by compressing the lip corners, but does not necessarily push the lips forward, as is the case for English /w/; while protrusion pushes the lips forward, opening and everting them to show the soft inner surfaces, as in English /ʃ, tʃ, ʒ, dʒ/ and /r/. Again like Laver (1980), Brown (1981) essentially uses horizontal compression to define lip rounding, which is notably absent from her description of the 'protruded' consonants /ʃ, tʃ, ʒ, dʒ/ and importantly for the present study, /r/. However,

in a very recent articulatory study on sound change triggered by American English /r/, Smith et al. (2019) observed that /ʃ/ lip rounding is different from /r/ lip rounding. Speakers produced /ʃ/ with open protruded (‘outrounded’) lips, while /r/ involved vertical movement by the upper and/or lower lip, sometimes with a narrow lip aperture (‘inrounded’). However, both /ʃ/ and /r/ exhibited inter-speaker variability in the shape and area of the labial constriction.

The lips are of particular interest in Anglo-English as labiodental variants (e.g., [v]) are becoming increasingly common (Docherty & Foulkes, 2001; Marsden, 2006). It is generally implied that labiodental variants have emerged by speakers retaining the labial component of /r/ at the expense of the lingual one (Docherty & Foulkes, 2001; Foulkes & Docherty, 2000; Jones, 1972), although there is a lack of articulatory data. Docherty & Foulkes (2001) hypothesise that this change in progress may be the result of the heavy visual prominence of the labial gesture for /r/, which may have led to the labial taking precedence over the lingual articulation. Lindley & Lawson (2016) observed one English participant who produced labiodental /r/ with no observable tongue body gesture. However, another English participant presented labiodentalisation accompanied by a tip-up tongue configuration, leading them to suspect that the change in progress from [ɹ] to [v] may be phonetically gradient, in line with Docherty & Foulkes (2001)’s hypothesis. However, as far as we are aware, no articulatory study has yet accounted for the exact contribution of the lips to the production of Anglo-English /r/, and as Docherty & Foulkes (2001) note, this may result in a ‘skewed view of the physical basis of this variant’ (p. 183).

### 1.3. *Acoustic properties of /r/*

Despite the diversity of possible tongue shapes observed for [ɹ], the acoustic profile of these different tongue configurations is remarkably indistinguishable, at least with regards to the first three formants (Espy-Wilson et al., 2000). /r/ is characterised by a low F1, a low F2, and an extremely low F3 (Mielke et al., 2016). Formant values from American English /r/ reported in the literature across tongue shapes, phonetic contexts and sexes range from 300-500 Hz for F1, 900-1300 Hz for F2, and 1300-2000 Hz for F3 (Delattre & Freeman, 1968; Espy-Wilson, 1992; Espy-Wilson & Boyce, 1999; Westbury et al., 1998; Uldall, 1958; Zhou et al., 2008). In rhotic Englishes, prevocalic /r/ presents lower formant values than postvocalic /r/, which is generally assumed to be the result of the presence of lip rounding in prevocalic /r/ (Delattre & Freeman, 1968; Lehiste, 1962; Zawadzki & Kuehn, 1980). Beyond F3, consistent acoustic differences have been found in higher formants in American and Scottish English; notably the difference between F4 and F5 has been found to be larger in retroflex than in bunched /r/ (Lennon et al., 2015; Zhou et al., 2008). This difference does not appear to be perceptibly salient as it has been shown that American English listeners are unable to distinguish between bunched and retroflex /r/ (Twist et al., 2007). On the other hand, there is evidence to suggest that lingual variation is socially distributed in postvocalic /r/ in certain varieties of Scottish English. Lawson et al. (2011) observed that middle-class speakers used bunched articulations, while working-class speakers used more retroflex ones in the eastern Central Belt of Scotland, and as a result, they argue that this articulatory variation must be in some way perceptible and exploited by listeners to index socio-economic class. In another small-scale study, Lawson et al. (2014) asked listeners to mimic speakers from audio recordings of middle- and working-class speakers and in some cases, mimicry participants adapted their tongue shape for /r/. This social distribution of tongue shape variants may have motivated sound changes, including the merger to schwa in bunchers (Lawson et al., 2013) and derhoticisation in retroflexers (Lawson et al., 2011). However, it is still unclear what made the bunched and

retroflex configurations perceptibly distinct enough to result in socially distributed variation in the English of the eastern Central Belt of Scotland.

The relationship between acoustics and the articulation of American English /r/ has received a lot of attention. It is generally agreed that the most salient acoustic feature of /r/ is its extremely low F3. Theoretical models have associated this low F3 with a large front cavity volume, i.e., between the palatal constriction and the lips (Alwan et al., 1997; Fant, 1960; Stevens, 1998), although for Stevens (1998), the various tongue configurations used for /r/ do not lower F3 but introduce an extra resonance,  $F_R$ , in the frequency range normally occupied by F2 along with a drop in amplitude of F3 proper. Espy-Wilson et al. (2000) used MRI-derived vocal tract dimensions in American English /r/ and found that the front cavity is indeed large enough to lower F3. Their tube models indicate that the front cavity includes a lip constriction formed by the tapering gradient of the teeth and lips - with or without rounding - and a large volume cavity behind it that includes sublingual space, which acts to increase the volume of the cavity. The sublingual space is the space between the tongue tip and the lower teeth that is introduced when the tongue tip or blade is raised towards the post-alveolar region (Hamann, 2003). Unlike tip-up /r/, the tongue tip is down in bunched /r/ and therefore has negligible sublingual space (Zhang et al., 2003). Espy-Wilson et al. (2000) found that the addition of a sublingual space lowers F3 by approximately 200 Hz. Differences have been observed between the size of the front cavity in different lingual configurations. Alwan et al. (1997) used MRI- and EPG-derived vocal tract dimensions, and in one American English speaker, the anterior cavity was larger for retroflex /r/ than bunched (6.1 cm<sup>3</sup> and 4.5 cm<sup>3</sup>, respectively). This difference may be due to the smaller sublingual space in bunched /r/, although Alwan et al. (1997) do not explicitly make this suggestion.

The consistency in formant values observed for /r/ has given rise to the suggestion that trading relations may exist between the different articulatory manoeuvres which reciprocally contribute to the lowering of F3. Dependence on one of these articulatory manoeuvres would be accompanied by less of another, and vice versa (Tiede et al., 2010). Guenther et al. (1999) found trading relations between palatal constriction location, constriction degree, and constriction size for American English /r/. Alwan et al. (1997) posit a trading relation between sublingual space for tip-up /r/ and a more posterior palatal constriction for tip-down /r/ (as discussed in Espy-Wilson et al., 2000). Extending the front cavity - and thus increasing its volume - could also be achieved through the addition of lip protrusion. Yet, to the best of our knowledge, trading relations involving lip protrusion have yet to be investigated. Given the trading relations already observed for /r/, it does not seem unlikely then that different degrees of lip protrusion may accompany different tongue configurations. Indeed, trading relations have been observed between the lips and tongue in other speech sounds, such as in the vowel /u:/ (Perkell et al., 1993). As a result, in line with previous work on trading relations in /r/, we predict that bunched configurations will be accompanied by more lip protrusion than retroflex ones, due to the smaller sublingual space in bunched /r/. To our knowledge, two existing studies have indeed observed a positive correlation between lip protrusion and bunching in both Anglo-English (Lindley & Lawson, 2016) and American English (Tiede et al., 2010), although both studies were small-scale, and explanations as to why have yet to be given.

In this paper, we give a detailed articulatory account of Anglo-English prevocalic /r/ on a relatively large scale (24 speakers). We aim to determine first and foremost whether Anglo-English /r/ can be produced using multiple tongue shapes, as has been found in other varieties. If this is indeed the case, we will assess

whether different tongue shapes are accompanied by different degrees of labiality. Finally, we will compare the labial configurations for /r/ and /w/ and will attempt to relate articulation to the change in progress towards labiodentalisation currently underway in Anglo-English /r/.

## 2. Materials and methods

### 2.1. Materials

The data we present here come from a study comparing hyperarticulated and non-hyperarticulated productions of /r/ (King & Ferragne, 2019). In this paper, we present data only from non-hyperarticulated tokens. Stimuli were made up of 16 minimal pairs contrasting /r/ and /w/ word initially. /r/ and /w/ were followed by the following lexical set vowels: FLEECE, GOOSE, KIT, DRESS, TRAP, STRUT, THOUGHT, LOT. 14/16 stimuli had a coda consonant and all words were produced in isolation. As the experimental paradigm limited the number of repetitions in the non-hyperarticulated context, only one repetition of each stimulus was recorded per speaker. We therefore present data from 384 tokens. A complete list of stimuli is presented in Appendix A.

Simultaneous articulatory and acoustic data were obtained using Articulate Assistant Advanced (AAA) software (Articulate Instruments Ltd., 2014). Tongue images were recorded at a rate of circa 121 frames per second (fps) using a high-speed SonixRP ultrasound system. Participants wore a headset to ensure the ultrasound probe remained in a stable position relative to the head (Articulate Instruments Ltd., 2008). Two NTSC micro-cameras were attached to the headset, capturing front and profile lip videos at a rate of circa 60 fps. An Audio-Technica AT803 microphone was also attached to the headset. Audio files were digitised as LPCM mono files with a 22050 Hz sampling rate and 16-bit quantization. Technical details concerning this particular ultrasound system and associated video and audio synchronisation are described in Wrench & Scobbie (2016). We recorded each speaker swallowing water in order to obtain an outline of the palate (Epstein & Stone, 2005). Speakers were also recorded biting on a plastic bite plate, which was used to image each speaker’s occlusal plane (Scobbie et al., 2011; Lawson et al., 2019). The palate and occlusal plane were subsequently traced in AAA.

### 2.2. Participants

29 native speakers of Anglo-English were recorded at Queen Margaret University, Edinburgh. Speakers were recruited through advertising on the university Research Recruitment Digest communications service. Participants self-identified as speaking with an English accent and the first author, who is a native Anglo-English speaker, made sure that this was indeed the case by speaking to the participants prior to recording them. Before participating, speakers signed an informed consent form and completed a background questionnaire. Ethical approval had previously been obtained from Queen Margaret University Research Ethics Committee. Data collection sessions lasted no more than 30 minutes for which participants were financially compensated. Some speakers’ data were excluded due to ultrasound data visualisation issues (n=4) and one English-Punjabi bilingual was excluded because Punjabi also has retroflex consonants in its inventory. We present data from the remaining 24 speakers (22 F, 2 M) aged between 18 and 55 ( $M=30.08 \pm 11.26$ ) who come from all over England (south west: n=1; south east: n=6; midlands: n=3; north west: n=7, north east: n=7). 19 speakers had lived in Scotland for at least one year. The inclusion of the word *war* in the stimuli allowed us to classify the participants as rhotic and non-rhotic. All speakers were non-rhotic apart

from the one speaker from the south west of England, where rhotic accents do indeed occur (Wells, 1982), although they are reportedly becoming less rhotic (Trudgill, 1999). Incidentally, this subject is one of the oldest speakers in the dataset (54 years old).

### 2.3. Acoustic analysis

The acoustic data were exported as wav files from AAA and analysed in Praat (Boersma & Weenink, 2019). Determining the point at which to segment /r/ from the following vowel is challenging. Although Lawson et al. (2010) suggest that for postvocalic /r/, the most reliable means to determine the dividing line between the two is by considering amplitude changes, in our prevocalic /r/ data, we observed large amounts of amplitudinal variation both within and across speakers. We were therefore unable to find a sufficient technique that could be applied to all speakers. As a result, /r/ and the following vowel were manually annotated as a whole. Praat's Burg algorithm was used to obtain formant values. Formant parameters were manually adjusted in order to reach an optimal match between formant estimation and the underlying spectrogram. The first three formants (F1-F3) were extracted at the point of minimal F3 for /r/ (as in Guenther et al., 1999). Unfortunately, formants higher than F3 were too weak to be accurately tracked. F1-F3 were also extracted at the midpoint of a steady state of the following vowel, avoiding obvious transitions to and from the surrounding consonants.

### 2.4. Articulatory analysis

#### 2.4.1. Ultrasound tongue imaging

One ultrasound frame was selected per recording depicting the maximal constriction of the anterior lingual gesture for /r/ prior to any obvious movement into the following vowel. This was achieved by holistically examining the raw ultrasound images one by one in a sequence. For each selected image, the tongue contour was traced semi-automatically in AAA and manually corrected when necessary. The resulting contours were rotated to each speaker's individual occlusal plane, which aided tongue shape classification, specifically with regards to the position of the tongue tip. Figure 1 depicts our rotation technique: all contours are rotated so that the occlusal plane (blue line) tracing is horizontal.

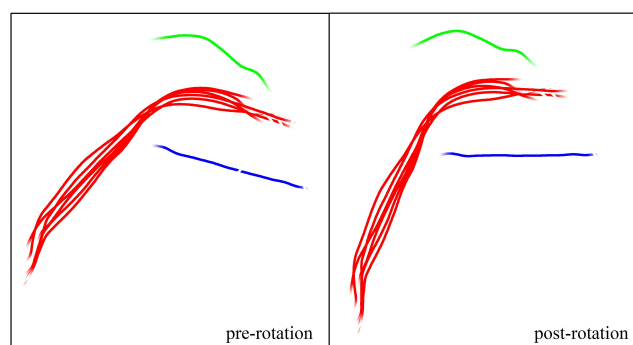


Figure 1: Example of rotation to the occlusal plane. The tongue tip is on the right. The hard palate is traced in the top curve. All contours are rotated so that the occlusal plane (bottom line) is horizontal.



Both the raw ultrasound frames and the rotated tongue contours were used to classify tongue configurations for /r/ on a continuum largely inspired by the one presented in Lawson et al. (2013) for Scottish English, which depicts four distinct shapes: Mid Bunched, Front Bunched, Front Up and Tip Up (pp. 199–200). Our classification differs in that it includes a fifth configuration: an ‘extreme’ retroflex involving curling up of the tongue tip, which has previously been associated with Anglo-English (as discussed in 1.1). The classification originally proposed by Lawson et al. (2013) grouped the curled-up and the non-curled-up tip-up /r/ together. Ultrasound images give some indication of the curling up of the tongue tip, which is described below. However, we do not know to what extent the identification of these articulations is constrained by speaker anatomy. In some cases, it is possible that the jaw shadow obscures the tongue tip, which would make visualising ‘real’ retroflexion challenging. It is therefore possible that the number of curled-up articulations is underestimated in our analysis<sup>2</sup>. The articulations of each configuration in our classification are described below, and Figure 2 presents raw ultrasound images of typical examples of each configuration from our dataset.

- Mid Bunched (MB): the middle of the tongue is raised towards the hard palate, while the front, blade and tip are low.
- Front Bunched (FB): the front of the tongue has a distinctly bunched configuration which results in a dip in the tongue’s surface behind the bunched section. The tip and blade remain lower than the rest of the tongue front.
- Front Up (FU): the front, blade and tip are raised and the tongue surface forms a smooth convex curve.
- Tip Up (TU): the tongue tip is pointing up resulting in a straight and steep tongue surface.
- Curled Up (CU): the overall tongue shape is concave and the tip is curled up. Curling up of the tongue tip results in a near-parallel orientation of the tongue surface to the ultrasound scanlines, producing artefacts in the ultrasound image (Scobbie et al., 2013). We tend to observe a bright white region above where the tongue tip is expected (Mielke et al., 2016) and a discontinuity in the tongue contour where the tongue tip is curled up (Bakst, 2016).

In order to facilitate the task of classifying tongue configurations, the decision tree presented in Figure 3 was produced and used throughout the classification process. The first author classified tongue shapes three times throughout the course of one year to ensure accuracy. Although discrepancies in the three classifications were rare, such cases were reexamined and the most common configuration of the three was selected<sup>3</sup>.

If we employ the traditional retroflex-bunched classification, the Mid and Front Bunched configurations have a low tongue tip and the primary constriction is located between the front to mid tongue body (Lawson et al., 2011), so we can consider them to be bunched. Although retroflexion has traditionally been described as an articulation involving the curling up of the tongue tip (e.g., Catford, 1977), Hamann (2003) notes that

---

<sup>2</sup>We would like to thank an anonymous reviewer for highlighting this point.

<sup>3</sup>An anonymous reviewer suggested having several researchers perform the classification procedure and to calculate a measure of inter-rater agreement. While this technique has not been implemented here, we would also recommend future studies to use such a technique, along with the use of a decision tree, such as the one presented in Figure 3.

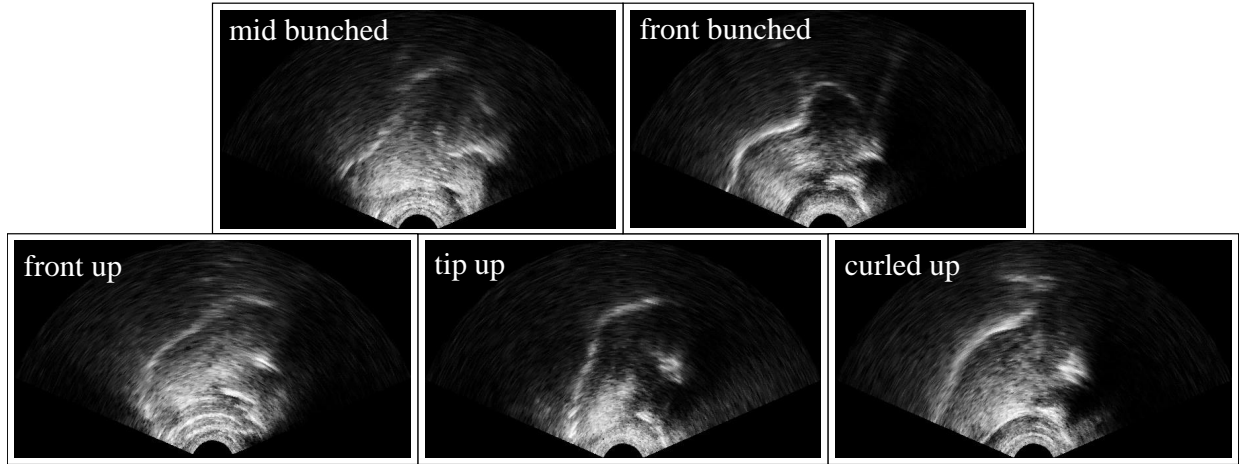


Figure 2: Raw ultrasound frames showing typical examples of each of the five /r/ configurations. The tongue tip is on the right side of the image. The top two images are bunched, while the bottom three are retroflex. The final retroflex configuration exhibits curling up of the tongue with a bright white line where the tongue tip is expected.

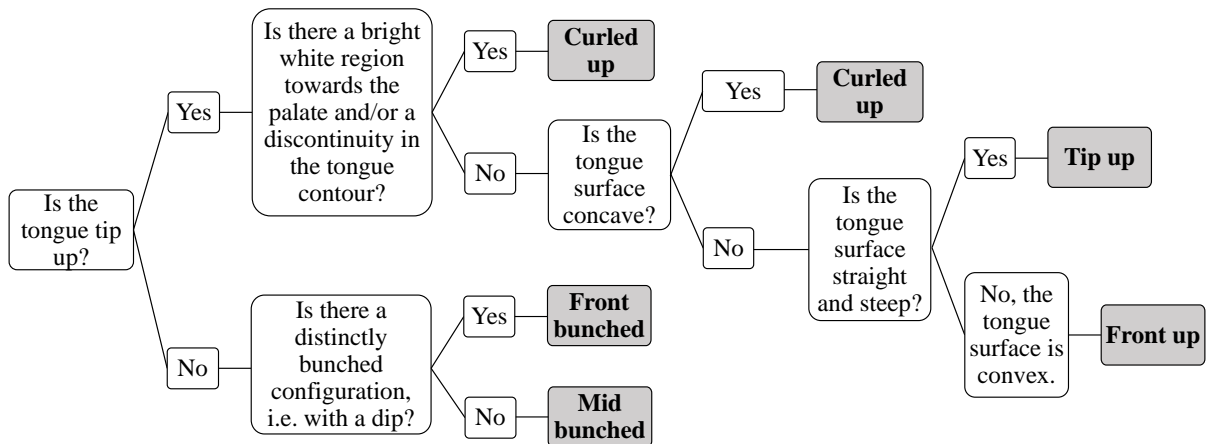


Figure 3: Decision tree used to classify tongue shapes into five distinct categories for /r/ from ultrasound data.

this property is violated by a large number of segments traditionally considered retroflex in many languages because the tongue tip often fails to curl up. Instead, she proposes the combination of four articulatory characteristics to define retroflex segments, namely apicality, posteriority, sublingual cavity, and retraction. As such, any sound articulated with the tongue tip behind the alveolar region and involving a displacement of the tongue back towards the pharynx or velum would be considered retroflex by her definition. As bunched /r/ has also been shown to include tongue root retraction (Delattre & Freeman, 1968; Proctor et al., 2019) and the drawing inwards of the tongue body away from the lips (Alwan et al., 1997), the main criterion we considered to define retroflexion for /r/ is the raising of the tongue tip, which results in the addition of a sublingual space. The tongue tip and/or tongue front are raised towards the post-alveolar region in the last three configurations of our classification (FU, TU, CU), and so, we therefore consider them to be retroflex. Although in some raw ultrasound images, the primary constriction (i.e., the highest point of the tongue) in some Front Up configurations may appear to be the tongue dorsum (as in the Front Up image presented in Figure 2), when the corresponding tongue contour is rotated to the occlusal plane, the tongue tip does generally appear to be the primary constriction, or at least pointing up, an example of which can be observed in Figure 1. Interestingly, the Front Up tongue shape has been considered to be bunched and not retroflex in other classifications. For example, the equivalent ‘blade raised’ configuration described in Mielke et al. (2016) is classified as bunched. However, the authors observe that these tokens are often ambiguous with respect to tongue tip or tongue blade angle and they do consider classifying them as retroflex. It appears then that the Front Up configuration lies somewhere in the middle of the bunched-retroflex continuum.

In the present study, our classification would place the variant with the highest, most curled-up tongue tip, the Curled Up configuration, at one end of the continuum. Curled Up is followed by the Tip Up and Front Up variants respectively. Deciphering which tongue shape is the most bunched category between Mid Bunched and Front Bunched is less evident. Although by visualising the tongue contour tracings in speakers who present both configurations revealed that the tongue tip is generally lower in the Mid Bunched than the Front Bunched configuration, the Front Bunched category presents the most obvious bunching of the tongue i.e., with a dip in the tongue surface (as can be seen in Figure 2). Furthermore, the very tip of the tongue is not always visible from ultrasound images and so we err on the side of caution regarding the accuracy of tongue tip tracings. It is hoped that results from this study may provide further insights into which bunched configuration is the most extreme of the two.

#### 2.4.2. Lip protrusion

Lip protrusion was calculated from profile lip videos in AAA. One image corresponding to a neutral lip configuration (with the lips closed) prior to speech was visually selected per speaker. The image corresponding to maximum lip protrusion was visually identified for each production of /r/ and /w/ by holistically examining sequential video frames. Lip protrusion was measured by calculating the difference between maximum protrusion and the speaker’s neutral lip protrusion. To obtain quantitative data, a fiducial line (i.e., a fixed line used as a basis of reference and measure) was positioned to intersect the lip corner during each speaker’s neutral image. This fiducial had previously been scaled (in centimetres) to a physical ruler positioned along the mid-line of the stabilisation headset and ran parallel to the upper and lower edges of the video pane. Each speaker was assigned one lip corner fiducial which was used for all his/her protrusion measures. For the same neutral lip image, a line was positioned to touch the lower and upper lip edge, intersecting the neutral lip corner fiducial. Using AAA, we calculated the distance from the origin of the

fiducial to where the lip edge line crossed, yielding a value (in centimetres) for the neutral lip position. We employed the same technique to obtain values for the maximum protrusion distance for /r/ and /w/ using the previously selected maximum protrusion images. The neutral lip distance measurement was subtracted from the maximum protrusion distance for /r/ and for /w/ yielding final protrusion values, as depicted in Figure 4.

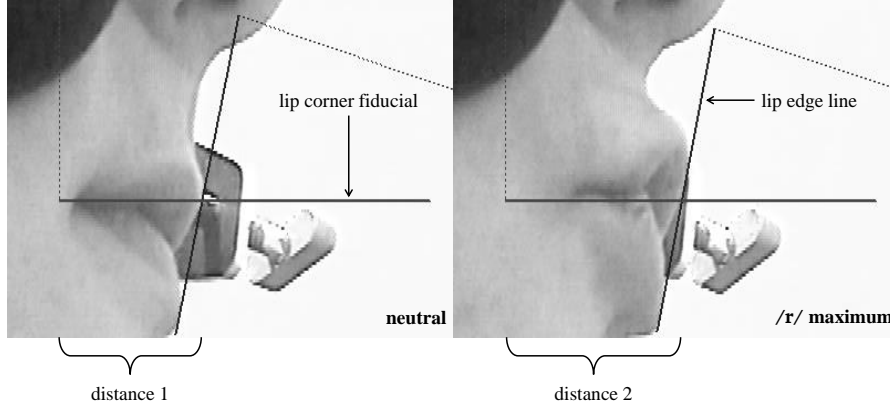


Figure 4: Lip protrusion measure. Distance 1 is subtracted from distance 2.

#### 2.4.3. Lip aperture and spreading

As the frontal and profile view lip cameras were synchronised, the corresponding frontal view images to the ones selected for the protrusion measure in the profile view (as presented in 2.4.2) were used to measure lip aperture and spreading. Aperture and spreading were measured during maximum protrusion for /r/ and /w/ and were compared to the values obtained during each speaker's neutral lip setting. Lip measurements were inspired by those presented in Garnier et al. (2012) and Mayr (2010), where spreading is measured at the lip corners, and aperture is measured from the middle of the top lip to the middle of the bottom lip. For lip spreading, a fiducial line was positioned to coincide with the quasi-horizontal line which is naturally formed between the top and bottom lip when the lips are closed in a neutral position. This horizontal fiducial ran parallel to the upper and lower edges of the video pane. A vertical line was then positioned at each lip corner intersecting the horizontal fiducial, as presented in Figure 5. Using AAA, we calculated the distance between the left and right lip corner along the horizontal lip fiducial in the neutral front image and in /r/ and /w/. To quantify lip aperture, another lip fiducial was positioned to vertically dissect the lips approximately at their mid-point at the philtrum dimple in their neutral setting. This vertical fiducial ran parallel to the left and right edges of the video pane. A horizontal line was positioned at the vermilion border of the outer edge of the top and bottom lip intersecting the vertical fiducial, as presented in Figure 5. Using AAA, we calculated the distance between the top and bottom lip along the vertical lip fiducial in the neutral front image and in /r/ and /w/. Each speaker was assigned one horizontal fiducial and one vertical fiducial which were used for all his/her lip aperture and spreading measures. Deviations from the neutral lip setting were measured by subtracting the measurements for the neutral lip image from the measurements for /r/ and /w/ (as presented in Figure 5). Although AAA produced values in centimetres, unlike for our lip protrusion measure presented in 2.4.2, no scaling device was used for the frontal lip view in our data. The

measurements are therefore not in world units. As a result, the values were transformed into the percentage of change relative to each speaker’s neutral lip setting dimensions.

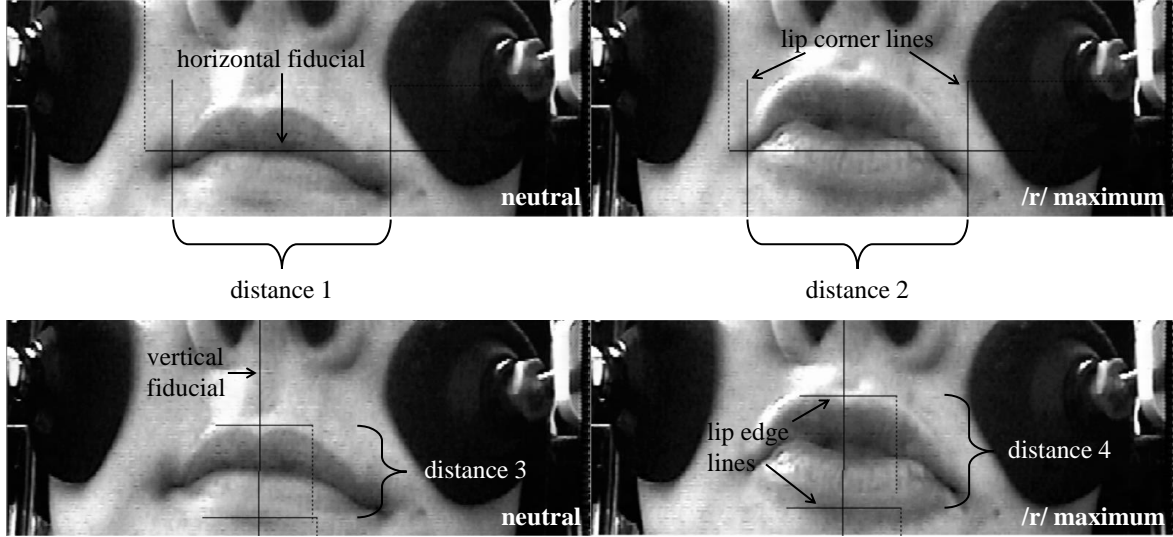


Figure 5: Frontal view lip measures. For lip spreading, distance 1 is subtracted from distance 2. For lip aperture, distance 3 is subtracted from distance 4.

#### 2.4.4. Statistical analysis

Statistical analysis was implemented in R (R Core Team, 2018) using the `lmer()` function of the *lme4* package (Bates et al., 2015) to perform a series of linear mixed-effects models. We tested the significance of main effects to model fit using likelihood ratio tests with the `mixed()` function in the *afex* package (Singmann et al., 2015). Model residuals were plotted to test for deviations from homoscedasticity or normality. The *lmerTest* library (Kuznetsova et al., 2017) was used to calculate indications of significance within the final models, which uses values derived from Satterthwaite (1946)’s approximations for the degrees of freedom. The resulting p-values are provided in the model summary tables. Plots of the predicted effects from final models were generated with the *sjPlot* package (Lüdtke, 2018).

### 3. Results

#### 3.1. Classification of tongue shapes

Visual classification of tongue configurations yielded the results presented in Table 1. Out of the 24 speakers, 7 produced only bunched /r/ configurations, 14 produced only retroflex, and 3 used both. Our data therefore contradict traditional descriptions of Anglo-English /r/ in that speakers do not only produce /r/ with a tip-up articulation. However, we observed double the number of speakers producing only retroflex /r/ compared to speakers producing only bunched.

Subject code	Age	Sex	/r/ coding	Shape
05	22	F	MB	bunched
08	26	F	MB	
17	27	F	MB	
10	44	M	FB MB	
03	22	F	FB	
11	29	F	FB	
22	23	F	FB	
29	18	F	MB FB FU TU CU	bunched & retroflex
14	23	F	MB FB CU	
18	23	F	FB FU CU	
02	22	F	FU	retroflex
23	33	F	TU FU	
16	25	F	TU FU	
13	54	F	TU	
12	20	F	CU FU	
15	25	F	CU TU FU	
19	28	F	CU TU FU	
27	37	F	CU TU FU	
28	29	F	CU TU FU	
07	22	F	CU TU	
09	21	F	CU TU	
21	41	F	CU TU	
25	55	F	CU TU	
04	53	M	CU	

Table 1: Observed tongue configurations in twenty-four subjects divided into three categories ordered from most bunched to most retroflex.

In order to discern any patterns regarding the geographical origin of speakers and their tongue configuration for /r/, the map presented in Figure 6 was produced. To make any real claims concerning the relationship between tongue shape and speaker origin, we would require more regionally-stratified data. However, from the present dataset, we note that two subjects (08 & 21) who come from the same town in the North West, Chester, use bunched and retroflex /r/ respectively. The only discernible pattern in our data concerns the subjects who use both retroflex and bunched /r/, as all three come from the South East, although other speakers from the same region were observed using either retroflex or bunched shapes. It is interesting to note that labiodental variants have been established as an accent feature of non-standard accents from the same region (Foulkes & Docherty, 2000). However, we stress that to make any claims regarding the relationship between tongue shape patterns and the development of labiodental variants in different regions would require more geographically balanced data.

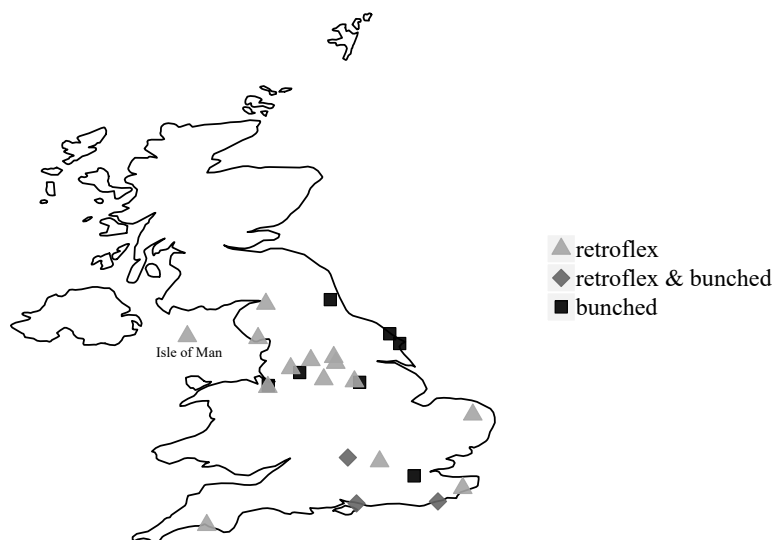


Figure 6: Map of speaker origin as a function of tongue configuration for /r/.

If we take a more detailed look at tongue configuration going beyond the simplistic retroflex-bunched distinction, based on our classification using five distinct shapes as presented in 2.4.1, we observe 9 out of the 24 subjects using one configuration exclusively, 6 of which are bunchers. In fact, all bunchers but one use one tongue configuration across all contexts. The remaining 15 speakers use multiple configurations. One buncher (speaker 10) uses the Front Bunched configuration in all vowel contexts except before the FLEECE vowel, where the Mid Bunched shape is used instead. Among the 17 retroflexers in the dataset, 13 of them use the extreme Curled Up configuration at least some of the time, which has previously been associated more with Anglo-English than American English. However, only one speaker (speaker 04) produces this extreme Curled Up variant exclusively, leading us to suspect that the following vowel may have a coarticulatory influence on retroflexion in most speakers, which has also been observed in American English (as discussed in 1.1).

In order to discern any patterns regarding tongue shape and the following vowel, we first need to establish what constitutes a close front and a open back vowel in Anglo-English. If we agree that F2 is an acoustic

correlate of tongue anteriority and F1 of tongue height, vowel plots should give us some indication of the relative frontness and openness of the vowels in the system. First and second formant values were extracted at the midpoint of a steady state of the vowel in /r/-initial words in Hertz. Formant values were scaled by means of Lobanov normalisation (Lobanov, 1971). Figure 7 shows ellipses to one standard deviation from the Lobanov normalised values. One striking observation is the frontness of the GOOSE vowel which is a known feature of UK accents, especially in Southern British English (e.g., Ferragne & Pellegrino, 2010; Harrington et al., 2011; Lawson et al., 2015). In terms of F2, GOOSE is by far the most variable of all the vowels in our dataset, with some tokens approaching the space occupied by FLEECE while others have an F2 closer to that of LOT. As previously discussed, articulatory studies have shown that the GOOSE vowel, while still rounded, can no longer be considered a back vowel in many varieties of English (e.g., Harrington et al., 2011; King & Ferragne, 2018; Lawson et al., 2015). Our formant data indicate that while some productions of the GOOSE vowel are fronted, others remain relatively back. This may be a result of having a large number of subjects from the North of England in our dataset (n=16) who have previously been shown to present less GOOSE-fronting than southerners (Ferragne & Pellegrino, 2010; Lawson et al., 2015). The STRUT vowel is also rather variable with some tokens having much higher F1 values than others, which presumably reflects dialectal differences concerning the FOOT-STRUT split. The backest vowel of the system is THOUGHT and the frontest is FLEECE. If retroflexion is favoured by back rather than front vowels, we would expect *raw* to exhibit more retroflexion than *reed*. However, if retroflexion favours open vowels over close vowels, we would expect /r/ preceding the TRAP vowel in *rack* to induce the most retroflexion, as it is the most open vowel in our dataset.

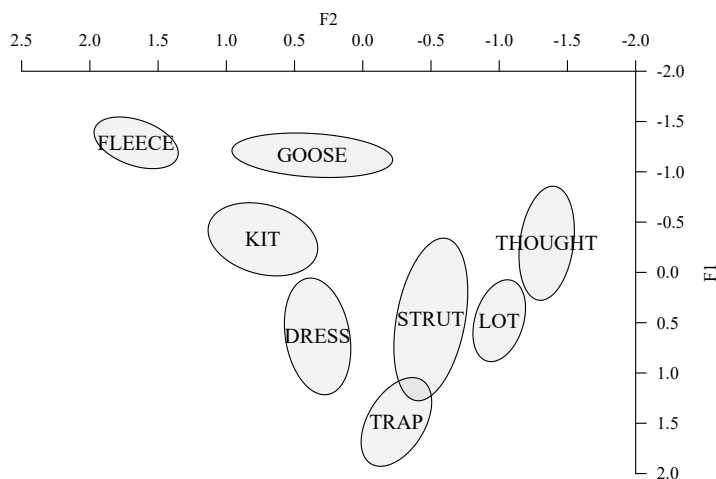


Figure 7: Lobanov-transformed vowel plot with one standard-deviation ellipses.

To examine to what extent the following vowel affects retroflexion, we considered the data from speakers who use at least one of the three retroflex configurations (n=17). Exclusively bunched /r/ users (n=7) were therefore excluded from this analysis. The proportion of each of the five /r/ configurations was plotted as a function of the following vowel in Figure 8. As predicted, the FLEECE vowel has the least retroflexion with less than 3% of the tokens presenting the extreme Curled Up variant. We observe that in the speakers who use



both retroflex and bunched variants, the bunched tokens are only used in /r/ followed by the frontest vowels of the system (FLEECE, GOOSE, KIT, DRESS). It may be that in these speakers, retroflexion is incompatible with front vowels and as a result, bunched configurations are used instead. The most retroflexion was observed preceding the LOT vowel with around 75% of tokens presenting the extreme Curled Up tongue configuration. Our data seem to be consistent with previous work on American English in that retroflexion is favoured by open back vowels. Although the THOUGHT vowel is the backest vowel of the system, LOT favours retroflexion more, perhaps because it is more open. However, TRAP is more open than STRUT but presents less retroflexion, perhaps because STRUT is generally further back. It seems then that both tongue position and height of the neighbouring vowel affect the tongue configuration used for /r/.

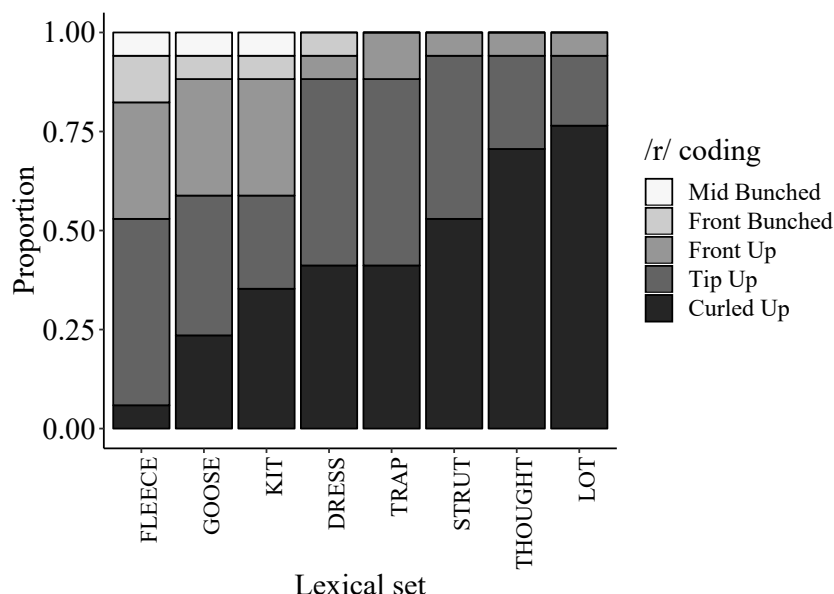


Figure 8: Proportion of /r/ tongue configurations as a function of the following vowel in retroflex users.

For visualisation purposes, Figure 9 presents tongue contour tracings for each speaker’s /r/ production at the point of maximal constriction preceding the FLEECE vowel (solid line) and the LOT vowel (dashed line) ordered from most bunched to most retroflex. Asterisks correspond to speakers who were coded as using more than one of the five tongue configurations. Even in speakers who are not considered to present multiple tongue shapes for /r/, we observe differences in tongue position between the two contours. The tongue is generally more anterior preceding FLEECE than it is preceding LOT, which is almost certainly a result of co-articulation. This observation may have an influence on the extent of accompanying lip protrusion. As we have already noted in 1.3, extending the front cavity results in lowering of F3 for /r/. Assuming that the front cavity is smaller for /r/ followed by the FLEECE vowel than it is for /r/ followed by LOT, in order to maintain a stable acoustic output for /r/ across all vowel contexts, speakers may compensate by using varying amounts of lip protrusion. /r/ followed by the FLEECE vowel may exhibit more protrusion than more open, back vowels, although we do not yet know to what extent the labial properties of neighbouring vowels have a coarticulatory influence on the lips for /r/.

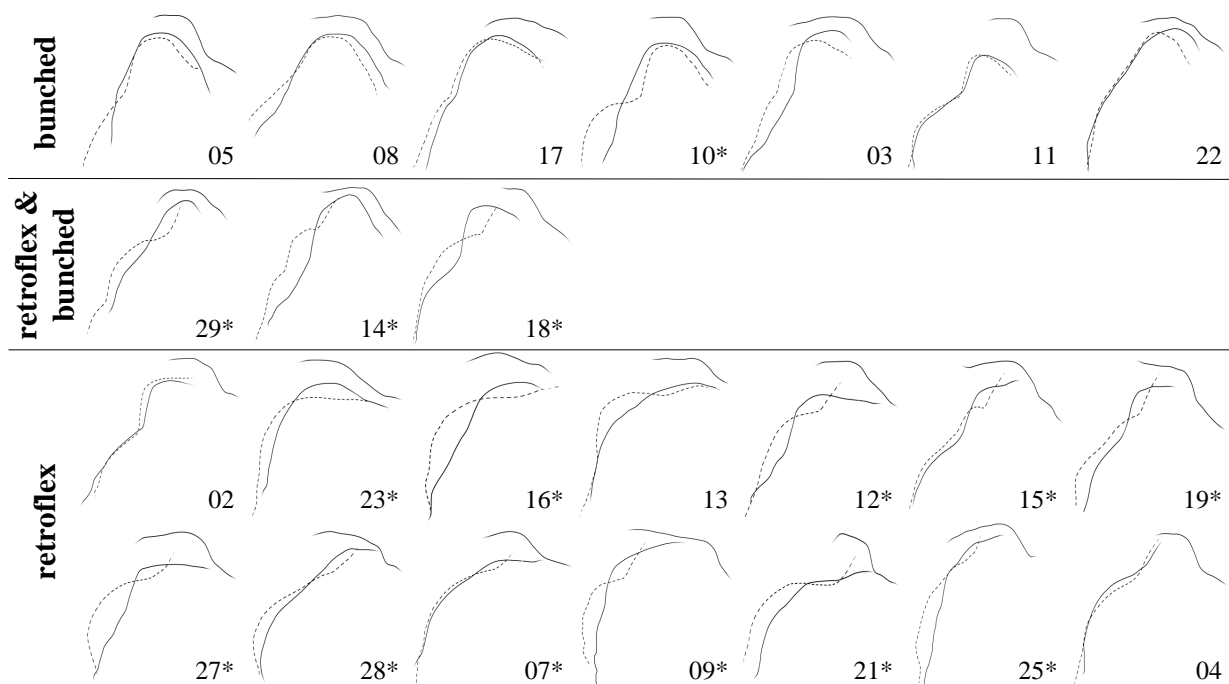


Figure 9: Tongue contour tracings ordered from most bunched to most retroflex for /r/ preceding the FLEECE (solid line) and the LOT vowel (dashed line). Speakers who use more than one of the five tongue configurations are indicated with an asterisk. The tongue tip is at the right side of the image. The palate is traced in the top curve for each speaker.

### 3.2. The influence of tongue shape on lip protrusion

In the three speakers who produced both retroflex and bunched /r/ configurations, the bunched variants had on average more lip protrusion than retroflex ones, as presented in the plots in Figure 10, which include the mean and standard deviation where possible (speaker 18 only produced one bunched token). This result therefore suggests that the degree of lip protrusion may be dependent on tongue shape, with bunched tongue shapes exhibiting more accompanying protrusion than retroflex ones.

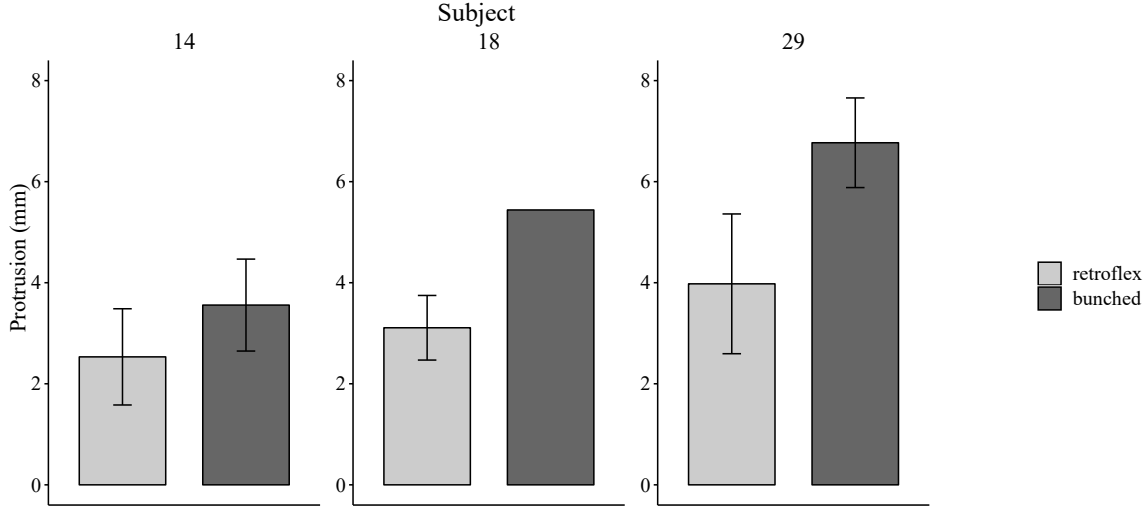


Figure 10: Mean and standard deviation protrusion values in the three speakers who produce both retroflex and bunched tongue configurations in millimetres.

In order to assess whether different tongue configurations are accompanied by different degrees of lip protrusion for /r/ in all speakers, a linear mixed-effects regression analysis was performed. The fixed factors were /r/ Coding (CU, TU, FU, FB, MB) and Vowel (FLEECE, GOOSE, KIT, DRESS, TRAP, STRUT, THOUGHT, LOT) and the random structure included by-Speaker random intercepts<sup>4</sup>. There was a statistically significant main effect of both tongue configuration ( $\chi^2(4) = 29.74, p < 0.001$ ) and following vowel ( $\chi^2(7) = 34.28, p < 0.001$ ) on lip protrusion. The final model output is presented in the model summary in Table 2.

As Table 2 indicates, the bunched tongue configurations (FB and MB) are predicted to have significantly more lip protrusion than the extreme Curled Up retroflex. Although FB is predicted to have more protrusion, by changing the reference level to FB and rerunning the model, we found no significant difference between FB and MB. There was no significant difference between the Curled Up retroflex and the other two retroflex configurations (TU & FU). Figure 11 presents the predicted effects of tongue configuration for /r/ on lip protrusion. We observe that the three retroflex configurations pattern together with the least protrusion, as do the two remaining bunched ones, with the most protrusion. As discussed in 3.1, the Front Up configuration seems to lie somewhere in the middle of the retroflex-bunched continuum with regards to its lingual characteristics. However, we notice that with regards to lip protrusion, Front Up strongly patterns with the Curled Up and Tip Up retroflex configurations. This result further justifies our decision to consider the Front Up configuration a retroflex and not a bunched shape.

<sup>4</sup>The inclusion of by-item varying intercepts resulted in a singular fit, presumably because, given the limited dataset, the main effect of vowel captures all the item variance, as pointed out by an anonymous reviewer.

Predictor	Estimate	Std. Error	<i>t</i> value	<i>p</i> value
(Intercept)	2.15	0.47	4.61	< 0.001
/r/ Coding TU	-0.005	0.26	-0.02	0.99
/r/ Coding FU	-0.37	0.37	-1.00	0.32
/r/ Coding FB	2.03	0.42	4.79	< 0.001
/r/ Coding MB	1.40	0.56	2.51	0.02
Vowel GOOSE	0.13	0.26	0.51	0.61
Vowel KIT	-0.67	0.26	-2.60	0.01
Vowel DRESS	-0.74	0.27	-2.75	0.01
Vowel TRAP	-0.48	0.27	-1.80	0.08
Vowel STRUT	-0.11	0.27	-0.39	0.70
Vowel THOUGHT	0.15	0.28	0.55	0.59
Vowel LOT	0.38	0.29	1.32	0.19

$$Protrusion \sim rCoding + Vowel + (1|Speaker)$$

Table 2: Output of a linear-mixed effects regression model of lip protrusion. The intercept corresponds to a CU tongue configuration preceding the FLEECE vowel

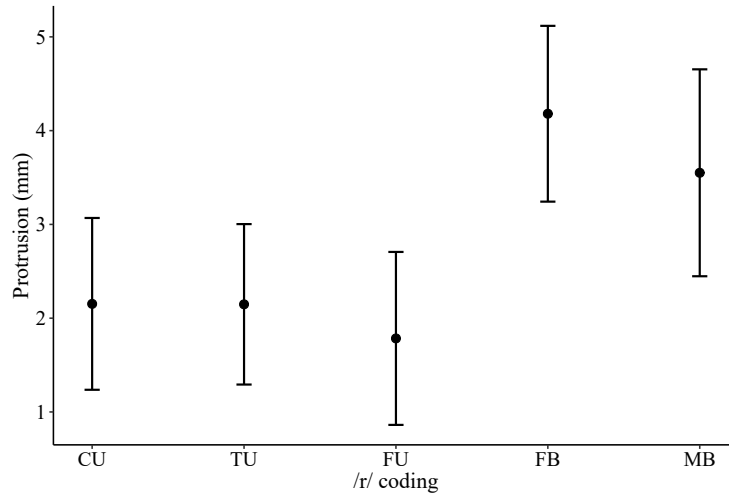


Figure 11: Predicted effects of tongue configuration on lip protrusion. Error bars are 95% confidence intervals.

With regards to the effect of the following vowel on lip protrusion for /r/, the model predicts that the KIT and DRESS vowels have significantly less protrusion than the FLEECE vowel. No significant difference is predicted between the FLEECE vowel and the remaining vowels in the dataset (GOOSE, TRAP, STRUT, THOUGHT, LOT). Figure 12 presents the predicted effects of the following vowel on protrusion in /r/.

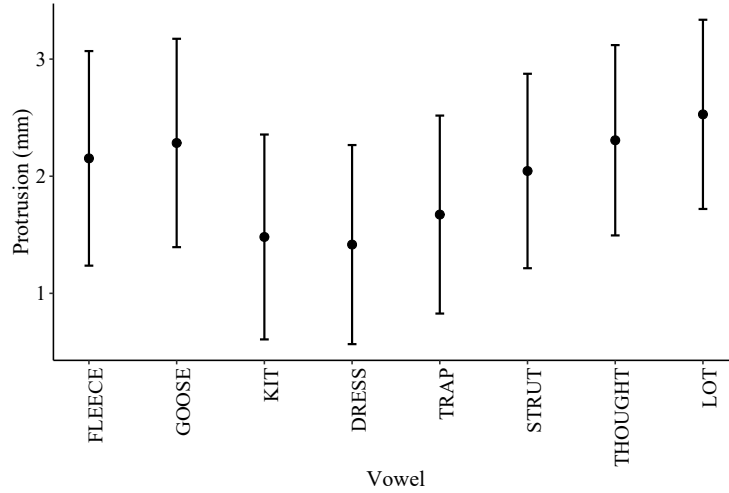


Figure 12: Predicted effects of following vowel on lip protrusion in /r/. Error bars are 95% confidence intervals.

### 3.3. Acoustics

As our dataset contains limited data from male speakers ( $n=2$ ) and as it is well established that speaker sex influences formant values, we only consider data from the remaining female speakers ( $n=22$ ) in our acoustic analysis. Across all productions of /r/ in women, the following mean formant values and their standard deviations (in Hz) were observed:

- F1:  $421.36 \pm 65.11$
- F2:  $1236.18 \pm 223.61$
- F3:  $1881.14 \pm 198.07$

Mean formant values are consistent with the range of values observed in previous studies on /r/ in American English (as presented in 1.3). Table 3 shows mean formant values (in Hz) according to tongue shape. Previous research on rhotic Englishes has not found a significant difference in F3 between the different possible tongue configurations for /r/. However, the mean formant values in our dataset do suggest that there may be differences across tongue shapes, notably with regards to FB, which has a lower mean F3 than the other four shapes. This difference is also apparent from the box plots of raw F3 values for each of the five tongue configurations presented in Figure 13. The median value of FB is lower than all the other tongue configurations and although the interquartile range is small, FB has the most outliers.

To test whether there are statistically significant differences in F3 for /r/ between the different tongue configurations and the following vowel, we performed a linear mixed-effects analysis. The fixed factors were

/r/ coding	F1	F2	F3
CU	435 $\pm$ 71	1 158 $\pm$ 212	1 851 $\pm$ 184
TU	419 $\pm$ 71	1 253 $\pm$ 247	1 914 $\pm$ 186
FU	442 $\pm$ 66	1 318 $\pm$ 209	1 960 $\pm$ 217
FB	399 $\pm$ 46	1 254 $\pm$ 227	1 761 $\pm$ 184
MB	411 $\pm$ 54	1 279 $\pm$ 147	2 026 $\pm$ 116

Table 3: Mean formant values and standard deviation (in Hertz) for all tongue shapes from most retroflex to most bunched in women.

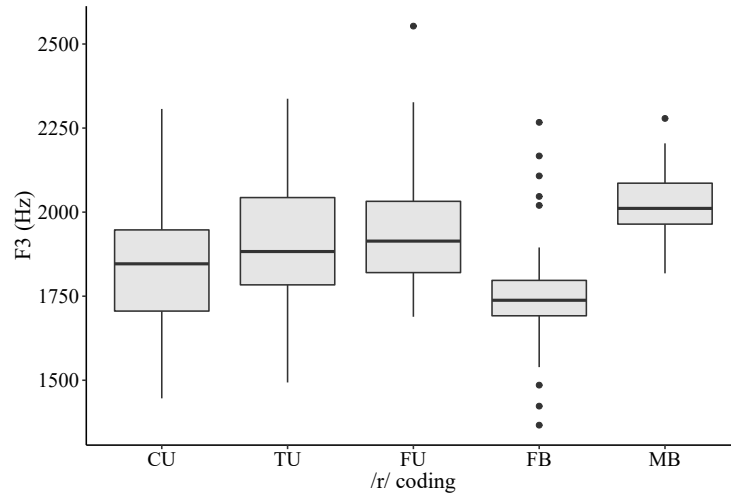


Figure 13: Box plots of raw F3 values for each of the five tongue configurations. The boxes (here and in all subsequent box plots) represent the interquartile range containing the middle 50% of values. Whiskers extend to the highest and lowest values, excluding outliers (in circles). A line across the box indicates the median.

/r/ Coding (CU, TU, FU, FB, MB) and Vowel (FLEECE, GOOSE, KIT, DRESS, TRAP, STRUT, THOUGHT, LOT) and the random structure included by-Speaker random intercepts. Likelihood ratio tests revealed that there was a statistically significant effect of the following vowel on F3 ( $\chi^2(7) = 52.13, p < 0.001$ ) but not of tongue configuration ( $\chi^2(4) = 4.32, p = 0.36$ ). The final model output is presented in the model summary in Table 4. All vowels are predicted to have a significantly lower F3 than the FLEECE vowel. The lowest F3 values are predicted to occur in /r/ followed by the back vowels in THOUGHT and LOT. Furthermore, our results are in line with previous work on English /r/ because tongue configuration was not a statistically significant factor, contrary to what the raw mean values would indicate. When individual variation is taken into account, any apparent differences in F3 between tongue configurations disappear. Indeed, the model’s marginal  $R^2$ , which is the variance described only by the main effects is 25.03%. The conditional  $R^2$ , which is the variance described by the main *and* the random effects is much higher at 61.48%<sup>5</sup>. The model also predicts speaker intercepts to range from 1 838 to 2 294 Hz.

Predictor	Estimate	Std. Error	<i>t</i> value	<i>p</i> value
(Intercept)	2 037.16	49.68	41.01	< 0.001
/r/ Coding TU	19.83	34.19	0.58	0.57
/r/ Coding FU	62.99	49.56	1.27	0.21
/r/ Coding FB	9.22	52.91	0.17	0.87
/r/ Coding MB	128.32	69.55	1.84	0.07
Vowel GOOSE	-151.02	36.89	-4.09	
Vowel KIT	-156.63	37.35	-4.19	
Vowel DRESS	-172.31	38.36	-4.49	
Vowel TRAP	-240.69	38.22	-6.30	< 0.001
Vowel STRUT	-226.11	39.14	-5.78	
Vowel THOUGHT	-259.69	40.45	-6.42	
Vowel LOT	-271.39	40.99	-6.62	

$F3 \sim rCoding + Vowel + (1|Speaker)$

Table 4: Output of a linear-mixed effects regression model of F3. The intercept corresponds to a CU tongue configuration preceding the FLEECE vowel

### 3.4. Labial articulation of /r/ and /w/

One speaker was excluded from this analysis because the camera angle in the frontal lip view made it difficult to view her top lip. As lip spreading and aperture were not measured in world units, all three lip dimensions were transformed into the percentage of change relative each speaker’s neutral lip setting. Table 5 presents mean percentage change for lip protrusion, spreading and aperture and their standard deviations for the 23 speakers analysed according to phoneme. On average, /r/ and /w/ involve an increase in lip protrusion and aperture compared to a neutral lip setting, although protrusion and aperture are greater in /w/ than in /r/. The most striking difference between /r/ and /w/ lies in lip spreading. While /r/ virtually does not change from the neutral setting (less than 0.1% on average), there is nearly 12% less spreading in /w/ than in the neutral setting, indicating that the lips are compressed. All three dimensions exhibit variation, which is probably due to inter-speaker differences. Lip aperture was particularly challenging to

<sup>5</sup>Conditional and marginal  $R^2$  were calculated using the `r.squaredGLMM()` function in the MuMIn package (Barton, 2018).

measure as the vermilion border of the top and bottom lip is not always evident in some speakers, which is perhaps reflected in the particularly high variability observed in this measure in comparison to the other two. This variability can also be seen in the box plots in Figure 14.

Phoneme	Protrusion	Spreading	Aperture
/r/	13.18% $\pm$ 10.25	0.07% $\pm$ 3.77	16.20% $\pm$ 14.92
/w/	18.69% $\pm$ 11.94	-11.92% $\pm$ 8.34	24.51% $\pm$ 19.65

Table 5: Mean and standard deviation percentage changes from a neutral lip posture for lip protrusion, spreading and aperture for /r/ and /w/.

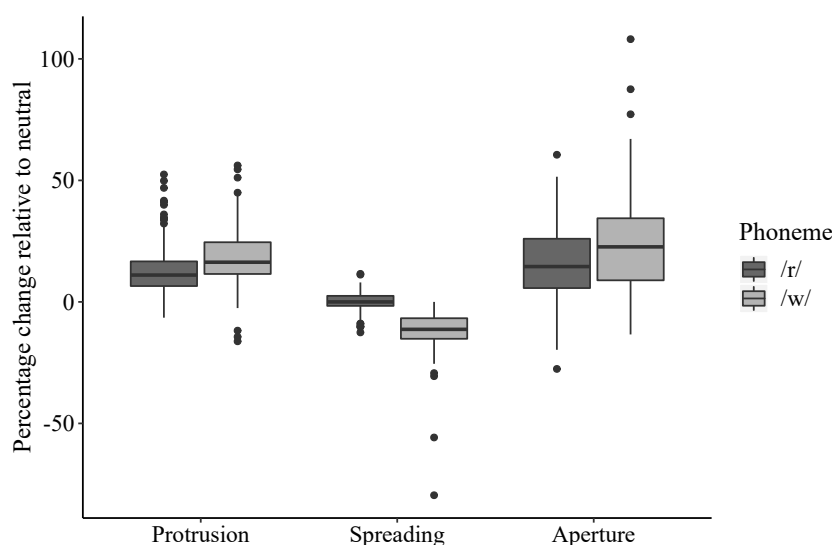


Figure 14: Box plots of percentages change from the neutral lip setting in protrusion, spreading and aperture for /r/ and /w/.

To test whether there are statistically significant differences between the labial posture of /r/ and /w/, we performed a generalised linear mixed-effects regression analysis with the phoneme (/r/ or /w/) as the binary outcome variable (/r/ coded as 0, /w/ as 1). The fixed factors were percentage change from the neutral lip setting in Protrusion, Spreading and Aperture, which were mean centred to improve model fit. The random structure included by-Speaker and by-Vowel varying intercepts. Likelihood ratio tests revealed that Spreading was the only statistically significant main predictor of phoneme ( $\chi^2(1) = 455.01, p < 0.001$ ). The other lip dimensions were not significant (Protrusion:  $\chi^2(1) = 0.81, p = 0.37$ ; Aperture:  $\chi^2(1) = 1.08, p = 0.30$ ). The final model output presented in Table 6 indicates that for an average speaker, the log-odds of observing a /w/ are -77.39 lower when lip spreading increases. These results suggest that /w/ has significantly less spreading, i.e., more horizontal compression, than /r/.

Figure 15 presents example frontal view images of /r/ and /w/ from 12 subjects grouped according to their tongue shape for /r/ (bunched or retroflex). Images were taken from productions of /r/ and /w/ followed by the FLEECE vowel, i.e., from the words *reed* and *weed*. Each subject’s left hand image corresponds to their lip posture for /r/. In nearly all subjects, the lip configurations are visibly different for /r/ and /w/, which is consistent with Brown’s observations that their lip postures differ (Brown, 1981).



Predictor	Estimate (log-odds)	Std. Error	<i>t</i> value	<i>p</i> value
(Intercept)	4.66	3.60	1.29	0.20
Protrusion (centred)	5.42	6.81	0.80	0.43
Spreading (centred)	-77.39	31.33	-2.47	0.02
Aperture (centred)	-4.89	4.67	-1.05	0.30
<i>Phoneme ~ Protrusion + Spreading + Aperture + (1 Speaker) + (1 Item)</i>				

Table 6: Output of a linear-mixed effects logistic regression model predicting phoneme (/r/ vs. /w/) with the intercept corresponding to /w/.

Impressionistically, horizontal contraction seems to tense the lips which results in the appearance of numerous vertical wrinkles across the red parts of the lips. These wrinkles are generally absent or much less apparent for /r/. Furthermore, the shape of the mouth opening generally differs for /r/ and /w/. For /r/, the lip opening has a slit-like elliptical shape, while for /w/ the opening is smaller and circular. Visualising the data therefore indicates that /r/ has a lip posture which corresponds to Catford’s description of exolabial articulations, while /w/ is closer to endolabial ones (as discussed in 1.2). We note that speaker 04, a retroflexer, is the only speaker whose lip configurations for /r/ and /w/ are somewhat similar: both have a small circular lip opening with a certain degree of wrinkling of the lip surface. Incidentally, this subject uses the most amount of horizontal contraction on average for /r/ according to our quantitative analysis, although horizontal compression was still larger for /w/.

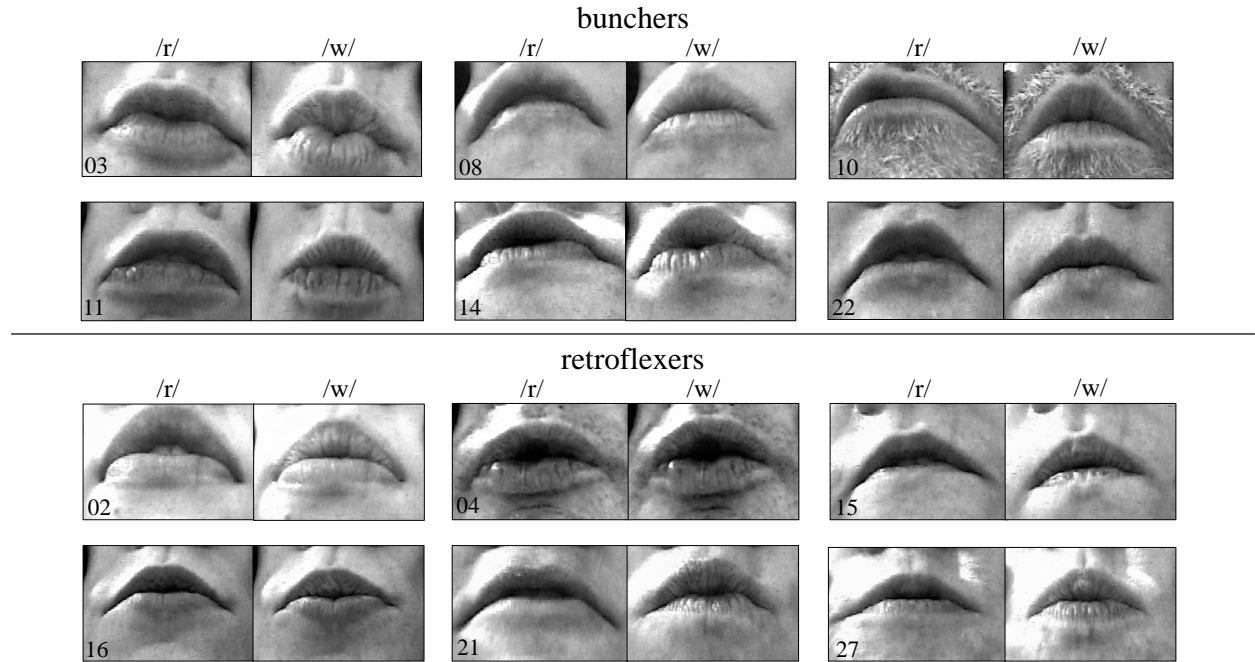


Figure 15: Frontal view lip images from 6 bunchers and 6 retroflexers for /r/ (left image) and /w/ (right image). Images were taken from /r/ and /w/ productions followed by the FLEECE vowel

### 3.5. Summary of results

Putting together the various analyses from this section, the following findings emerge. Firstly, Anglo-English /r/ may be produced with a range of tongue shapes from curled-up retroflex (CU) to tip down bunched (MB), although retroflexion is more common than bunching. 3 subjects who come from the South East of England produce both retroflex and bunched configurations, while the remaining 21 subjects who come from all over England use either retroflex or bunched shapes. Given the lack of geographically-stratified data presented here, we cannot comment on any potential regional patterns regarding tongue shape for /r/. In retroflex users, our results suggest that the degree of retroflexion is related to the quality of the following vowel. The close front FLEECE vowel, appears to be the least compatible with retroflexion, contrary to the open back LOT vowel. In the three speakers who presented both retroflex and bunched tongue shapes, bunching was only utilised in conjunction with the frontest vowels of the system. Although speakers who use exclusively bunched shapes tend to have acquired one distinct tongue shape for /r/, one speaker produces a different, arguably more bunched tongue shape (with an even lower tongue tip) in the context of /r/ followed by the FLEECE vowel. Furthermore, tongue contour tracings revealed that even in speakers who use one distinct shape for /r/, the following vowel has a co-articulatory influence because the tongue is generally more anterior for /r/ followed by the front FLEECE vowel than /r/ followed by the back LOT vowel.

Our analysis suggests that the degree of lip protrusion for /r/ may be related to both tongue shape and the following vowel. According to our statistical analysis, bunched tongue shapes have significantly more lip protrusion. Productions of /r/ followed by the rounded vowels in LOT, THOUGHT and GOOSE are predicted to have the most lip protrusion of all the vowels, suggesting there is a co-articulatory influence of the labial properties of the following vowel on /r/. However, no significant difference in lip protrusion is predicted between /r/ followed by the FLEECE vowel and /r/ followed by the rounded vowels in LOT, THOUGHT and GOOSE, which is unexpected given that the FLEECE vowel is non-rounded. Finally, our results suggest that what distinguishes the lip posture for /r/ from that of /w/ is the degree of horizontal compression at the lip corners. While the lip corner dimension for /r/ does not vary on average from that of a neutral lip posture, the space between the lip corners decreases by nearly 12% on average for /w/, indicating a contraction of the lip corners compared to a neutral lip setting. Lip protrusion and lip aperture were not significant predictors of phoneme, /r/ versus /w/. Qualitatively, frontal lip images indicate that /r/ is generally produced with exolabial rounding while /w/ is endolabial.

## 4. Discussion

### 4.1. Articulation of Anglo-English /r/

As is the case for English /r/ in other varieties, Anglo-English presents a range of possible tongue shapes for /r/ from Mid Bunched to Curled Up retroflex. However, the production of Anglo-English /r/ differs from the results from recent studies on American English in that retroflexion is much more common in Anglo-English. For example, out of 27 subjects, Mielke et al. (2016) only observed 2 producing exclusively retroflex tokens in both pre- and post-vocalic /r/, compared to our 14/24 subjects in prevocalic /r/. Although their classification would consider our Front Up configuration to be bunched and not retroflex, if we do the same, our Anglo-English data still have far more exclusively retroflex users (25%) than the American English data (<8%). The difference in results may also reflect the fact that our data are limited to word-initial /r/,

whereas Mielke et al. (2016) also included prevocalic /r/ in onset clusters. However, Mielke et al. (2016) observed the highest rates of retroflexion to occur in the same prevocalic syllable-initial context used in the present study. There does therefore appear to be a difference between American English and Anglo-English /r/: Anglo-English /r/ is far more likely to be produced with retroflexion.

More frequent retroflexion has also been observed in non-rhotic New Zealand English. In a large-scale ultrasound study of 62 New Zealand English speakers, nearly 20% of subjects produced exclusively retroflex tongue shapes (Heyne et al., 2018). Like Mielke et al. (2016), Heyne et al. (2018) also considered the equivalent of our Front Up classification to be bunched and not retroflex. If we do the same, the percentage of exclusively retroflex users in Anglo-English (25%) and New Zealand English (nearly 20%) are remarkably consistent. It appears then that exclusively retroflex tongue shapes are up to three times more frequent in non-rhotic than in rhotic Englishes. Heyne et al. (2018) speculate that as New Zealand English speakers very rarely produce /r/ in postvocalic environments, where bunching is heavily favoured, speakers are less likely to acquire bunched /r/ as an alternative articulation strategy if they have already mastered retroflexion. Our Anglo-English data seem to support this hypothesis. Future studies could consider to what extent the production of /r/ varies in children acquiring rhotic and non-rhotic Englishes.

Although retroflexion is generally more frequent in non-rhotic than in rhotic English speakers, the rate of retroflexion is influenced by co-articulation with neighbouring segments. In the present study, retroflexion is favoured by open back vowels versus close front ones, in a similar fashion to American English (Ong & Stone, 1998; Mielke et al., 2016; Tiede et al., 2010). The incompatibility of retroflexion with close front vowels, notably in the FLEECE, KIT and GOOSE vowels, is manifested through the use of less extreme retroflex variants, i.e., less curling back of the tongue tip, less tongue tip raising, and more bunching. This shift from extreme retroflexion towards more bunched configurations in close front vowel contexts further strengthens the argument that the possible tongue shapes for /r/ are on a continuum rather than the initial suggestion of dichotomous categories (Uldall, 1958). The fact that retroflexion is not compatible with close front vowels is perhaps not surprising as it has been suggested that retroflex sounds are always produced with a retracted tongue body (Hamann, 2002) and as a result, vowels which are also produced with a retracted tongue body, i.e., back vowels, are more compatible. However, bunched /r/ has also been associated with a retraction of the tongue. For example, Delattre & Freeman (1968) discuss the narrowing of the vocal tract in the pharyngeal region and much more recently, a retraction of the tongue body towards the lower rear pharyngeal wall was observed in all word-initial rhotics in a real-time magnetic resonance imaging study of four native American English speakers (Proctor et al., 2019). In the present study, speakers who present both retroflex and bunched shapes produce bunched tokens only in the context of a close front vowel, particularly with the FLEECE vowel. As both retroflex and bunched configurations are retracted, retraction cannot be the only articulatory property which makes retroflexion incompatible with front vowels. As Hamann (2003) suggests, the tongue shape for [i] which involves the tip being tucked under the lower front teeth is inherently incompatible with that of retroflexion. Unlike in retroflexes, the tongue tip remains relatively low in the mouth for bunched /r/, which is perhaps why bunching is more compatible with high front vowels than retroflexion. In one buncher (speaker 10), /r/ preceding all vowels except for the FLEECE vowel were produced with a Front Bunched configuration. /r/ before FLEECE, however, was produced with a Mid Bunched configuration. We observed from tongue contour tracings that the Mid Bunched configuration has a lower tongue tip to the Front Bunched one in speakers who present both bunched shapes, which would

thus explain why the Mid Bunched shape with a lower tongue tip is preferred in the context of the FLEECE vowel. It therefore seems natural to consider the Mid Bunched category to be the most bunched tongue configuration, despite the fact that bunching, which is generally associated with a dip in the tongue surface, is less apparent than in the Front Bunched shape. We therefore conclude that our continuum ranges from tip down Mid Bunched, most compatible with high close vowels, to tip-up Curled Up retroflex, most compatible with low open ones.

A novel finding of this study is that the degree of accompanying lip protrusion may be influenced by tongue configuration. Bunched tongue configurations are predicted to have significantly more lip protrusion than retroflex ones. As discussed in 1.3, retroflex consonants, by definition, include the addition of a sublingual space, which increases the volume of the front cavity, thus lowering the third formant. Bunched /r/ involves the tongue tip being positioned relatively low in the mouth and therefore presumably creates less space underneath the tongue tip. The difference we observe regarding the degree of lip protrusion could thus be a compensation strategy used by bunchers to lengthen the front cavity in order to obtain the same sized front cavity and therefore, the same acoustic output as retroflexers. Indeed, we observed no statistically significant difference across tongue configurations in F3.

Our analysis also indicates that the use of lip protrusion as a compensation strategy may go beyond the bunched-retroflex distinction. Although our results generally support Gimson (1980)’s observation that /r/ productions in the context of rounded vowels present more lip protrusion than in the context of non-rounded vowels, labial coarticulation cannot account for the fact that in the context of the close front FLEECE vowel, /r/ is predicted to have significantly more lip protrusion than in the context of the more open non-rounded vowels such as those in KIT and DRESS. Labial coarticulation does also not account for the lack of a statistically significant difference in lip protrusion between /r/ followed by the FLEECE vowel and the rounded vowels in LOT, THOUGHT and GOOSE. Visualising tongue contour tracings revealed that /r/ preceding the FLEECE vowel is generally produced with a more anterior tongue position than /r/ preceding LOT, no doubt due to lingual co-articulation. As this fronting of the tongue will presumably result in the shortening of the front cavity, speakers may again compensate for this shortening by increasing lip protrusion, thus extending the front cavity, regardless of underlying tongue shape. A limitation to our analysis is that in the present dataset, place of articulation and rounding are partly confounded: the only non-rounded back vowel is the STRUT vowel, which may actually be realised as the rounded [ʊ] in speakers who do not present the FOOT-STRUT split, i.e., in linguistic Northerners, who as it happens, make up the majority of the dataset (n=16). Despite our reservations, compensation strategies for co-articulation with front vowels in retroflexes have been observed in other languages. For example, the vowel /i/ was rounded preceding retroflexes in Wembawemba, an extinct Indigenous Australian language, but not in other vowel contexts (Flemming, 2013). It is interesting to note that despite the higher degree of lip protrusion, /r/ preceding the FLEECE vowel still results in significantly higher predicted F3 values than /r/ preceding all other vowels in the dataset. It seems then that increased lip protrusion does not necessarily result in complete compensation for lingual co-articulation with the FLEECE vowel.

We stress that although our data point towards a possible articulatory compensation strategy involving the use of lip protrusion to extend the front cavity for /r/, more articulatory data, ideally from a more robust imaging technique which would provide vocal tract dimensions i.e., magnetic-resonance imaging, is evidently required. Indeed, another limitation to our study is the fact that the sublingual space is not visible

from ultrasound data. Furthermore, there may well be a three-way trading relation between the size of the sublingual space, palatal constriction location and degree of lip protrusion, which falls outside the scope of this paper. Although we have focused on Anglo-English, we see no reason why the use of lip protrusion as a compensation strategy for /r/ could not be extended to other varieties of English, which could also be the object of further study.

Given the significant differences in lip protrusion we have observed between retroflex and bunched tongue configurations, future studies could consider whether this difference is perceptibly salient to an interlocutor in both the auditory and visual domains. Furthermore, although some clues may lie in higher formant values, without the use of advanced and rather expensive instrumental techniques capable of imaging or tracking the tongue, researchers are not yet capable of telling a bunched /r/ from a retroflex one. Visualising the lips, however, can be accomplished with ease, and could therefore be an alternative, more cost-effective strategy. However, we again stress the need for further research verifying our claim that bunched /r/s are inherently more protruded than retroflexes.

#### *4.2. Accounting for the labial gesture in Anglo-English /r/*

Quantitative analysis of the profile and frontal lip images indicates that what distinguishes the lip postures for /r/ and /w/ is the horizontal dimension (i.e., lip corner to corner) of the interlabial space. Lip protrusion and lip aperture were not significant predictors of phoneme category, /r/ or /w/, although both dimensions were higher on average for /w/. We observed that the horizontal dimension of the lips for /r/ remains very similar to a neutral lip posture. However, the space between the lip corners decreases by nearly 12% on average from the neutral setting for /w/. Impressionistically, Catford's account of endolabial versus exolabial articulations (as discussed in 1.2) seems to rather accurately describe the different lip postures we observe for /r/ and /w/ in nearly all of our subjects. /w/ is endolabial because the lip corners are pushed to the centre forming a round shaped opening between the lips. /r/ is exolabial because, rather than bringing the lip corners to the centre, they are compressed vertically, which creates an elliptical shaped lip opening.

Our results therefore indicate that the lip postures for /r/ and /w/ may be phoneme specific in Anglo-English. Although it is well-established that lip rounding and lip protrusion cause formant frequencies to decrease because they increase the length of the vocal tract (Stevens, 1998; Vaissière, 2007), the exact acoustic consequences of the different lip postures for /r/ and /w/ are not clear. While the main acoustic correlate of /r/ is generally associated with a low F3 in close proximity to F2, the labio-velar approximant /w/ is characterised by a high F3 and a low F2 (Espy-Wilson, 1992). Catford explains that front vowels are usually exolabial, i.e., without horizontal compression, in order to avoid over-lowering the second formant and hence preserve their front quality (Catford, 1977, p.173). Similarly, Stevens notes that in the case of a backed tongue position, the condition of minimum F2 is achieved only if the lips are rounded and a narrow opening is formed (Stevens, 1998, pp. 280-281). We suggest then that by limiting their use of horizontal compression for /r/, Anglo-English speakers avoid over-lowering the second formant, thus conserving the proximity between the second and third formant for /r/ and ensuring a maximal perceptual contrast between /r/ and /w/.

Somewhat unexpected differences have been observed in the perception of approximants between American and Anglo-English listeners. In Dalcher et al. (2008), American and English participants judged whether copy-synthesised sounds with manually adjusted formant values were more like /r/ or /w/. A significant difference was observed for a stimulus which had a third formant typical of /r/ and second formant typical of

/w/. American speakers identified this stimulus as /r/ 90% of the time, while Anglo-English speakers only identified it as /r/ 59% of the time. Dalcher et al. (2008) argue that the reason for such a disparity may be due to Anglo-English speakers being exposed to labiodental variants without a canonically low F3, unlike American English speakers. As a consequence, they speculate that F3 alone is no longer a sufficient cue to distinguish /r/ from /w/ in Anglo-English and that the F2 boundary between /r/ and /w/ may have become sharper in Anglo-English speakers. The fact that the vast majority of the Anglo-English speakers presented in this paper use a lip configuration that potentially prevents them from over-lowering F2 (i.e., exolabially, with very little horizontal compression) seems to support Dalcher et al. (2008)’s hypothesis. Although all our speakers had an observable tongue body gesture with low F3 values typical of /r/, given the pressure to differentiate /r/ and /w/ beyond F3 due to exposure to high-F3 variants, Anglo-English speakers may find themselves in a delicate articulatory balancing act, having to make trade-offs between keeping F3 low without over-lowering F2. As F2 is less of a concern, we predict that American English speakers would be freer to use more variable, more /w/-like lip postures for /r/ in order to enhance r-saliency. The findings from a very recent study on American English support this hypothesis. Labial postures presented in Smith et al. (2019) were much more variable across speakers, with more instances of endolabial articulations reported for /r/ than in our Anglo-English data.

An alternative explanation for the observed difference in labial configurations between /r/ and /w/ in Anglo-English could be that by using distinctive articulatory cues, speakers are able to enhance the perceptual contrast between the two sounds in the visual domain. Indeed, speech has been shown to be visually optimised in cases where pressure to maintain a phonological contrast is high. For example, Havenhill & Do (2018) observed that in American English, the visual lip rounding cue enhances perception of the /ɑ/-/ɔ/ contrast, and Traunmüller & Öhrström (2007) found that in Swedish, listeners rely on visual cues in the perception of /i/-/y/. Future research could consider whether the different visual cues for /r/ and /w/ are perceptibly salient to Anglo-English speakers.

Finally and perhaps somewhat ironically, the lip posture we have described for /r/ in Anglo-English, which is potentially used by speakers to enhance F3 lowering all the while avoiding over-lowering F2, seems to share similar features to labiodental articulations. In order to protrude the lips without horizontal compression, the lower lip is raised towards or even beyond the level of the top front teeth, which is described as vertical compression of the lip corners by Catford (1977). The lips are also everted revealing the soft inner surfaces, as discussed in Brown (1981). The inner surface of the lower lip is thus in close proximity with or perhaps even touching the upper front teeth. In some speakers, the upper front teeth are even visible during their /r/ production (e.g., speakers 02, 08, 11, 15, 21 and 22 presented in Figure 15). On the other hand, horizontal compression of the lips corners in /w/ draws the corners of the mouth together away from the front teeth along the occlusal plane, making contact between the lips and front teeth almost impossible. Indeed, the teeth were never visible in any of the /w/ tokens. We speculate then that the lip posture observed for /r/ in Anglo-English may result in the approximation of the lower lip and the top teeth, or labiodentalisation. Labiodental variants could thus continue to emerge if the labial gesture takes precedence over the lingual one, as suggested by Docherty & Foulkes (2001), particularly if the labial gesture is visually prominent. As a result, like Dalcher et al. (2008), we also predict an increase in labiodental /r/ in Anglo-English.

## 5. Conclusions

Articulatory data presented in this paper have shown that Anglo-English /r/ is not only produced with retroflexion but presents similar lingual variation to that observed in rhotic Englishes with tongue shapes ranging from tip down bunched to curled-up retroflex. However, retroflexion is three times more frequent in Anglo-English than American English, which may be a direct consequence of the absence of postvocalic /r/ productions in Anglo-English, a context which reportedly favours bunching, as discussed by Heyne et al. (2018). Although some speakers present one configuration exclusively, in others, tongue shape may be directly related to the following vowel with tip-up variants favouring low open vowel contexts and tip down ones favouring high close ones. A novel finding of this study is that the degree of accompanying lip protrusion may be directly related to the size of the front cavity in Anglo-English with smaller front cavities presenting the most lip protrusion. Tip-down tongue shapes, which have less space underneath the tongue than tip-up ones, appear to compensate for their smaller cavity volume through increased lip protrusion. Lingual co-articulation with neighbouring front vowels may reduce the size of the front cavity for /r/ regardless of tongue shape, for which speakers also seem to compensate via increased lip protrusion. We therefore conclude that lip protrusion is an articulatory mechanism used to enhance the acoustic saliency of /r/. Pressure to maintain a perceptual contrast between /r/ and /w/ due to increased exposure to high-F3 labiodental variants of /r/ in Anglo-English may have resulted in the development of a specific labial gesture for /r/, which enables speakers with an observable tongue body gesture to maintain a low F3 without over-lowering F2. Over-lowering of F2 could cause perceptual uncertainty as the acoustic cue that distinguishes a high-F3 /r/ from /w/ may be F2 (Dalcher et al., 2008). In Englishes where high-F3 variants are not reported, the frequency of F3 remains the most prominent acoustic cue for /r/ (Dalcher et al., 2008), which we predict allows speakers more freedom to vary the accompanying lip gesture for /r/, which may account for the differences observed between the labial gesture in the present study and that presented in Smith et al. (2019) on American English. Finally, in avoiding over-lowering F2 due to increased exposure to labiodental /r/, the lip posture in speakers who do not use labiodental /r/ (i.e., with an observable tongue body gesture) has perhaps inadvertently become more labiodental. Following Dalcher et al. (2008), we also predict a further increase in labiodentalisation in Anglo-English /r/. The cue for /r/ in Anglo-English will continue to shift to F2 to such an extent that speakers will attend less to F3, provoking them to retain the labiodental component of their articulation at the expense of the lingual one.

## Acknowledgements

We would like to thank Eleanor Lawson, Jim Scobbie, Alan Wrench, Steve Cowen and the rest of the Clinical Audiology, Speech and Language Research Centre at Queen Margaret University for kindly allowing us to use their facilities and for their assistance with data collection. We particularly wish to thank Eleanor Lawson for her help with the analysis of lip camera data. We express our gratitude to Ioana Chitoran for her invaluable input throughout this project. Finally, we thank the assistant editor, Marianne Pouplier, and three anonymous reviewers for their comments and suggestions.

## Funding sources

The first author received a travel grant from Paris Diderot University for data collection.

## Competing interests statement

The authors have no competing interests to declare.

## Appendix A.

Lexical set	/r/-initial	/w/-initial
FLEECE	reed	weed
GOOSE	room	womb
KIT	ring	wing
DRESS	red	wed
TRAP	rack	whack
STRUT	run	won
THOUGHT	raw	war
LOT	rot	what

Table A.7: Test-words and corresponding lexical sets

## References

- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh, UK: Edinburgh University Press.
- Alwan, A., Narayanan, S., & Haker, K. (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. *The Journal of the Acoustical Society of America*, 101, 1078–1089. doi:10.1121/1.418030.
- Articulate Instruments Ltd. (2008). *Ultrasound Stabilisation Headset Users' Manual, Revision 1.4*. Edinburgh, UK: Articulate Instruments Ltd.
- Articulate Instruments Ltd. (2014). *Articulate Assistant Advanced Ultrasound Module User Manual, Revision 2.16*. Edinburgh, UK: Articulate Instruments Ltd.
- Ashton, H., & Shepherd, S. (2012). *Work on Your Accent*. London: Collins.
- Badin, P., Sawallis, T., & Lamalle, L. (2014). Comparaison des stratégies articulatoires d'un locuteur bilingue anglais-français: Données et modèles préliminaires. In *XXXièmes Journées d'Etudes Sur La Parole (JEP2014)*. Le Mans, France. URL: <https://hal.archives-ouvertes.fr/hal-01228883>.
- Bakst, S. (2016). Differences in the relationship between palate shape, articulation, and acoustics of American English /r/ and /s/. *UC Berkeley Phonology Lab Annual Report*, 12, 216–224.
- Barton, K. (2018). MuMIn: Multi-model inference. URL: <https://cran.r-project.org/web/packages/MuMIn/index.html>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. doi:10.18637/jss.v067.i01.
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer. Version 6.0.50, Accessed March 2019 from <http://www.praat.org/>.
- Brown, G. (1981). Consonant rounding in British English: The status of phonetic descriptions as historical data. In R. Asher, & E. J. Henderson (Eds.), *Towards a History of Phonetics* (pp. 67–76). Edinburgh, UK: Edinburgh University Press.
- Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh, UK: Edinburgh University Press.
- Catford, J. C. (1988). *A Practical Introduction to Phonetics*. Oxford, UK: Clarendon Press.
- Dalcher, C. V., Knight, R.-A., & Jones, M. J. (2008). Cue switching in the perception of approximants: Evidence from two English dialects. *University of Pennsylvania Working Papers in Linguistics*, Vol. 14: Iss. 2, Article 9.
- Dediu, D., & Moisik, S. R. (2019). Pushes and pulls from below: Anatomical variation, articulation and sound change. *Glossa: A Journal of General Linguistics*, 4(1), 1–33. doi:10.5334/gjgl.646.
- Delattre, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics*, 6, 29–68. doi:10.1515/ling.1968.6.44.29.



- Docherty, G. J., & Foulkes, P. (2001). Variability in (r) production - instrumental perspectives. In H. Van de Velde, & R. van Hout (Eds.), *'R-Atics: Sociolinguistic, Phonetic and Phonological Characteristics of /r/* (pp. 173–184). Brussels, Belgium: Université Libre de Bruxelles.
- Ehrlich, S., & Avery, P. (2013). *Teaching American English Pronunciation-Oxford Handbooks for Language Teachers*. Oxford University Press.
- Epstein, M. A., & Stone, M. (2005). The tongue stops here: Ultrasound imaging of the palate. *The Journal of the Acoustical Society of America*, 118, 2128–2131. doi:10.1121/1.2031977.
- Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English. *The Journal of the Acoustical Society of America*, 92, 736–757. doi:10.1121/1.403998.
- Espy-Wilson, C. Y., & Boyce, S. E. (1999). A simple tube model for American English /r/. In J. J. Ohala, Y. Hasegawa, D. Granville, & A. C. Bailey (Eds.), *Proceedings from the 14th International Congress of Phonetic Sciences* (pp. 2137–2140). San Francisco, CA: University of California, Berkeley. URL: [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14\\_2137.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_2137.pdf).
- Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S., & Alwan, A. (2000). Acoustic Modeling of American English /r/. *The Journal of the Acoustical Society of America*, 108, 343–356. doi:10.1121/1.429469.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague, Netherlands: Mouton.
- Ferragne, E., & Pellegrino, F. (2010). Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association*, 40, 1–34. doi:10.1017/S0025100309990247.
- Flemming, E. S. (2013). *Auditory Representations in Phonology*. Routledge. doi:10.4324/9781315054803.
- Foulkes, P., & Docherty, G. J. (2000). Another chapter in the story of /r/: 'Labiodental' variants in British English. *Journal of Sociolinguistics*, 4, 30–59. doi:10.1111/1467-9481.00102.
- Garnier, M., Ménard, L., & Richard, G. (2012). Effect of being seen on the production of visible speech cues. A pilot study on Lombard speech. In *INTERSPEECH-2012* (pp. 611–614).
- Gimson, A. (1980). *An Introduction to the Pronunciation of English*. London, UK: Arnold.
- Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., & Perkell, J. S. (1999). Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *The Journal of the Acoustical Society of America*, 105, 2854–2865. doi:10.1121/1.426900.
- Hagiwara, R. (1995). Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics*, 90, 1–187.
- Hamann, S. (2002). Retroflexion and Retraction revised. *ZAS working papers*, 28, 13–25.
- Hamann, S. (2003). *The Phonetics and Phonology of Retroflexes*. Ph.D. thesis LOT Utrecht, The Netherlands.
- Hancock, M. (2003). *English Pronunciation in Use*. Cambridge University Press.
- Harrington, J., Kleber, F., & Reubold, U. (2011). 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. doi:10.1017/S0025100310000265.
- Havenhill, J., & Do, Y. (2018). Visual Speech Perception Cues Constrain Patterns of Articulatory Variation and Sound Change. *Frontiers in Psychology*, 9, 728. doi:10.3389/fpsyg.2018.00728.
- Heyne, M., Wang, X., Derrick, D., Dorreen, K., & Watson, K. (2018). The articulation of /ɹ/ in New Zealand English. *Journal of the International Phonetic Association*, (pp. 1–23). doi:10.1017/S0025100318000324.
- Jones, D. (1972). *An Outline of English Phonetics*. (9th ed.). Cambridge, UK: Cambridge University Press.
- King, H., & Ferragne, E. (2018). /u/-fronting in English: How phonetically accurate should phonological labels be? In *16èmes Rencontres Du Réseau Français de Phonologie*. Paris, France.
- King, H., & Ferragne, E. (2019). The contribution of lip protrusion to Anglo-English /r/: Evidence from hyper- and non-hyperarticulated speech. *Proceedings of Interspeech 2019*, (pp. 3322–3326). doi:10.21437/Interspeech.2019-2851.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82. doi:10.18637/jss.v082.i13.
- Ladefoged, P., & Disner, S. F. (2012). *Vowels and Consonants*. John Wiley & Sons.
- Laver, J. (1980). *The Phonetic Description of Voice Quality: Cambridge Studies in Linguistics*. Cambridge, UK: Cambridge University Press.
- Lawson, E., Mills, L., & Stuart-Smith, J. (2015). Variation in tongue and lip movement in the goose vowel across British Isles Englishes. In *10th UK Language Variation and Change*. York, UK.
- Lawson, E., Scobbie, J. M., & Stuart-Smith, J. (2011). The social stratification of tongue shape for postvocalic /r/ in Scottish

- English. *Journal of Sociolinguistics*, 15, 256–268. doi:10.1111/j.1467-9841.2011.00464.x.
- Lawson, E., Scobbie, J. M., & Stuart-Smith, J. (2013). Bunched /r/ promotes vowel merger to schwa: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics*, 41, 198–210. doi:10.1016/j.wocn.2013.01.004.
- Lawson, E., Stuart-Smith, J., & Rodger, L. (2019). A comparison of acoustic and articulatory parameters for the GOOSE vowel across British Isles Englishes. *The Journal of the Acoustical Society of America*, 146, 4363–4381. doi:10.1121/1.5139215.
- Lawson, E., Stuart-Smith, J., & Scobbie, J. M. (2014). A mimicry study of adaptation towards socially-salient tongue shape variants. *University of Pennsylvania Working Papers in Linguistics*, 20, 12. URL: <https://repository.upenn.edu/pwpl/vol20/iss2/12>.
- Lawson, E., Stuart-Smith, J., & Scobbie, J. M. (2018). The role of gesture delay in coda /r/ weakening: An articulatory, auditory and acoustic study. *The Journal of the Acoustical Society of America*, 143, 1646–1657. doi:10.1121/1.5027833.
- Lawson, E., Stuart-Smith, J., Scobbie, J. M., Yaeger-Dror, M., & MacLagan, M. (2010). Analyzing liquids. In M. De Paolo, & M. Yaeger-Dror (Eds.), *Sociophonetics: A Student's Guide* (pp. 72–86). London, UK: Routledge.
- Lehiste, I. (1962). *Acoustical Characteristics of Selected English Consonants*. Ann Arbor, MI: University of Michigan Communication Sciences Laboratory.
- Lennon, R., Smith, R., & Stuart-Smith, J. (2015). An Acoustic Investigation of Postvocalic /r/ Variants in Two Sociolects of Glaswegian. In *18th International Congress of Phonetic Sciences*. Glasgow, UK. URL: <http://eprints.gla.ac.uk/109759/>.
- Lilly, R., & Viel, M. (1977). *La Prononciation de l'anglais: Règles Phonologiques et Exercices de Transcription*. Paris, France: Hachette.
- Lindley, N., & Lawson, E. (2016). An articulatory investigation of Anglo-English prevocalic /r/. In *BAAP Colloquium*. Lancaster, UK.
- Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America*, 49, 606–608. doi:10.1121/1.1912396.
- Lüdecke, D. (2018). sjPlot - Data Visualization for Statistics in Social Science. doi:10.5281/zenodo.1310947.
- Magloughlin, L. (2016). Accounting for variability in North American English /ɹ/: Evidence from children's articulation. *Journal of Phonetics*, 54, 51–67. doi:10.1016/j.wocn.2015.07.007.
- Marks, J. (2007). *English Pronunciation in Use Elementary Book with Answers, with Audio*. Cambridge University Press.
- Marsden, S. (2006). A sociophonetic study of labiodental /r/ in Leeds. *Leeds Working Papers in Linguistics and Phonetics*, (pp. 153–172).
- Mayr, R. (2010). What exactly is a front rounded vowel? An acoustic and articulatory investigation of the NURSE vowel in South Wales English. *Journal of the International Phonetic Association*, 40, 93–112. doi:10.1017/S0025100309990272.
- Mielke, J., Baker, A., & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language*, 92, 101–140. doi:10.1353/lan.2016.0019.
- O'Connor, J. D. (1967). *Better English Pronunciation*. Cambridge, UK: Cambridge University Press.
- Ong, D., & Stone, M. (1998). Three-dimensional vocal tract shapes in /r/ and /l/: A study of MRI, ultrasound, electropalatography, and acoustics. *Phonoscope*, 1, 1–13.
- Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study. *The Journal of the Acoustical Society of America*, 93, 2948–2961. doi:10.1121/1.405814.
- Proctor, M., Walker, R., Smith, C., Szalay, T., Goldstein, L., & Narayanan, S. (2019). Articulatory characterization of English liquid-final rimes. *Journal of Phonetics*, 77, 100921. doi:10.1016/j.wocn.2019.100921.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. URL: <https://www.R-project.org/>.
- Roach, P. (1983). *English Phonetics and Phonology: A Practical Course*. (2nd ed.). Cambridge, UK: Cambridge University Press.
- Satterthwaite, F. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, 2, 110–114. doi:10.2307/3002019.
- Scobbie, J. M. (2006). (R) as a variable. In K. Brown (Ed.), *The Encyclopaedia of Language and Linguistics*. (pp. 337–344). Oxford, UK: Elsevier volume 10. (2nd ed.).
- Scobbie, J. M., Lawson, E., Cowen, S., Cleland, J., & Wrench, A. A. (2011). A common co-ordinate system for mid-sagittal articulatory measurement. *CASL Research Centre Working Paper: WP-20*, (pp. 1–6).
- Scobbie, J. M., Lawson, E., Nakai, S., Cleland, J., & Stuart-Smith, J. (2015). Onset vs. coda asymmetry in the articulation of English /r/. In The Scottish Consortium for ICPHS 2015 (Ed.), *Proceedings of the 18th International Congress of*

- Phonetic Sciences*. Glasgow, UK: The University of Glasgow. URL: <http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPhS0704.pdf>.
- Scobbie, J. M., Punnoose, R., & Khattab, G. (2013). Articulating five liquids: A single speaker ultrasound study of Malayalam. In L. Spreafico, & A. Vietti (Eds.), *Rhotics: New Data and Perspectives* (pp. 99–124). Bozen-Bolzano: BU Press.
- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2015). afex: Analysis of Factorial Experiments. *R package version 0.13–145*, . URL: <http://CRAN.R-project.org/package=afex>.
- Smith, B. J., Mielke, J., Magloughlin, L., & Wilbanks, E. (2019). Sound change and coarticulatory variability involving English /ɹ/. *Glossa: a journal of general linguistics*, 4(1), 1–51. doi:10.5334/gjgl.650.
- Stevens, K. N. (1998). *Acoustic Phonetics* volume 30. Cambridge, MA: MIT press.
- Sweet, H. (1877). *A Handbook of Phonetics* volume 2. Clarendon Press.
- Tiede, M. K., Boyce, S. E., Espy-Wilson, C. Y., & Gracco, V. L. (2010). Variability of North American English /r/ production in response to palatal perturbation. In *Speech Motor Control: New developments in basic and applied research* (pp. 53–68). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199235797.003.0004.
- Tiede, M. K., Boyce, S. E., Holland, C. K., & Choe, K. A. (2004). A new taxonomy of American English /r/ using MRI and ultrasound. *The Journal of the Acoustical Society of America*, 115, 2633–2634. doi:10.1121/1.4784878.
- Traunmüller, H., & Öhrström, N. (2007). Audiovisual perception of openness and lip rounding in front vowels. *Journal of Phonetics*, 35, 244–258. doi:10.1016/j.wocn.2006.03.002.
- Trudgill, P. (1999). *The Dialects of England*. Blackwell.
- Twist, A., Baker, A., Mielke, J., & Archangeli, D. (2007). Are "Covert" /ɹ/ Allophones Really Indistinguishable? *University of Pennsylvania Working Papers in Linguistics*, 13, 16. URL: <https://repository.upenn.edu/pwpl/vol13/iss2/16/>.
- Uldall, E. (1958). American 'molar' R and 'flapped' T. *Revista do Laboratório de Fonética Experimental da Faculdade de Letras da Universidade de Coimbra*, 4, 103–106.
- Underhill, A. (1994). *Sound Foundations: Living Phonology* volume The teacher development series. Oxford: Heinemann.
- Vaissière, J. (2007). Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages. In M. Solé, P. Beddor, & M. Ohala (Eds.), *Experimental Approaches to Phonology* (pp. 54–71). Oxford, UK: Oxford University Press.
- Wells, J. (1982). *Accents of English* volume 1. Cambridge, UK: Cambridge University Press.
- Westbury, J. R., Hashi, M., & Lindstrom, M. J. (1998). Differences among speakers in lingual articulation for American English /ɹ/. *Speech Communication*, 26, 203–226. doi:10.1016/S0167-6393(98)00058-2.
- Wilson, I. L. (2006). *Articulatory Settings of French and English Monolingual and Bilingual Speakers*. PhD Thesis University of British Columbia, Vancouver, Canada.
- Wrench, A. A., & Scobbie, J. M. (2016). Queen Margaret University ultrasound, audio and video multichannel recording facility (2008-2016). *CASL Research Centre Working Paper: WP-24*, (pp. 1–14).
- Zawadzki, P. A., & Kuehn, D. P. (1980). A cineradiographic study of static and dynamic aspects of American English /r/. *Phonetica*, 37, 253–266. doi:10.1159/000259995.
- Zhang, Z., Boyce, S., Espy-Wilson, C., & Tiede, M. (2003). Acoustic strategies for production of American English 'retroflex' /r/. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1125–1128). Barcelona, Spain: Universitat Autònoma. URL: [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15\\_1125.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_1125.pdf).
- Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Holland, C., & Choe, A. (2008). A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/. *The Journal of the Acoustical Society of America*, 123, 4466–4481. doi:10.1121/1.2902168.