Not quite the same: The social stratification and phonetic conditioning of the FOOT-STRUT vowels in Manchester[1]

DANIELLE TURTON

Lancaster University

MACIEJ BARANOWSKI

University of Manchester

ABSTRACT

The FOOT-STRUT split, which has its origins in 17th century English, is notable for never having occurred in the speech of Northerners in England; thus, *stood~stud* are still homophones. The present study analyses 122 speakers from Manchester in the North-West of England. Although the vast majority of speakers have no distinction between the two vowels in minimal-pair production and judgement tests, vowel height is correlated with socio-economic status: the higher the social class, the higher the F1 of STRUT. Surprisingly, in statistical models the predictor of vowel class remains significant. This means that, for a speech community without the split, there remains an effect in the expected direction: STRUT vowels are lower than FOOT vowels in the vowel space (i.e. they have a higher F1). We suggest that co-articulatory effects of surrounding consonants explain this acoustic difference, as they have significant lowering/heightening effects on F1 but are not fully captured by our statistical model. We argue that the perplexing nature of the historical split can be partially accounted for in this data, as the frequency of co-occurring phonetic environments is notably different for FOOT than in STRUT, resulting in cumulative effects of

co-articulation. We also present evidence of age-grading which suggests that middle class speakers may develop a phonetic distinction as they age.

1. INTRODUCTION

The FOOT-STRUT split is one of the most prominent sound changes dividing the linguistic systems of the North and South of England. Estimated to have taken place during the 17th century (Wells 1982: 197, Lass 2000: 89), the historical split never occurred in the North of England, meaning that speakers today have homophones in the following words:
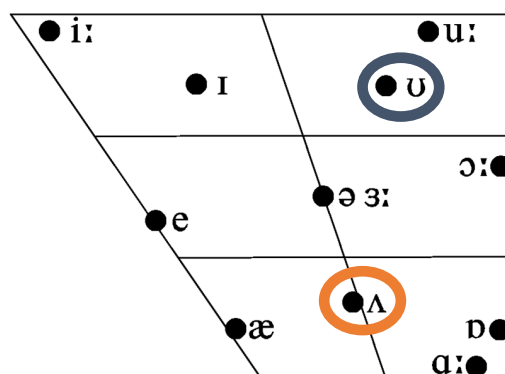
put ~ putt

stood ~ stud

could ~ cud

look ~ luck

In these varieties, FOOT is represented by the high back lax vowel /ʊ/, and there is a five-vowel system of short vowels, excluding /ʌ/ (see Figure 1 for the monophthongs of Received

Pronunciation as *Figure 1: The representation of the FOOT-STRUT vowels in the vowel space for Received Pronunciation*

traditionally

described).[2] As far as we know, the only other areas which report /ʊ/ for both FOOT and STRUT are some parts of Ireland, such as the local variety spoken in Dublin (Wells 1982; Hickey 1999; Lonergan 2015) and the areas of North-East Wales which border with the North-West of

England, such as Wrexham (Morris 2013). It is important to note that the fourth pair listed above, *look ~ luck,* is not a relevant FOOT-STRUT minimal pair for all Northerners. The pronunciation of the *–ook* words such as *book, cook* with the GOOSE vowel [u:] is found in areas such as Liverpool, Newcastle and Stoke-on-Trent, but is restricted to the speech of the oldest generations in Manchester (Baranowski and Turton 2015), the community under investigation here.

The present study aims to explore the social stratification of the FOOT-STRUT vowel(s) in the accent spoken in Manchester in the North-West of England. A secondary aim is to test the phonemic status of this vowel, and whether all speakers truly do show a complete lack of split in both their production and in judgement elicitation tasks. Overall, the paper seeks to contribute to our theoretical understanding of vowel categories and distinctions, as well as addressing some of the methodological considerations relevant to measuring distinctions. From a diachronic perspective, this investigation seeks to utilise present-day data and synchronic patterns to illuminate potential historical pathways of diachronic, affording researchers a window on how splits originate. Moreover, we seek to address how constraints on the phonological grammar, such as the creation of a new phonemic category, places limits on the range afforded to variation and sociolinguistic change.

2. THE FOOT-STRUT VOWEL(S) IN ENGLISH

2.1 The history of the split

The early history of the FOOT-STRUT split is unclear, with initial reports dating back to the late 16th century, but more solid evidence appearing in the 17th century (Lass 2000: 89). The split stemmed from Middle English short /u/ unrounding in most contexts to something approximating [ɤ], with [ʊ] being retained after labials. Minkova (2014: 245) points out that the phonetic rationale for this is clear: the labiality feature is kept in the transition from onset to nucleus. This would seemingly result in an allophonic distinction, and Wells (1982: 197)

suggests this is potentially how the split began. However, the situation becomes much more complicated when we consider the intermediate stages resulting in the present-day realisations with minimal pairs, thus demonstrating that an allophonic distribution is not the case in present-day English. Not only are there clear exceptions to the above rule (today *fun, mud* have [ʌ], *cushion* has [ʊ][3]), these exceptions spread with the shift of Middle English /oː/ to /u/ during the Great Vowel Shift. The *–oo-* set contains many examples of [ʊ] after non-labials e.g. *looking good*. Some words in the /oː/ set shortened early on, and thus joined the /u/ category in the Great Vowel Shift, in time for the split. These tend to have /ʌ/ in Present Day English e.g. *blood, gums*. Those lexical items which shortened later tend to have /ʊ/, but these words varied regionally[4], which is why many Northerners today have /uː/ for *–ook* words (Wells 1982: 197-8).

The trajectory to the rather messy present-day situation is outlined in Figure 2 adapted from Wells (left panel; 1982: 198) and Lass (right panel; 2000: 89). Their interpretations of the trajectory are slightly different from one another. Lass argues that evidence from 17th century reports demonstrates that Middle English /u/ started to shift while it was still [u], that is, before it centralised and lowered to [ʊ], although it is probable that this tense/lax distinction was already phonetically variable at this point. Wells includes more phonetic gradience in the change. However, what is clear from the available records, such as the word lists in Ellis (1875: 1001-1017), is that the evidence of the split from 17th century orthoepists is variable in nature and moreover is based predominantly on reports from well-off speakers based in the London area. Thus, it is important to emphasise that the overall picture that we do have is based on limited evidence. Nevertheless, what *is* clear is that the split is phonologically and phonetically complex and that during the historical shift to the present-day pattern, vowels shifted as part of a broader mechanism, with lexical items flitting from one category to another. This, alongside gradient phonetic change and regional variability, gives us the system we have today.
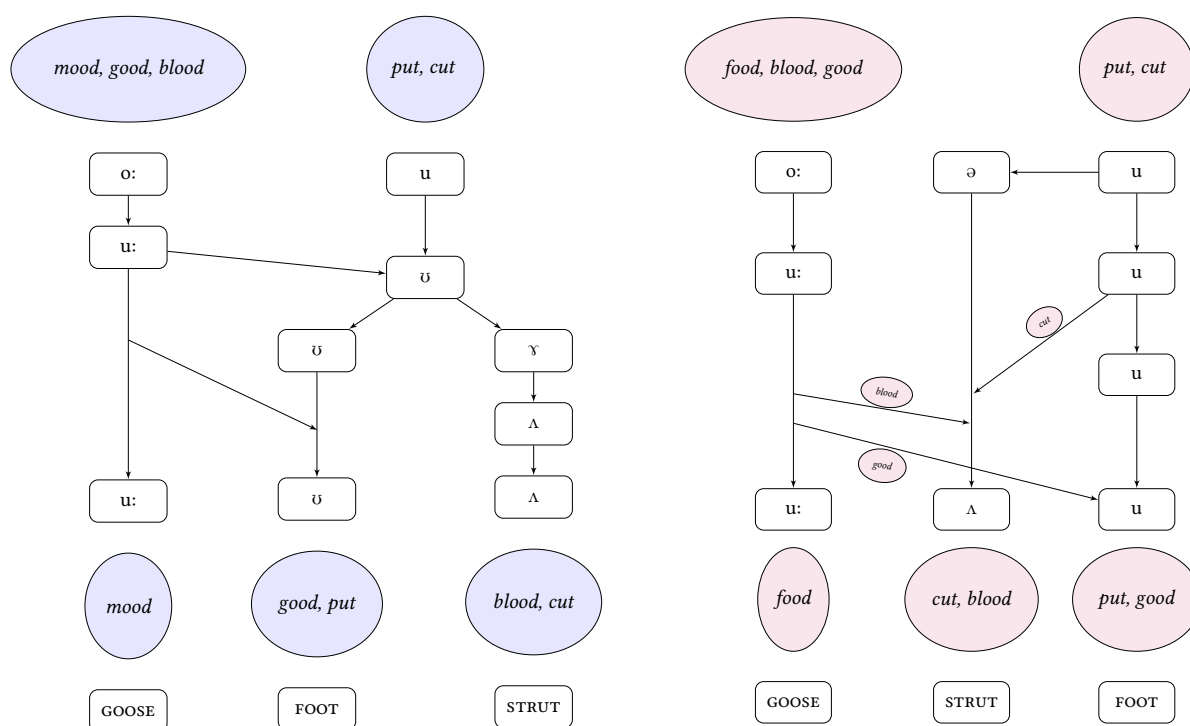
Figure 2: Historical trajectory of split adapted from Wells (1982) on the left and from Lass (2000) on the right

From the perspective of Northern authors, Beal (2012) reports no evidence of the split, and quotes 18th century sources which confirm that it is not found in the North. Although the split never did occur in the North of England, understanding its diachronic trajectory may give us insight into why some of these changes occurred in the first place in the South. Of course, the historical linguistic context is completely different today, and it is unlikely that if the FOOT-STRUT split was indeed spreading northwards in 21st century England, it would be through the same mechanism. However, some understanding of the potential phonetic and lexical pathways of change may help us predict whether this split is likely to continue spreading (albeit very slowly) and the mechanisms by which this might happen.

2.2 Present Day English

Today, regional accents in the North of England[5] retain the lack of distinction. In a large-scale online survey, MacKenzie, Bailey & Turton (2016, in prep) found that 82% of speakers north of the Midlands claim that *foot* and *cut* rhyme, whilst Midlands speakers were at 52%. Maguire's (2009) survey reports similar figures (72% non-distinct in the North, which includes the Midlands). A map of the overall results can be found in Figure 3 and demonstrates that the isophone has moved further North compared to its position in the *Linguistic Atlas of England* (Orton et al. 1978). Evidence from Britain et al.'s (2016) English Dialects App also suggests that the split may be spreading, with the boundary being further North and the percentage of speakers with the split increasing all over the country (see also Leemann et al. 2018). Note that
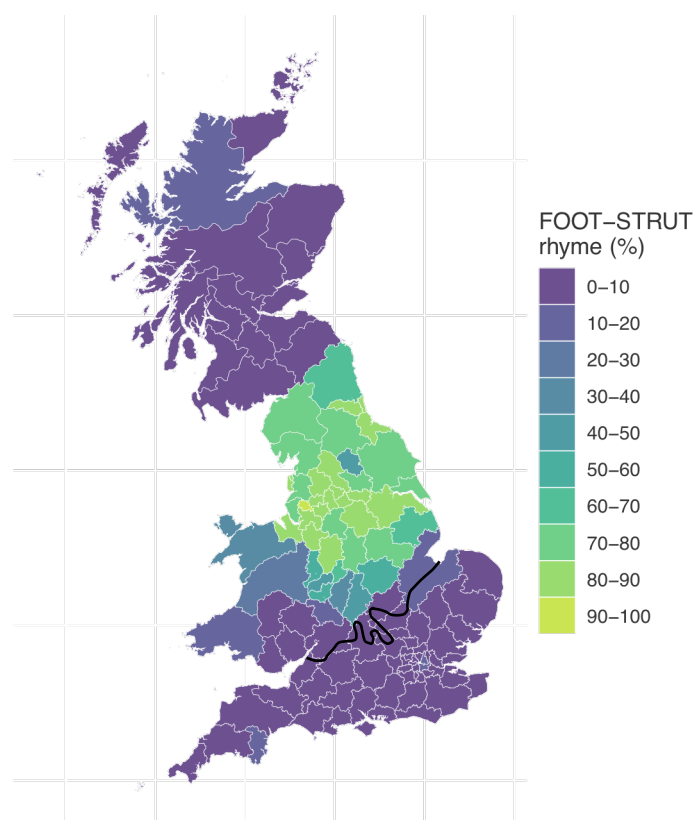
these two recent surveys inherently elicit judgements from younger speakers, who will be much more likely to use mobile phone apps or take online surveys. In contrast, the data from the *Linguistic Atlas of England* mapped 1950s data from the *Survey of English Dialects* (SED), which purposely targeted the opposite demographic: Non-mobile Older Rural Males (NORMs). Thus, this change may not be as drastic as it first seems.

Figure 3: Map of FOOT-STRUT distinction from MacKenzie et al. (2016, forthcoming) with superimposed major isophone from the Linguistic Atlas of England (Orton et al. 1978) based on 1950s SED data.

If the distinction is indeed spreading today, the most likely scenario is that this is happening through lexical diffusion i.e. moving through the lexicon gradually, but abruptly on the phonetic dimension. That is, speakers in the Midlands area are flipping appropriate words one by one into the STRUT lexical set. Trudgill (1986: 59) suggests just this: that a FOOT-STRUT distinction is spreading northwards, but only on a word-by-word basis. Reports from Wells (1982) would also seem to support this view, as he cites Midlands speakers as being variable, with words such as *rubber* having [ʊ], but *butter* having [ʌ] (Heath 1980, cited in Wells 1982: 352). As the split is, synchronically, highly unpredictable by phonological conditioning, lexical diffusion is the only mechanism by which it could spread.

In terms of the phonetics and phonemic distribution of varieties which do not have the true split, very little quantitative empirical work exists. Britain (2015) describes the phonetic contexts in the Fens which seem to favour a more [ʊ]-like variant vs. a more [ʌ]-like realisation, as shown in Table 1 (see also Britain 2013, 2014 for a detailed overview of FOOT-STRUT in the Fens).

More [ʊ]-like variants:	More [ʌ]-like variants:
After /w, p, b/ or generally nearby rounded consonants (*wonder, pub, bun*)	before and after /v/ (*love, vulture*)
Word initially (*under*)	after /θ/ (*thunder*)
before stops (especially /g, p, t, n/) (*hug, up, cut, run*)	before /ð/ and /l/ (*other, dull*)
before /ʃ/ (*rush*)	

Table 1: Britain on the phonetic tendencies of FOOT-STRUT variation in the Fens (2015: 410)

Chambers and Trudgill (1998:106-112) separate varieties which show intermediate forms in terms of variance and gradience into 'mixed lects' and 'fudged lects'. Mixed lects are varieties which sometimes use [ʊ] and sometimes use [ʌ], somewhat interchangeably (although undoubtedly there would be predictive factors). They map data from East Anglia and the East Midlands area showing the variability of this vowel in SED informants, showing that in some areas both variants are acceptable for the same lexical item (see also Upton 1995). Fudged lects are varieties which have a phonetic realisation somewhere between [ʊ] and [ʌ]. There is plenty of evidence in the descriptive dialectology literature that many speakers of English without a phonemic distinction have these fudged variants, what Hughes, Trudgill & Watt (2005: 60) call a 'phonetic compromise' between [ʊ] and [ʌ], approximating schwa. As Beal (2010: 13) explains, it is often middle-class speakers of Northern English who may produce a compromised [ə]-like form. However, Beal is clear that presence of this lowered vowel is unlikely to indicate a phonemic contrast between /ʊ/ and /ʌ/ but rather one phoneme with an intermediate phonetic realisation.

One could infer from these descriptions that fudged variants are the result of speakers aiming for an intermediate phonetic target somewhere between the two distinct forms that they have heard in the wider (supra-local) community. Whilst this is entirely possible, the other interpretation is that in some locations, fudged forms date back to the original shift of [ʊ] to [ʌ]. Ellis (1889: 17) describes 'the intermediate country' (which reflects areas between the northernmost boundary of [ʌ] and the southernmost boundary of [ʊ]) as being mixed or transitional. Available dialect surveys from the mid-20th century may also support this. Although the SED only reports fudged variants in Norfolk, the American fieldworkers in *The Lowman Survey* (Kurath & Lowman 1970:17) map such realisations from Norfolk to Somerset and Dorset (transcribed as [ʌ̟ ~ʌˆ]). It is possible that American ears are more sensitive to the variation in British STRUT (which traditionally is lower than in American English; although see

Lindsey 2019), and interestingly it can be noted that the SED fieldworker for Norfolk was one of only two American fieldworkers on the project, W. Francis Nelson (Orton 1962: 16). Thus, it may be the case that these fudged forms likely represent fossilised transitional forms which have always been the norm in some areas of England but not necessarily noted by British ears. This story is probably not the case, however, for fudged variants in the North which, as Beal (2010) indicates, are associated with middle-class speech. Such variants are reported for the middle classes in Hull (Williams and Kerswill, 1999:146), middle-class females in Newcastle (Watt and Milroy, 1999: 28) and in self-conscious speech in Derby (Docherty & Foulkes 1999: 28), and are more likely to have arisen from speakers being aware of their own local [ʊ]-like variant in contrast with a more prestigious Southern [ʌ]-like variant. Such fudged forms are certainly not typical of Manchester speech, but as this study focusses on vowel height (the F1 dimension acoustically), further analysis of F2 and vowel centralisation would be required to investigate the true picture of potential middle-class fudging in the city.

There is little in the way of instrumental phonetic study of the FOOT-STRUT vowels, and almost no published work from a variationist perspective. Ferragne & Pellegrino's (2010: 14-20) analysis of numerous varieties of English based on a single speaker report a complete lack of split for East Yorkshire English, Lancashire, Liverpool and Newcastle in their *hood~hud* pairs. More recently, Strycharczuk et al. (2019) analysed reading passage data from the English Dialects App Corpus representing 141 speakers from seven cities across the North of England, including Manchester. They found that around a quarter of their Northern informants had some kind of split. This effect was more typical of female speakers, and those who had moved around more i.e. more mobile informants. Just four of the 25 Mancunians in their corpus were found to exhibit some kind of split, below the average for Northerners in this sample.

Intriguingly, Flynn (2012) and Braber & Flynn (2015), who present one of the only sociolinguistically-informed acoustic phonetic investigations of the FOOT-STRUT vowels in a Northern variety (Nottingham in the East Midlands), report that although the vowels sound the same, there is a small instrumental difference. Braber & Flynn (2015:370) point out that this unusual result warrants further research. This is something the present investigation will be able to address.

If the vowels truly are non-distinct, we may expect instances of hypercorrect FOOT to something approximating [ʌ] from speakers who interpret a lower vowel as what Chambers & Trudgill refer to as a more 'statusful RP form' (1998:42). Wells (1982: 353) cites *sugar, butcher, cushion* as common words in which to find speakers 'poshing-up' and producing hypercorrect FOOT[6]. Sankoff (2004) reports that her own grandmother, a Northern English speaker (from the Greater Manchester area) living in Canada, had variable hypercorrection in many FOOT words such as *put* and *book*, a strategy seemingly implemented to avoid the 'hated' [ʊ] sound. The motivation for mobile speakers to change later in life is discussed in the next section in more detail. Here in Manchester, the authors have also heard speakers hypercorrect *foot, book,* often to something approximating LOT rather than an RP-like form. Almost always in our experience, this is heard in teachers or university lecturers, and when great attention is being paid to speech. Thus, the sociolinguistic interview format may not be the best place to find such forms.

Labov (1994: 347) states that the importance of Received Pronunciation (henceforth RP) as a reference accent will mean that some individuals from the North will have succeeded in striving to acquire a true distinction. The results from more recent dialect surveys support this (MacKenzie et al. 2016, in prep; Britain et al. 2016; Strycharczuk et al. 2019). Labov does note, however, that these speakers are in the minority and begin learning the rules in their pre-adolescent years. This may be at boarding school or perhaps from southern parents, although

this southern influence would likely only survive in middle-class areas or in the case of determined individuals. Wyld's (1936: 3-4) descriptions support this, as he asserts that speakers acquire the distinction through 'opportunity and experience', signifying that some individuals make a concerted effort to achieve an RP STRUT. His claims that RP is a regionless dialect spoken in the public[7] school system would seem to hold for the FOOT-STRUT vowels in the North. See Halfacre (2019) for instrumental evidence of a true distinction in privately educated Northern speakers.

2.3 Plasticity, change across the lifespan and age grading

If we indeed find speakers using hypercorrect FOOT, and individuals acquiring a split through education, it raises the question as to whether Northerners can acquire a split later in life. If not a phonemic one, perhaps they can recreate an approximation of the phonetics. Sankoff's (2004) paper analysing the speech of Nick and Neil from Michael Apted's *Up* documentary series attempts to do just this. Nick and Neil, both from the North of England, show a large increase in the use of [ʌ] realisations as they age. Although Sankoff demonstrates that Nick and Neil have changed their phonetics, she is cautious in stating whether or not they have achieved a phonemic split. Sankoff (2004:18) describes a scene in one episode where Nick, as an adult, slips back into his Northern [ʊ] forms when he becomes emotional whilst describing how he came to learn that his baby brother was deaf. Following Labov (1972) and the 'danger of death' question[8], Sankoff notes how states of high emotion are the best for eliciting the true vernacular and would possibly help reverse the effect of formality that the interviewees may have felt during the filming of the documentary series. This would suggest that Nick has not truly acquired the phonemic split like a native speaker, but very successfully learnt which words contain which phonetic realisation. Sankoff states that there is clear evidence of age

grading in Nick and Neil's speech, whilst pointing out that the two men are exceptional individuals who have lived unusual lives and moved around a great deal, as opposed to the majority of speakers who tend to stay put and do not change past adolescence.

Labov (1994: 348) notes how the role of the mass media should not be dismissed in advance for individuals who learn to acquire a Southern-like STRUT vowel as they age. Although studies of sound change have shown little evidence of the media producing a community shift, Labov states that the media could serve as a template for the FOOT-STRUT split in highly-motivated individuals who may be isolated within their Northern peer-group and aspire to speak the perceive educated standard. This may be further motivated by mobility, as Strycharczuk et al.'s (2019) speakers demonstrate, as well as various studies analysing the speech of Northerners living in the South of England, which we turn to now.

Evans & Iverson (2004, 2007) present a series of perception and production experiments looking at Northerners attending university in London. In their production data, they found that these Northerners would often produce the fudged variant of FOOT-STRUT, which the authors put down to speakers' motivation to lose their regional accent in order to fit in at university. In a related vowel judgement task (Evans & Iverson 2004), the authors compared Northerners living in London with those still living in the North.[9] Northerners living in London were highly capable of changing their acceptability of pronunciation if they were expecting a Southern voice as opposed to a Northern voice. That is, they would accept higher F1s as acceptable in STRUT words when listening to a Southern voice. Northerners remaining in the North of England did not do this. These results could initially point towards the university students acquiring a new phonetic form, which may after time result in a phonemic split, even though in production speakers opt for the fudged form. However, recent results from Chiu & Evans (2018) demonstrate that, indeed, if the phonemic contrast is not learnt early in life, it can never be acquired perfectly.

In a series of production and perception tasks, Chiu & Evans (2018) compare Northern Freshers (i.e. students just arriving in London for university), with Northern students who have been in London for two years, and with Southern students. Using EEG, a 3-way alternative forced choice discrimination task, and a vowel production task, the authors find some important results for studies of lifespan change. They find that Northerners who have been in London for two years perform, seemingly, as well as Southerners on the production of FOOT-STRUT distinction. However, on the EEG data, even though long-term Northerners are very good at detecting vowel changes (much better than the Freshers), they show considerably different Event Related Potential responses when compared to Southerners, who have acquired this phonemic distinction from birth. Chiu & Evans conclude that whilst Northerners living in the South of England can change their production very successfully given time in the area, these changes are not fully complete.

Kerswill (1996: 200) proposes a hierarchy of difficulty for language change across the lifespan, whereby new lexical items are the least difficult to learn (and can be done throughout life), regular changes, such as Neogrammarian shifts and mergers are somewhere in the middle and seem possible in adulthood, and new phonemic distinctions have to be acquired by adolescence. The fact we are able to learn new words later in life is non-controversial, but a growing body of work also supports the idea that regular sound change may be possible in adulthood. Harrington's (2006) famous study of the Queen's *happ*Y-tensing demonstrated this. Similarly, Conn and Horesh (2002) show that the fronting of the GOAT vowel may be acquired by adults moving to Philadelphia, a GOAT-fronting dialect, from Detroit, a non-fronting region. The ability to acquire regular sound changes in adulthood is also demonstrated in more recent work by Bowie (2019). Thus regular shifts that are phonetically gradual and lexically abrupt certainly seem possible post-adolescence for some speakers. This seems to hold for categorical variable rules: as long as the speaker has variability in the first place, they can shift their rates

of frequency as they age (see MacKenzie 2017 on phonological variation and Raumolin-Brunberg 2005 on morphological variation). Recent work by Hartshorne, Tenenbaum & Pinker (2018) demonstrates that for certain syntactic phenomena, the window may close at "the crux of adulthood" i.e. 17.4 years. They note that this critical period may differ for the phonology and, perhaps, in some individuals.

Indeed, new phonological oppositions seem to be amongst the most difficult and thus less likely to be acquired post-adolescence. Like Labov, Kerswill (1996) suggests that phonological oppositions such as FOOT-STRUT would have to be acquired by 13 years of age in order for speakers to truly demonstrate the extra phonemic category. This is one of many articles in the second dialect literature arguing for a critical period for more difficult forms like FOOT-STRUT. Chambers's (1992) seminal study of Canadian children moving to the South of England also gives us a window into the hierarchy of difficulty of different changes. Chambers demonstrates how more simple changes, which do not require the learning of a new phonemic category and are orthographically distinct (e.g. the loss of t-voicing), are acquired much more successfully than those which involve a new phonemic distinction (e.g. the LOT-THOUGHT distinction) or are lexically unpredictable (e.g. the TRAP-BATH distinction.) In Kerswill's hierarchy, the learning of a new phonemic distinction is surpassed only by lexically irregular phonological rules such as the Philadelphia æ-tensing pattern whereby /æ/ tenses in a highly unpredictable set of environments (Labov 1994). Payne's (1980) study of a Philadelphia suburb demonstrated that the pattern is so difficult to learn that only children with both parents from the area did so successfully, and Sneller's (2018) recent work in Philadelphia has shown that many speakers in the community are moving towards the simpler nasal pattern of æ-tensing. From the perspective of the phonological grammar, if this evidence suggests anything, it is that the odds are stacked against speakers in acquiring a FOOT-STRUT distinction later in life. How this will play out for some members of the speech community where the social

pressures are stacked in favour of a split is something which we can seek to investigate to some extent here.

2.4 The unmerging of the FOOT-STRUT vowel in Northern English varieties?

The existing literature suggests that the FOOT-STRUT vowel(s) are conditioned by contextual phonetic/phonological factors (such as neighbouring segments) and social factors (such as socio-economic class and mobility), whilst posing interesting questions for plasticity and the possibility of speakers acquiring new phonemic contrasts. Moreover, if the evidence from new dialect map data is to be believed, social pressures of the RP-like distinction may eventually outweigh the structural pressures against creating a new phonemic category. This goes some way to explaining why the spread of FOOT-STRUT distinction goes against Herzog's corollary to Garde's Principle: that mergers expand geographically at the expense of distinctions (Herzog 1965; Labov 1994, 2007). Indeed, Labov himself (1994:342-343) asserts that it is an error to assume that mergers can never be reversed, and outlines the social conditions that would have to be present for such a rare development to occur; the FOOT-STRUT pair is not a merger, of course, but we can apply the same principles to some extent. As Eckert & Labov (2017) note more recently, the reversal of mergers is an exception to the rule and they cite just one example of such a process: Baranowski's (2007: 236) report of NEAR-SQUARE reversal in Charleston, South Carolina. Charleston is unusual in that many aspects of its phonology are led by a prominent upper class, demonstrating that the social pressures have to be particularly strong to exert such a force on the underlying system. Labov (1994: 342-343) states that for Herzog's Principle to be tested, there must be an overt campaign to 'bestow prestige on the distinction' but also that, as a general rule, 'mergers and splits have no social affect associated with them' (see also Baranowski 2013 for a further discussion of the role of social factors in mergers). In

terms of FOOT-STRUT, we agree with Labov on this note: it is not the presence of two separate phonemic categories that upwardly mobile speakers see as prestigious, rather it is a lower vowel articulation generally. However, it is important to underline that we do not expect our particular dataset to illuminate the debate of whether or not the FOOT-STRUT distinction is spreading North, simply because Manchester is not in close enough proximity to the FOOT-STRUT border. Although we may find aforementioned 'motivated individuals' in Manchester who are attempting to become split (perhaps like Chiu & Evans 2018 and Evans & Iverson's 2007 Northern Freshers), the issue of diffusion more widely in the British Isles is really best studied by researchers working far closer to the FOOT-STRUT border (e.g. Evans & Iverson 2004; Braber & Flynn 2015). These are the regions more likely to have mixing of speakers with and without the distinction, and it is this face-to-face contact which would cause wider community spread. The question of how and whether this progression will affect an area much further north, such as Manchester, is something this paper seeks to address, which leads to our research questions:

i. What is the social stratification of the FOOT-STRUT vowel in Manchester?

    Regardless of whether speakers have the split, is the STRUT vowel lower (i.e. higher in F1) the higher the social class?

ii. What is the phonemic status of FOOT-STRUT in Manchester as a Northern dialect? Is there a complete lack of split?

3. METHODOLOGY

This study forms part of a larger project into variation and change in Manchester. We define the speech community primarily as the area within the M60 ring road, including the

neighbourhoods of Sale, Wythenshawe, and Stockport just south of the M60 (which form part of the same uninterrupted urbanised area). Crucially, we do not include the entire urban conurbation of Greater Manchester, which includes satellite towns such as Rochdale, Oldham, or Bolton. The linguistic systems of these areas are quite distinct from Manchester itself (Baranowski & Turton, 2015) and warrant separate sociolinguistic study.

Data for this study is based on the acoustic analysis of the speech of 122 informants, stratified by age, gender, social class, and ethnicity (see Table 2). Of those informants, 91 speakers identify themselves as White British; the other 26 represent the two largest ethnic minorities in Manchester— Pakistani (18 speakers) and Black Caribbean (13 speakers). The informants all grew up in Manchester and, in the case of the White speakers (with the exception of the upper-middle class), at least one of their parents is native to the Manchester area as well.

Speakers are divided into five occupational levels for statistical models and graphs (1-5: lower working, upper working, lower-middle, middle-middle, upper-middle) using traditional methods of occupation[10], but this is simplified into working vs. middle class in Table 2. For a detailed debate on whether occupation is the best way to operationalise social class, see Baranowski & Turton (2018). Although social class is occasionally visualised as a continuous variable in the figures, it is always treated as a categorical variable for the purposes of the regression.

Sociolinguistic interviews were conducted, focussing around the topic of growing up in Manchester in order to elicit narratives of personal experience (Labov, 1984). This spontaneous speech was supplemented with word list reading and minimal-pair tests for a range of vocalic and consonantal contrasts. Minimal-pair tests for the words *book~buck* and *crux~crooks* were collected for 112 of the speakers by asking speakers to read each pair and state whether they thought the words were the same or different. The production of this was

judged by two local students. There were five speakers who pronounced the –*ook* words such as *book, look, cook* with the GOOSE vowel instead of the FOOT vowel. They are not included in the minimal pair analysis here[11], and have had these words excluded from the dataset for casual speech.

The interviews were recorded on Sony PCM-M10 recorders with Audio-Technica ATR3350 lavaliere microphones, at a sampling rate of 44.1 kHz at 24 bits. Speakers' complete vowel systems were measured in Praat in terms of F1 and F2. For 25 speakers the point of measurement was selected by hand prior to the advent of forced-alignment tools, with the remaining 97 speakers being processed through the Forced-Alignment Vowel Extraction (FAVE) Suite (Rosenfelder et al. 2014). Recent research into these forced-alignment tools indicates that there is no significant difference between FAVE placed boundaries and human annotators (MacKenzie & Turton 2019). FAVE codes the FOOT vowel as UH1 and the STRUT vowel as AH1 and filters out unstressed tokens.

	working class		middle class	
	female	*male*	*female*	*male*
young **[8 to 30]**	13	15	17	18
middle **[31-55]**	6	9	13	7
old **[56+]**	8	7	6	4

Table 2: Social stratification of speakers. Note that in plots and statistical models, class is operationalised as five factors and age is a numeric variable

This process resulted in a total of 11,492 tokens. Formant values were normalised using

Lobanov's (1971) method and scaled back to Hz. Tokens elicited by word list and minimal

pairs were removed from the dataset for the statistical models in order to focus on spontaneous

Predictor	**Factor levels/details**
age	continuous, centred around mean
sex	*female*
	male
social class	*lower working class (1)*
	upper working class (2)
	lower middle class (3)
	middle middle class (4)
	upper middle class (5)
ethnicity	*Black*
	Pakistani
	White
vowel	STRUT
	FOOT
preceding segment	*voiced stops, /w/ and /ʃ/*
	other
following place	[N.S]
following manner	*obstruent* (apart from affricates)
	affricate
	sonorant
following voice	voiceless

	voiced
syllable	open syllable
	closed syllable
	complex coda monosyllable
	complex coda + more syllables
word frequency	continuous, SUBTLEX, Zipf-scaled and centred around mean
duration	continuous, log transformed
interaction: age*sex*class	
random slope: speaker and vowel	

Table 3: All considered predictors and their final factor levels. Baseline levels are italicised.

speech (8,831 tokens). Eight speakers with the true split (confirmed in minimal-pair tests; see Section 4.1 and 4.2) were also removed from the final model resulting in a final number of 8,237 tokens. Word frequency is operationalised as a centred Zipf-scaled frequency from the SUBTLEX-UK corpus (van Heuven et al. 2014).

We used mixed-effects linear models in R (lme4 package; Bates et al. 2015) to model normalised F1, with a t-value of ±2 indicating a significant effect in the output tables. In model selection, we follow Barr et al. (2013) in testing our random effects structure, that is, starting off by including the maximal random effects structure as justified by the experimental design and reducing where appropriate. We decided on a random effects structure which included a by-speaker random slope for vowel category (i.e. FOOT vs. STRUT) to account for any individuals who may have been motivated to acquire a split independently, as discussed in the previous sections. Table 3 shows the predictors factored into model selection. Following place of articulation was found to have next to no effect on F1, and was removed from the final model after a likelihood ratio test confirmed its inclusion was not making a significant contribution to

the model. All continuous variables were centred around the mean. Duration was log transformed to avoid skew and to ensure the predictor conformed more closely to the normal distribution. Graphs are visualised using ggplot2 (Wickham 2009), and F1 values are summarised over speaker and word, to prevent speakers or words that have higher number of tokens than others from skewing the overall effect.

Note that the best model includes and interaction of age*sex*class, but we reserve the presentation of this model for Appendix A, presenting the model without the interaction in the main body of the paper. This is because a model with the interaction shows the coefficients for the baseline of the interaction predictors only (i.e., lower working class females of the mean age).[12] Thus, we deem the model without the interaction as a better overview of the variation in this dataset, but discuss the results of the interaction in Section 4.4.

We attempted various combinations in collapsing the preceding segment and following manner categories. We tried categorising with the in-built Plotnik (Labov 2011) categories which are outputted with FAVE-align, but this did not seem to be the best way of capturing the variation in F1. When we had settled on all other predictors of the model, we ran additional models with preceding sound specified individually, in order to observe trends. We used the coefficients of this model to categorise similar sounds. A clear trend emerged in terms of sounds which favoured a higher vowel (lower F1) for preceding segment, which were the voiced stops, as well as /w/ and /ʃ/ which have a high tongue body and (some) lip rounding. All other sounds have a fairly wide pattern of variation which, for the most part, also seem to act in tandem with the following segment to produce an effect on the option was to keep preceding category as unspecified as possible. We opted for two categories, separating the voiced stops, /w/ and /ʃ/ from everything else. Doing so resulted in a reduced Akaike information criterion (AIC) when compared to other possible combinations.[13] For following manner, there was a straightforward division between obstruents and sonorants, although

affricates were displaying some significantly different behaviour and thus were set as a third category from the rest of the obstruents. The final breakdown of the predictors and their factor levels can be found in Table 3. We plan to probe these phonetic conditions further in future work, but this may be best investigated with lab speech, as opposed to the sociolinguistic interview format.

4. RESULTS

4.1 Minimal-pair tests

Figure 4 shows the results of the minimal-pair tests conducted at the end of the sociolinguistic interview. There were two minimal-pair tests for the FOOT-STRUT vowels: the informants were asked to read the pairs *book-buck* and *crooks-crux,* and to confirm whether the words sounded the same or different to them. In the left panel of Figure 4 is the participant response to this question, i.e. whether the speaker themselves contended that the words were the same or different. The right panel shows whether the production was actually different. As explained in the methodology, speakers who have [uː] for *–ook* words, rather than [ʊ], have been removed. As Figure 4 shows, the vast majority of speakers report *book* and *buck* as being homophones in their judgement, which is largely matched by their production. The highest social classes are the exception to this, where these results would suggest that over half of respondents seem to have the split. There is one speaker in the upper-working-class group who reports these as being different, but produces them the same[14]. We revisit the production of *book-buck* in terms of the acoustics in Section 4.3.4.

The picture for the considerably less frequent minimal pair words *crooks* and *crux* is very similar, although interestingly we have one upper middle class speaker who reports them as being the same, even though she has *buck* and *book* as different in production and perception. This is an indicator of the complexity of the potential phonemic status of this vowel in Northern

Figure 4: Do 'book' and 'buck' sound the same to you? Results of the minimal-pair judgement tests by social class

Englishes. As discussed, the highest social classes may very well have some kind of phonemic distinction between the FOOT and STRUT vowels, acquired from private school, or in some cases from parents from the South. However, for very low frequency words, such as *crux*, speakers are likely to put this in the "wrong" lexical set, as they simply do not hear it enough to be able to correctly posit its representation. This result from just one speaker is enough to support Trudgill's (1986) claim that, if the FOOT-STRUT split is indeed moving northwards, it is doing it on a word by word basis. This is non-controversial evidence of an arguably rare example of the mechanisms behind lexical diffusion.

As expected, the vast majority of speakers report no FOOT-STRUT distinction and show no evidence for it in the minimal pair production. Those speakers who do seem to have some kind of distinction do not necessarily have all words in the same lexical set as speakers of RP (e.g. *crux* remains in FOOT for the upper middle class speaker mentioned above). Using the results from the minimal-pair tests, along with the criteria listed below in Section 4.2, we deduced that eight of our speaker set have a true split (all middle to upper middle class speakers).

4.2 FOOT-STRUT: general findings and models

Figure 5 shows a selection of working class speakers from the sociolinguistic interview data, who clearly have one phonemic category for the FOOT-STRUT vowels (FOOT words are shown in blue circles and STRUT words in red triangles). From the youngest to the oldest speakers, male and female, it is clear that no emerging split is in progress here. Contrast this with the upper middle class speakers in Figure 6. To reiterate, eight of our 122 speakers show evidence of having a phonemic split, with six being upper middle class. Speakers were classified as having the split if they a) reported a split in their judgement in the minimal pair test b) produced a split in the minimal pair test and c) showed a consistent split in their sociolinguistic interview. These speakers were removed from the linear mixed-effects regression in Table 4.

Figure 6 shows four upper middle class speakers. Mike T, Wendy J and Matthew P were categorised as split speakers using the criteria above. Wilma L was not categorised as a split speaker: she reports *book~buck* as sounding the same, produces them the same, and has enough phonetic overlap in her sociolinguistic interview tokens to indicate that a phonemic split on a level with RP is not present. It is notable, however, that her productions seem to be in the direction as expected with someone with the split: STRUT tokens (in red) are lower than FOOT tokens, in blue. However, the minimal pairs items *buck~book* are completely overlapping in phonetic space for Wilma L. We return to this point later in Section 4.3.4.
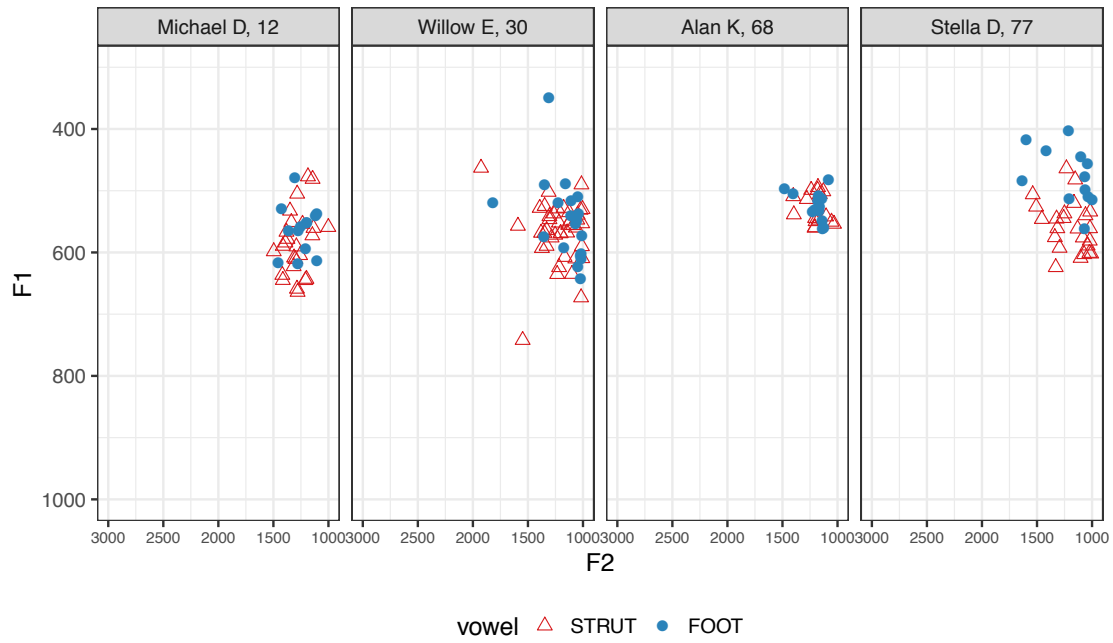
Figure 5: A selection of working class speakers, demonstrating no evidence of a split. STRUT words are triangles and FOOT words are circles.
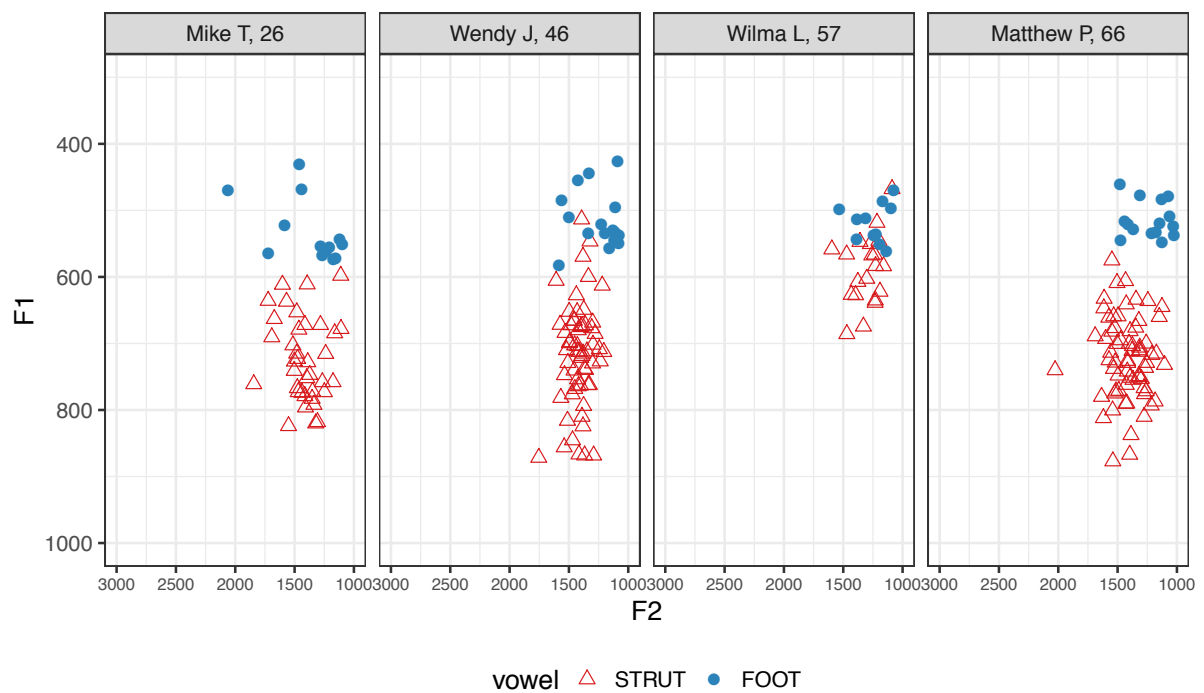

Figure 6: A selection of upper middle class speakers demonstrating that some are split and some are not. STRUT words are triangles and FOOT words are circles.

	estimate	std. error	t-value	N	mean F1
vowel class (baseline STRUT)				7442	588
FOOT	-28.6975	4.636	-6.19	4037	545
social class (baseline 1)				2020	557
2	1.699	5.641	0.301	1828	564
3	6.8122	5.5405	1.23	2365	569
4	5.2319	6.1017	0.857	1743	580
5	23.3936	13.7879	1.697	281	584
age (continuous)	-0.1561	0.1165	-1.34		
sex (baseline female)				4616	564
male	12.3675	3.9944	3.096	3621	572
ethnicity (baseline black)				997	559
pakistani	7.9339	7.0304	1.129	1378	580
white	-5.4057	5.7022	-0.948	5862	566
duration (log)	39.3824	1.8089	21.772		
preceding segment (baseline voiced stops, /ʃ/ and /w/)					
other	20.3222	3.4418	5.905		
following voice (baseline voiceless)				5312	577
voiced	-17.8416	3.4966	-5.103	6167	569
following manner (baseline obstruent)				8070	568
affricate	-24.501	7.6123	-3.219	534	593
sonorant	16.2214	3.6963	4.389	2875	582
syllable (baseline closed syllable)				6708	572
open syllable	-12.0491	3.5938	-3.353	2514	575
monosyllabic comp coda	-14.1149	4.1503	-3.401	1114	579
disyll+ complex coda	-22.0533	4.414	-4.996	1142	566

frequency (continuous)	-0.6614	1.5103	-0.438	
(Intercept)	93.1876	9.2026	10.126	

Table 4: Final simple model without significant interaction terms of age*class*vowel. Random effect of word (sd = 14.29) and random slope of speaker by vowel class (sd: STRUT = 40.6, FOOT = 36.37). T-values above ±2 are indicative of a significant effect when compared to the baseline factor in italics.

Table 3 shows the model from the mixed-effects linear regression on normalised F1 from the sociolinguistic interviews with the split speakers removed, leaving us with 8,237 tokens. Effects shown to be significant inclusions in the model by likelihood ratio test are social class, preceding segment, following manner, following voice, syllable category and, surprisingly, vowel class itself i.e. whether the word would be in the FOOT or STRUT category. Although age does not surface as significant, it does in an interaction with social class and vowel, but we save discussion of this for Section 4.4 and present the full model in Appendix A. A model including split speakers can be found in Appendix B. The next section will discuss each of the linguistic effects in turn, starting with the most surprising effect: vowel class.

4.3. Linguistic effects

4.3.1 Vowel class

If, as we have argued, Mancunians do not exhibit the FOOT-STRUT distinction, we would expect that the predictor of vowel class would not surface as significant in our statistical model. Or perhaps that it would be significant for some more middle class speakers who do not have the split proper, but have something approximating it. In fact, both the model and Figure 7 below demonstrates that there does seem to be an effect in the expected direction across the board, i.e. for all social classes.
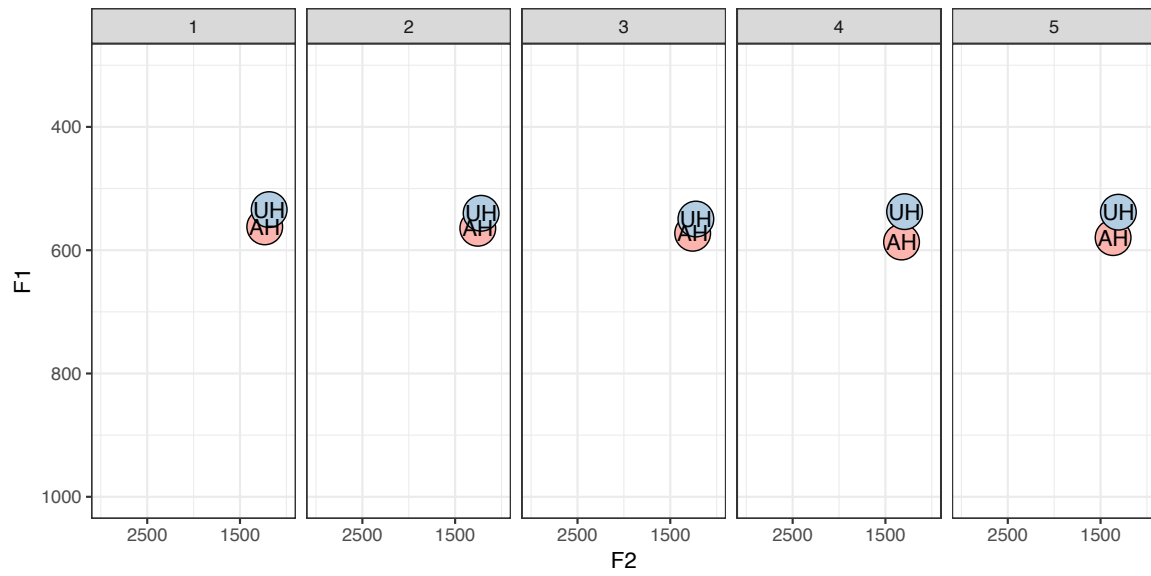
Figure 7: The two vowels across five socio-economic classes (speakers with true split removed). UH represents the *FOOT* class and AH the *STRUT* class.

As can be seen from Figure 8, although the difference is small (an effect size of just 28.7 Hz difference in the regression table in Table 3) it is still in the "right" direction. This difference

increases the higher up the social scale we go, as can be seen from zooming in on F1 in Figure 9.
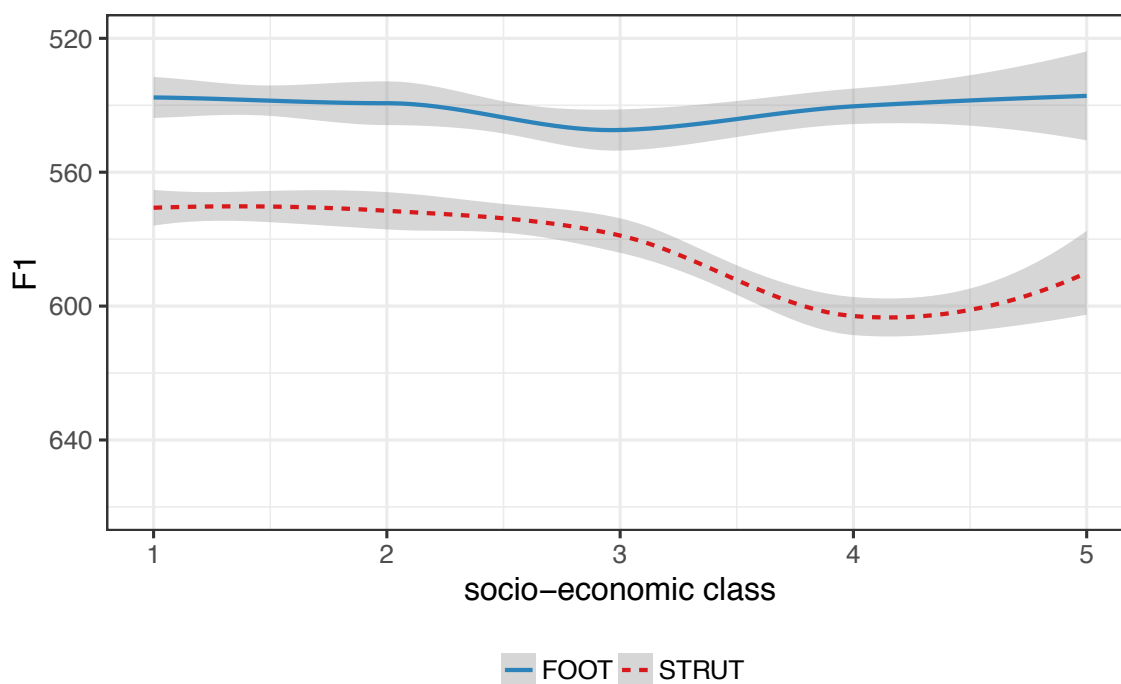

Figure 8: F1 smoothed across five social classes (grouped by speaker and word), split speakers removed

Figure 8 shows that, although the difference is small for the first three social classes (between 20-30Hz and arguably imperceptible) this increases for the two highest social classes, who have an average difference of around 60Hz. Note again that speakers with a true split have been removed in these graphs and the statistical model presented above (i.e. all the speakers included claim to make no distinction between the vowels in minimal pair tests), but the result holds for the speakers we have classified as having a non-distinct FOOT-STRUT. This is an intriguing result, as it is clear from our speaker judgements in minimal pair tests that these vowels do not represent separate phonemic classes for the vast majority of speakers. Next, we turn to phonetic environment effects as a way of explaining some of this variation.

4.3.2 Phonetic environment effects

The phonetic environment predictors in this model are preceding segment, following manner and following voice. Following place of articulation has no significant effect on the F1 of the vowel, and is not represented in this model. It is clear that certain segments have heightening or lowering effects on F1, although these effects are not always entirely straightforward. Preceding segment results in a more FOOT-like vowel for preceding voiced stops, /ʃ/, and /w/. Most of these sounds are rounded or have some form of lip-rounding, which matches the rounding of [ʊ]. Why preceding /d, g/ might result in a higher vowel (lower F1) and /t, k/ do not, however, is not exactly clear, but considering this is an effect size of 20Hz, it is not necessarily a result we are determined to find a clear interpretation for, although we would review this if the effect were to be replicated by other studies. Following manner results show that nasals and liquids result in a slightly higher F1 when compared to all obstruents, which is unsurprising given the known phonetic effects of nasals on F1, but is a very small effect (just 16Hz). Following affricates e.g. in words like *much, touch, Dutch* show particularly low F1s i.e. more FOOT-like vowels compared to other obstruents, which again may be due to the rounding associated with these segments. The following voicing effect is potentially surprising, with voiced segments resulting in lower F1. Given that vowels are shorter before voiceless consonants, we did expect to find F1 lowering in a larger proportion of the vowel when the following consonant is voiceless. An anonymous reviewer points out that voiced obstruents are associated with lowering of the larynx, which increases pharynx length, thus lowering F1 and that it is also possible that speakers compensate for the shorter time before voiceless obstruents by opening the jaw for the vowel faster than they do before voiced consonants. In turn, this could potentially lead to vowels reaching their target values more consistently before voiceless consonants and therefore showing higher F1 values before voiceless consonants than before voiced ones. This is certainly possible and something to be considered, but we are unable to find any corroboration of this kind of effect in the existing

literature. That, and the fact that most English speakers "voiced" obstruents are not truly voiced, means that we present this explanation with interest, but also caution. Finally, following syllable sequence shows that open syllables have lower F1s, but closed syllables of different kinds have higher F1s, an effect which has been found in the phonetics literature and has been shown to be a cross-linguistic tendency (e.g. Storme 2017). There is a strong effect of duration in that longer vowels have higher F1s. This, as already mentioned, is expected, given that the jaw has more time to open. There is no significant effect of word token frequency. Frequency may be expected to have a stronger effect on the F2 dimension where we would be observing reduction and centralisation.

One aspect of the phonetics which is difficult to account for in the statistical model is the combinatory 'tag-team' effects of certain segments flanking the vowel. For example, the summary from Britain's analysis of the Fens (Table 1) indicates that adjacent segments such as /v/ can increase F1, as can following /l/ or / ð/, but in the presence of a following /g/ or /ʃ/, the F1 may be lower. Some of these findings have been replicated here. Although there is clearly some coarticulation of segments with a high tongue body (e.g. /g/) or rounded lips (e.g. /p, b, w/), there is no neat label by which we can categorise these. Inputting individual segments as preceding/following would cause difficulty in running in the model due to having empty cells of numerous combinations. To refine our argument, it may be worth looking at the lexical incidence of FOOT-STRUT vowels and their surrounding environments.

4.3.3 The co-occurrence of FOOT-STRUT tokens with adjacent phonetic forms

This section seeks to demonstrate that although we may not be able to provide a clear encapsulation of the phonetic environment effects to fully account for the FOOT-STRUT split in our data (as vowel class remains a significant predictor), the lexical incidence of FOOT-STRUT vowels in English with particular phonetic forms is very clear. As we will now show, FOOT

vowels tend to occur with those sounds which lower F1 (thus resulting in a higher vowel), and STRUT vowels tend to occur with those sounds which result in a higher F1 (lower vowel). This argument may appear somewhat circular, but the point is this: the various heightening/lowering effects of F1 cannot be fully captured by our statistical model, but the lexical incidence of these various effects is difficult to reject.

The numbers from the following graphs have been calculated as follows. Word frequency per million words was calculated for each word in the dataset using the SUBTLEX corpus. We then plotted proportional tables based on surrounding segments and frequency per million words to work out how often certain segments occur with FOOT vowels and how often they occur with STRUT vowels. For example, Figure 9 shows how often (when taking into account frequency per million words) FOOT-STRUT vowels co-occur with a following plosive. We can think of this in the following way: in a given set of one million FOOT tokens, 75% have a following plosive, but for the equivalent number of STRUT tokens, less than 35% do. Recall that our regression shows that following plosives give a lower F1, which is most likely due to the fact that vowels tend to be shorter before plosives (e.g. Peterson & Lehiste 1960) and, as discussed above for duration, the jaw has less time to open resulting in a closer vowel.
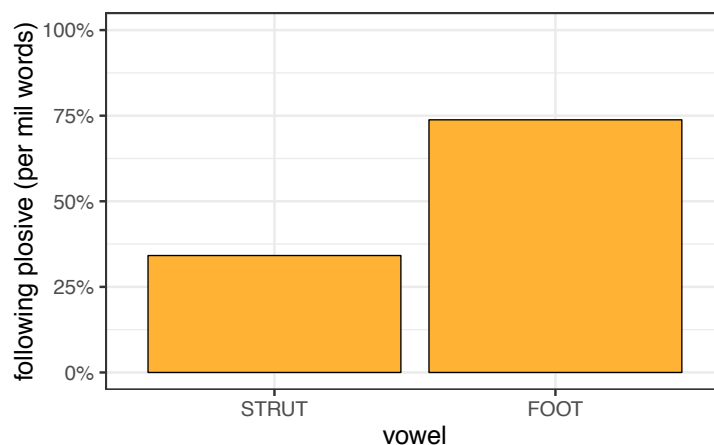

*Figure 9: Frequency of **FOOT-STRUT** lexical occurrence. Following plosives tend to be found much more with FOOT words when compared with STRUT words*

Figures 10 & 11 show the frequency of occurrence of following nasals and preceding /p, b, w/ respectively for each lexical set. Figure 10 shows that following nasals account for over 35% of STRUT tokens, but very few of FOOT. In fact, as Table 5 shows, the only FOOT lexical item in our dataset with a following nasal is the word *woman* (with a moderate to low frequency of 166 per one million words), a disyllabic word. Syllable structure is accounted for by our statistical model, and an open syllable is more likely to result in a lower F1 (higher vowel), whereas a following nasal results in a higher F1. Finally, Figure 11 shows that preceding rounded tokens take up a much higher proportion of FOOT vowels than STRUT vowels, accounting for almost 50% of FOOT items in our frequency calculations. By far the most frequent words for each set according to SUBTLEX figures for this context are *would* for FOOT and *but* for STRUT. However, note that *but* only occurs six times in our dataset because unstressed tokens have been filtered out. This is potentially a functional explanation for why *but* became part of STRUT despite its phonetic environment favouring a more FOOT-like realisation: it is almost always in a reduced position syntactically, and thus is realised with a more centralised pronunciation.[15] This explanation arguably breaks down at *should, would, could,* but then again these modals can and do occur in stressed position much more often than *but*. A closer investigation of the typical prosodic context these words tend to find themselves in would give us a clearer picture of whether this hypothesis is viable.
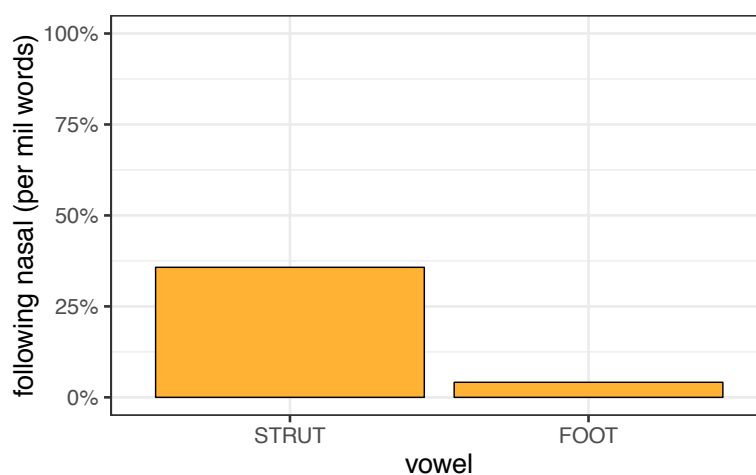
*Figure 10: Frequency of **FOOT-STRUT** lexical occurrence. Following nasals occur much more frequently with STRUT vowels*
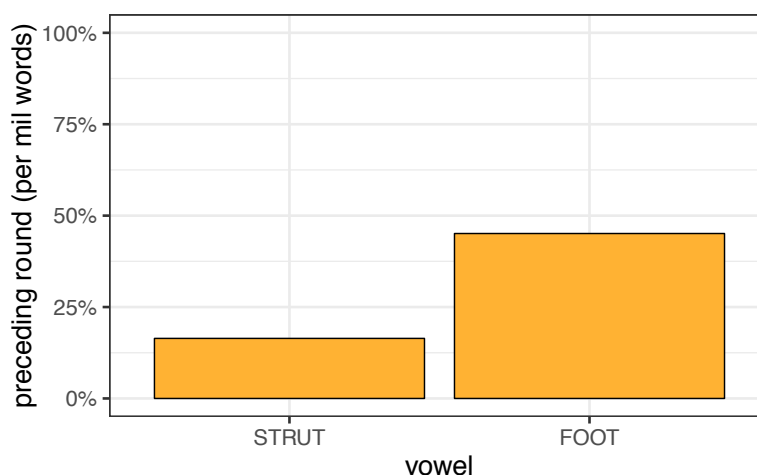

*Figure 11: Frequency of **FOOT-STRUT** lexical occurrence. Preceding /p, b, w/ occur much more frequently with FOOT vowels than STRUT*

Although these effects alone should easily be able to be accounted for by a statistical model, it is the combinatory effect of more than one preceding and following coarticulatory segment that is difficult to capture. For example, often a following plosive occurs with a preceding round segment (*would, put*) which may have a cumulative effect on the vowel height, but not all cross-sections of preceding and following segments exist (e.g. preceding round, following nasal in a closed syllable does not exist) so it is problematic to enter all of this detail into the statistical model. Table 5 demonstrates frequent and infrequent occurrences in the data with example words in the favouring and disfavouring contexts. In summary, FOOT words consistently co-occur with phonetic environments that favour a higher vowel. STRUT words

consistently co-occur with phonetic environments that favour a lower vowel. This is simply an observation about English.

Context	FOOT words		STRUT words	
	most frequent words	**frequency per million**	**most frequent words**	**frequency per million**
following plosive **and** preceding round	*would*	2820	*but*	5774
	put	972	*public*	278
	book	162		
	wood	60		
preceding round	*pull*	87	*budget*	232
	push	83	*above*	213
following plosive	*good*	2255	*couple*	261
	could	1863	*cut*	225
	look	1693	*luck*	177
	should	991		
preceding round **and** following nasal	*woman* [only example]	166	*won*	232
			wonderful	212
following nasal			*some*	1885
			come	1802
			something	1038
			done	976
			money	690

Table 5: Frequent contexts and word examples

Our results illuminate a situation that has received little attention in the field, that is the fact that classes of sounds that are not phonologically distinct can appear to be differentiated phonetically. This is not the first time an apparent separation in two seemingly non-distinct vowel classes has surfaced due to the coincidence of contexts in which the relevant words find

themselves in. Fridland et al. (2014: 345) find a similar situation in the durational measures of the LOT-THOUGHT vowels in their California and Nevada speakers, who otherwise clearly have the low-back merger. They point out that findings such as these may be the key to understanding some of the more perplexing aspects of language change, and provide an explanation for the impossibility of unmerging of lexical sets. In addition, Thomas (2019:133) notes a similar pattern in his Mexican American English speakers' realisations of the low-back LOT-THOUGHT vowels. Thomas argues that, although it may look like there is separation between the two sets acoustically for his speakers, the asymmetrical distributions of the vowel classes are responsible for the apparent distinction. Note that in American English, unlike British English, CLOTH words tend to fall into the THOUGHT set e.g. *off, soft* and thus this set has many frequent words before /f/. As Thomas explains, a following /f/ reduces F2, and these really only occur in the American English LOT set in a following syllable e.g. *officer.*[16] On the other hand, the American English LOT set tends to co-occur with contexts that promote a raised F2. Two of the most common words LOT words are *not, got* which not only have flanking consonants that may raise F2, but are also frequently in syntactically reduced positions, resulting in centralisation. Thomas (2019:133) explains the effect in the following way, which is pertinent to the present study also:

> [T]he coarticulatory factors of the neighbouring consonants are magnified [and] these factors conspire to reduce the efficacy of an index of separation […] For BOT and BOUGHT, determination of a phonological contrast should be made by comparison of similar phonetic contexts.

This serves as a warning for researchers using large-scale datasets to unveil patterns of distinction or merger in spontaneous speech based on statistical patterns only. As Thomas argues, comparison of the pairs in the same phonetic context is required to avoid such contextual effects skewing the results. This is something we can do in the present investigation,

thanks to the minimal pair tests, which we will turn to in the next section. The fact that previous scholars (Braber & Flynn 2015) have observed this same pattern in non-distinct FOOT-STRUT gives us added confidence in our argument.

At this point, we conclude for the time-being that the phonetic effects on F1 are more complex than the predictors captured in our model. As Figures 9-11 show, there are various conditioning effects which team up, meaning that words which fall into the FOOT set tend to be those for which adjacent segments induce a lower F1, and those which fall into the STRUT set tend to be those for which adjacent segments promote a higher F1. Further corroboration of this argument should come from production in minimal-pair tests, as a prediction arising from this would be that speakers will produce minimal pairs the same, as they have identical surrounding environments despite being in different lexical sets.

4.3.4 Acoustics of minimal pair tests

Above, we have argued that it is the cumulative effects of various phonetic effects which create apparent differences between FOOT and STRUT lexical sets in Mancunian speakers without a distinction. We demonstrate that patterns in English mean that speakers without a distinction may appear to have one in the phonetics due to the co-occurrence of phonetic segments with prescribed lexical sets. If this is truly the case, we should see that the acoustics of the minimal-pair tests, where the phonetic environment is identical, but the lexical set is different, should be the same if speakers truly do not have the split. As we shall show below, this argument is corroborated by the production of our minimal-pair tests, in which speakers produce identical tokens when the surrounding segments are the same.

Figure 12 plots the production of *book* and *buck* in acoustic space for a selection of working class and upper middle class speakers. The working class speakers in the top half of the plot (selected from lower and upper working) demonstrate that these words are identical in acoustic

space, with negligible differences for some speakers. The working class speakers are representative of the vast majority of speakers in our dataset for these pairs. The upper middle class speakers have been selected to demonstrate both the contrast and lack of contrast dependent on the speaker themselves. Speakers such as Jacob A (21 years old) and Wilma L (57 years old) have no evidence of a phonemic distinction from these pairs. However, these speakers are only representative of around 50% of our upper middle class contingent. Speakers such as Helen C (45 years old), Mike T (26 years old), and Matthew P (66 years old) have a clear phonemic distinction in their production (confirmed in casual speech; see Figure 6) as well as their own perception. This tallies with both Wyld's (1936) and Labov's (1994) claims that acquisition of the distinction in Northern speakers is most likely something that will appear in a minority of speakers, based on education, parental input or individual motivation. We discuss some of the possible mechanisms behind this below in the next section.
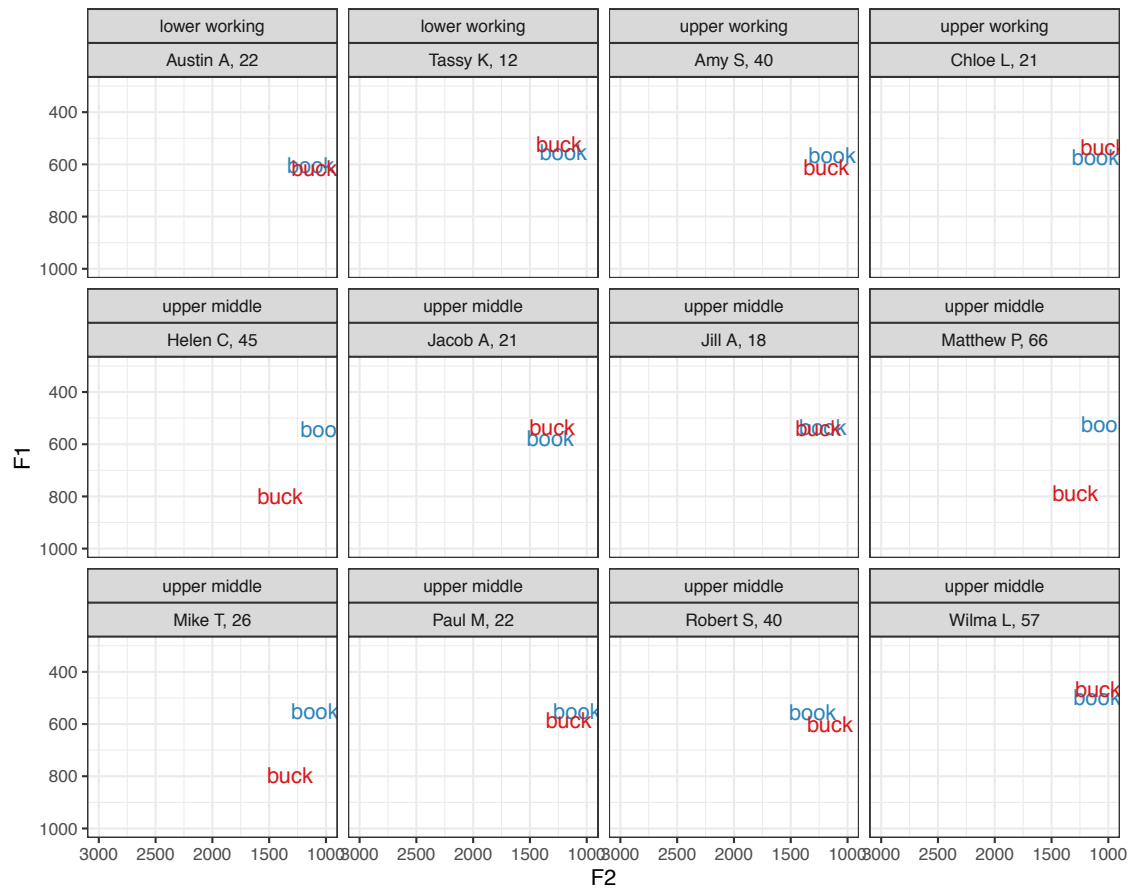
Figure 12: A selection of speakers minimal pair productions of *'book' and 'buck'* demonstrates that the vast majority of speakers have identical productions, but some upper middle class speakers have a true distinction.

4.4 Social effects

This section will primarily focus on the socio-economic class effects in the data, but first we briefly discuss additional social factors. There is a small effect of sex, with females showing a slightly higher F1 (lower vowel) than males, somewhat in line with Strycharczuk et al.'s (2019) results. Note that this is an overall finding in the acoustics of both vowels combined. We find no evidence that women are more likely to have a phonemic distinction, just that they have a lower vowel (this is confirmed by likelihood ratio test of a model which includes an interaction between vowel class and sex and one that does not).

Ethnicity does not seem to play a role in this distinction, although a model which includes an interaction between ethnicity and vowel class does perform slightly better than one without. Closer inspection of this shows that Pakistani and White speakers are more likely to have a difference between the two vowel classes than Black speakers. We approach this modest result cautiously, however, as our numbers for ethnic minority speakers are not yet at a level where we are able make strong conclusions about small statistical significances, particularly when we do not have appropriate representation of ethnic minority groups across the social class spectrum (for example, British Pakistani speaker Robert S is our only non-white upper middle class speaker). This will be an avenue for future study, considering the acquisition of local forms from ethnic minority groups, particularly in relation to their arrival in the area. Drummond (2012) demonstrates that Polish speakers arriving in Manchester retain a distinction between FOOT and STRUT, but the next question is whether the children of these speakers would do so. This is the situation we are in with our Black and Pakistani speakers who were born in Manchester (or have been lived there since the age of 3) but do not necessarily have parents from the area. However, many more factors will need to be taken into consideration, alongside a larger speaker set, before we are able to address these questions.

In terms of socio-economic class, the effects are clear: the higher the social class, the lower the vowel. This is a monotonic effect across the five social classes. What is more interesting is the increasing difference between the two lexical sets as we move up the social scale, even with the exclusion of the split speakers, as shown in Figures 7 and 8 above. The distance between the FOOT class and the STRUT class is bigger for the middle and upper middle classes (4-5) than for the lower working to lower middle classes (1-3), and the adjacent phonetic effects cannot explain all of the variation found here. More specifically, this difference between social classes cannot be accounted for by coarticulatory effects alone. For the first three socio-economic groups, this is stable, but jumps when we get to the upper middle

classes in Table 3, with the STRUT vowel being considerably lower (higher F1). This effect is accounted for in an additional model which includes a three-way interaction with both socio-economic class and vowel, but also age. This model is an improvement on the model in Table 4 (confirmed by likelihood ratio test) and can be found in Appendix A.

Figure 13 visualises the extremes of this interaction by focussing on the lowest and highest socio-economic groups (split speakers included). In the left panel, we can see the lower working class speakers who have a small difference between FOOT-STRUT vowels in the expected direction and are stable across all ages. We have argued that this small difference, around 25Hz here, is highly unlikely to map onto any phonemic distinction. We have also argued that the difference in the expected direction reflects the fact that STRUT words tend to have favouring environments for a higher F1 (lower vowel). This argument, however, cannot account for the fact that the highest social classes show a stronger effect than the lower ones, unless we are to believe they are somehow more susceptible to this coarticulation. However, Figure 13 shows that the youngest speakers in the upper middle classes actually show a pattern similar to the working classes. It seems to be older speakers that have a grasp of the distinction. This is reflected in the minimal pair individuals plot in Figure 12, where three of the four young upper middle class speakers (age 26 and below) have no evidence of a distinction. This result is suggestive of age grading, meaning that middle class speakers may acquire the distinction as they age. The fact that young upper middle class Mancunians do not have the distinction is, arguably, not surprising. The vast majority of Northerners are not aware of the FOOT-STRUT distinction growing up, and it is not salient to Northerners as being a Northern form. This is discussed by Trudgill (1986) and demonstrated by Evans & Iverson's (2004) perceptual judgement tasks from Northerners still living in the North. This is in contrast to the TRAP-BATH distinction found in RP and the South, possibly the most overtly salient form Northerners

would point to as being the difference between their speech and that of Southern British English speakers.

Figure 13: The lowest and highest socio-economic groups variation in FOOT vs. STRUT words across speaker age

Perhaps paradoxically though, or perhaps not, it is the STRUT vowel which some upwardly-mobile Northerners do move towards when moving to the South, or entering upwardly-mobile professions. As Trudgill (1986: 18) describes:

> It is... interesting to note that Northerners moving to the South and accommodating to Southern speech usually modify butter /bʊtə/ to /bʌtə/ or at least to /bətə/, but ... would rather drop dead than say /dɑːns/: the stereotype that this is a Southern form is too strong.

STRUT falls below the radar, in a way, as it is not a phonemic category for most in the first place. Most who do become aware of it, on some subconscious level or not, do so later in life, perhaps coming into contact with Southern speakers. This is unlike the Southern /ɑ:/ vowel in BATH, the same vowel which Northerners have in the PALM lexical set. As Sankoff (2004) puts it, Northerners can acquire the phonetics of STRUT with no social cost. Perhaps it is the case that Northerners do not see STRUT as Southern, but as educated. Social judgement and matched guise tests are possible ways to confirm this.

It may just be, as Wells (1982:353) says, interpreted as a 'poshing-up' of FOOT. We do not have any clear cases of hypercorrect FOOT in our sociolinguistic interviews, and as contended above, the sociolinguistic interview is not the ideal corpus to investigate this due to its target of the true vernacular. However, hypercorrect FOOT does occur in the speech of aspiring middle-class Northerners, in our everyday experience, ones who actually have little contact with the South. This would be consistent with Evans & Iverson's (2004) findings that Northerners in the South have a perceptual category for STRUT but Northerners in the North do not. Evans & Iverson (2007: 3815) express surprise that the Northerners in the North seem to have no awareness of a lower STRUT vowel, given the amount of exposure to RP and Standard Southern British English accents Northern speakers get from television. As discussed above, Labov (1994: 348) also asserts that mass media influence should not be dismissed for individuals wishing to acquire a FOOT-STRUT distinction later in life, and it is entirely possible that certain highly-motivated individuals may use the media as a template. However, Evans & Iverson's (2004, 2007) results indicate that something approximating a split cannot be achieved passively, and that speaker efforts will become more successful with greater exposure to the South and southern speakers.

Thus, it is not surprising that our young upper middle class speakers are not aware of the distinction yet. It may only be as they get older, become more mobile, enter the workplace,

or have more contact with Southern speakers at university that we might expect this distinction to start taking hold (see also Baranowski 2017 on GOAT-fronting in Manchester, another feature led by the highest-status groups in the city, arguably brought in from the South). Given enough exposure to speakers with the split, the distinction is relatively simple to work out from the spelling (although there are numerous exceptions e.g. *blood* vs. *good, but* vs. *put;* cf. Chambers 1992:700). The changing demographics of some areas of the city may change the point in time that young people are exposed to the split in future. For example, some areas of Manchester, such as Chorlton-cum-Hardy, have seen a massive influx of Southerners since the move of the BBC to Salford in 2012. This is an area ripe for investigation, from the effect of the southern children of BBC staff attending both private and comprehensive schools in the area, to the effect on the middle classes throughout South Manchester.

An alternative explanation could be that older speakers are of a generation where RP was promoted in schools, and most certainly in the kinds of prestigious private schools attended by the upper middle classes. Teachers are no longer as strict in this respect, and it is unlikely that younger students would be persecuted for lacking a split today as their parents' generation might have been. Even without a phonemic distinction, a suitably low STRUT variant may pass as acceptable. The only way to confirm this would be via a lifespan investigation, for which this variable would provide fruitful study. As discussed above, it seems possible that speakers are able to shift their gradient phonetics throughout their lifespan in a regular way, and reweight the frequency of occurrence of categorical variables. The evidence thus far seems to be against a speaker's ability to acquire a phonemic split in adulthood. Thus, the point still remains as to whether these younger speakers will be able to acquire a true split in later life. This may be down to individual speaker motivation and something Northerners do to sound more educated.

As summarised in Section 2.3, Chiu & Evans's (2018) results from Northerners living in London demonstrate the overall argument perfectly. Upwardly mobile Northerners can do a

very good job of mimicking a Southern FOOT-STRUT distinction, but when it comes down to it, if this phonemic distinction is not acquired at a young age, i.e. pre-adolescence, they can never truly look like the Southerners. They can 'fake it' fairly well, but they cannot reach the same grammatical representation as a true native, as the EEG results unveil. Such results, alongside the results from the present investigation, open the door for a series of studies on this variable in upwardly mobile speakers in the North. These could range from the sociolinguistic judgement of a lowered STRUT vowel, to perceptual tests of learning a phonemic distinction vs. learning a lowered vowel overall. Chiu & Evans (2018) have already gone some of the way to looking at this, but crucially they investigate speakers living in the South. How motivated speakers still based in the North perform on such tasks, or upper middle class speakers returning from university, are areas ripe for future research.

4.5 Preliminary observations: fronting and rounding

We have not discussed possible fronting, centralisation or rounding in the present investigation. In terms of rounding, some scholars have argued that [ʊ] is not the most accurate phonetic representation for present-day Northern speech, and that something unrounded approximating [ɤ] is more accurate. The variant found in Manchester certainly has some degree of rounding, but not as rounded and high as [ʊ] would suggest. Perhaps mid-back [ʊ̞] is the closest description to the Mancunian FOOT (see also Wells 1982: 352) but this remains an open question for now. Certainly some areas of the country (e.g. areas in the Midlands) auditorily sound more rounded than areas such as Manchester. In terms of F2 (fronting or centralisation), preliminary results show the same monotonic effect of centralisation across social-class, with upper middle class speakers having considerably fronter vowels in comparison to the lower working class (around 120Hz between the two extremes). These initial results show no significant interaction with vowel class, demonstrating that this is a centralisation effect across the board, rather than these higher social groups being more likely to show some kind of

phonemic distinction in frontness. However, these comments should be taken with caution at this stage, as they are based on preliminary inspection of the automatic measurements provided by FAVE and we have not investigated this in any detail at this point in time.

5. CONCLUSION

This investigation of the FOOT-STRUT vowel(s) in Manchester has demonstrated that the vast majority of speakers have no phonemic split between the two lexical sets, but that lower, more Southern-like, articulations of the vowel are found in the higher social classes. Speakers with a true phonemic distinction are in the minority (just eight of our speakers) and are restricted to the middle and upper middle classes. Even for these speakers, there is evidence that not all words are assigned to the expected lexical set (recall that *crux* was pronounced with the FOOT vowel for one of our speakers who otherwise demonstrated a true split). However, the statistical models show a small but robust phonetic effect in the expected direction: FOOT vowels are consistently higher than STRUT vowels (lower F1) for all social classes. We have argued that this is down to combinatory coarticulatory effects of the surrounding segments, which are not fully accounted for in our model, and examples from frequency of co-occurrence in English are presented from the SUBTLEX corpus to support this argument. Our results are a potential warning to researchers attempting to diagnose the presence or absence of mergers by purely statistical means, as claims of lack of merger may be premature and due to unaccounted-for phonetic effects. Nevertheless, the upper middle classes show a larger effect of lexical set, which cannot be explained away by coarticulation. We suggest that this is due to age grading, as young upper middle class speakers do not show the same phonetic differences in their vowel categories when compared to middle aged and older speakers. This result is consistent with upper middle class speakers being exposed to more Southern forms as

they age, perhaps coming into contact with speakers from other areas of the country. If the STRUT vowel carries connotations of educated speech, rather than Southern speech, they can acquire this form without necessarily being aware of abandoning their Northern vowel. Of course, perceptual judgement tests would be needed to test this further. We have cast doubt on the ability of such speakers to truly acquire a phonemic contrast, although the dataset presented here cannot confirm or deny such a fact.

This study has illuminated various avenues for future research, from a phonetic, sociolinguistic and cognitive perspective. There are many questions we raise here for which we are unable to answer with our current corpus of sociolinguistic interviews. In terms of our own dataset, the results of this study are based purely on the acoustics correlates of vowel height and we highlight preliminary results for alternative acoustic measures.

Overall, we have argued that the phonetic conditioning of the FOOT-STRUT split is as complex synchronically as it was diachronically. The possibility that vowel class remains significant should perhaps not be entirely surprising after all, as it is the best summary of the forces behind the historical split: it encapsulates the various phonetic effects responsible for the change.

Appendix A: model including interaction

	Estimate	Std. Error	t value
vowel class			
STRUT (baseline)			
FOOT	-24.5772	7.0364	-3.493
social class (baseline LWC)			
UWC	2.2785	8.7502	0.26
LMC	6.225	8.6595	0.719
MMC	30.3212	9.3125	3.256
UMC	86.8308	15.7492	5.513
age (continuous)	0.2695	0.2729	0.987
sex (baseline female)			
male	13.1091	3.8281	3.424
ethnicity (baseline black)			
pakistani	4.8568	7.0852	0.685
white	-8.8588	5.7268	-1.547
duration (log transformed)	37.4533	1.7475	21.432
preceding segment (baseline voiced stops, /ʃ/ and w)			
other	19.6206	3.2976	5.95
following voice (baseline voiceless)			
voiced	-18.754	3.3615	-5.579
following manner (baseline obstruent)			
affricate	-23.0728	7.2846	-3.167
sonorant	18.1112	3.5267	5.135
syllable (baseline closed syllable)			
open syllable	-11.99	3.4457	-3.48
monosyllabic complex coda	-14.0084	4.0084	-3.495
disyllabic+ complex coda	-22.281	4.2311	-5.266
			-0.293
frequency	-0.4235	1.4471	
(intercept)	87.764	9.9323	8.836

interactions			
vowel STRUT: class UWC	-0.5751	9.0255	-0.064
vowel STRUT: class LMC	1.118	8.7965	0.127
vowel STRUT: class MMC	-31.5219	9.4473	-3.337
vowel STRUT: class UMC	-90.3487	16.4705	-5.485
vowel STRUT: age	-0.378	0.2777	-1.361
class UWC: age	-0.3315	0.4282	-0.774
class LMC: age	-0.2541	0.5198	-0.489
class MMC: age	1.3554	0.5378	2.52
class UMC: age	0.7586	0.6799	1.116
vowel STRUT: class UWC: age	-0.1183	0.4497	-0.263
vowel STRUT: class LMC: age	0.131	0.5282	0.248
vowel STRUT: class MMC: age	-1.9818	0.5364	-3.695
vowel STRUT: class UMC: age	-0.2744	0.6973	-0.393

Appendix B: Model including speakers with distinction and interaction

	estimate	std. error	t-value	N	mean F1
vowel class (baseline STRUT)				6245	588
FOOT	-24.44338	7.0689	-3.458	2586	538
social class (baseline LWC)				2020	557
UWC	2.49954	8.71967	0.287	1828	564
LMC	6.65604	8.62803	0.771	2387	569
MMC	30.28878	9.30016	3.257	1838	582
UMC	86.72534	15.69276	5.526	758	628
age (continuous)	0.26425	0.27192	0.972		
sex (baseline female)				5011	570
male	13.33387	3.84677	3.466	3820	577
ethnicity (baseline black)				997	559
pakistani	6.64275	6.97995	0.952	1378	580
white	-6.6171	5.60025	-1.182	6456	574
duration (log)	37.50249	1.74697	21.467		
preceding segment (baseline voiced stops, /ʃ/ and /w/)				1918	549
other	19.61431	3.29745	5.948	6913	580

following voice (baseline voiceless)				3222	580
voiced	-18.50222	3.35555	-5.514	5609	569
following manner (baseline obstruent)				5807	567
affricate	-23.0415	7.29199	-3.16	527	594
sonorant	17.82306	3.53138	5.047	2497	583
syllable (baseline closed syllable)				4455	573
open syllable	-12.12694	3.44034	-3.525	2325	574
monosyllabic comp coda	-13.96852	4.00454	-3.488	939	578
disyll+ complex coda	-22.24233	4.22681	-5.262	1111	567
frequency (continuous)	-0.34126	1.43603	-0.238		
interactions					
vowel STRUT: class UWC	-0.1291	9.04845	-0.014		
vowel STRUT: class LMC	1.07393	8.8253	0.122		
vowel STRUT: class MMC	-31.56403	9.47915	-3.33		
vowel STRUT: class UMC	-90.39774	16.51921	-5.472		
vowel STRUT: age	-0.37723	0.27868	-1.354		
class UWC: age	-0.31797	0.42666	-0.745		
class LMC: age	-0.24784	0.51794	-0.479		
class MMC: age	1.37511	0.53549	2.568		
class UMC: age	0.76336	0.67747	1.127		
vowel STRUT: class UWC: age	-0.05116	0.44981	-0.114		
vowel STRUT: class LMC: age	0.13369	0.5299	0.252		
vowel STRUT: class MMC: age	-1.9806	0.53812	-3.681		
vowel STRUT: class UMC: age	-0.27433	0.69941	-0.392		
(Intercept)	85.73284	9.8529	8.701		

Appendix B: Final model including speakers with the split, and with significant interaction terms of age*class*vowel. Random effect of word (sd = 14.22) and random slope of speaker by vowel class (sd: STRUT = 25.40, FOOT = 27.52).

REFERENCES

Baranowski, Maciej. 2007. *Phonological variation and change in the dialect of Charleston, South Carolina.* (Publication of the American Dialect Society 92.) Durham, NC: Duke University Press.

Baranowski, Maciej. 2013. On the role of social factors in the loss of phonemic distinctions. *English Language and Linguistics* 17/2: 271-295

Baranowski, Maciej. 2017. Class matters: the sociolinguistics of GOOSE and GOAT in Manchester English. *Language Variation and Change* **29**:301–339.

Baranowski, Maciej & Danielle Turton. 2015. Manchester English. In *Researching Northern Englishes*, ed. Raymond Hickey, pp 292-316. Amsterdam and Philadelphia: John Benjamins.

Baranowski, Maciej & Danielle Turton. 2018. Locating speakers in the socioeconomic hierarchy: towards the optimal indicators of social class. Paper presented at New Ways of Analysing Variation (NWAV) 47, New York University.

Barr, Dale J, Roger Levy, Christoph Scheepers, & Harry J Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**: 255–278.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**:1–48.

Beal, Joan C. 2010. *An introduction to regional Englishes: dialect variation in England.* Edinburgh: Edinburgh University Press.

Beal, Joan C. 2012. "By those provincials mispronounced": the STRUT vowel in eighteenth-century pronouncing dictionaries. *Language and History* **55**(1): 5-17.

Bowie, David. 2019. Individual variation in the development of the Western Vowel System of Utah. *Linguistics Vanguard*, 5(2).

Braber, Natalie, & Nicholas Flynn. 2015. The East Midlands. In Researching Northern
    Englishes, ed. Raymond Hickey, pp 369-392. Amsterdam and Philadelphia: John
    Benjamins.

Britain, David. 2013. Space, diffusion and mobility. In J K Chambers, N. Schilling,
    Handbook of Language Variation and Change (second edition).The handbook of
    language variation and change, pp.471-500. Wiley.

Britain, David. 2015. Between North and South: The Fenland. In Researching Northern
    Englishes, ed. Raymond Hickey, pp 417-436. Amsterdam and Philadelphia: John
    Benjamins.

Britain, David, Adrian Leeman, Tam Blaxter, Marie-Jose Kolly, Sarah Grossenbacher,
    Melanie Calame & Daniel Wanitsch. 2016. Big bad dialectological data? Lessons
    from the English Dialect App. Paper presented at Newcastle University Linguistics
    seminar series, October 2016.

Chambers, J. K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University
    Press. Second Edition.

Chiu, Faith & Bronwen G. Evans. 2018. Learning and adjusting for a FOOT-STRUT split:
    production, discrimination and processing data from northern British English speakers
    in London. Paper presented at the 8th Northern Englishes Workshop, 27-28 March
    2018.

Conn, Jeff & Uri Horesh. 2002. Assessing the acquisition of dialect variables by migrant
    adults in Philadelphia: A case study. *University of Pennsylvania Working Papers in
    Linguistics*: 8(3), Article 5.

Eckert, Penelope & William Labov. 2017. Phonetics, phonology and social meaning.
    Journal of Sociolinguistics, **21**(4), 467-496.

Ellis, Alexander. 1889. *On Early English Pronunciation. With especial reference to Shakspere and Chaucer. Vol I-V. London: The Philological Society.*

Evans, Bronwen G. & Paul Iverson. 2004. Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *The Journal of the Acoustical Society of America* **115**, 352–361.

Evans, Bronwen G. & Paul Iverson. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America* **121**, 3814–3826.

Ferragne, Emmanuel, and François Pellegrino. 2010. Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics. *Journal of Phonetics* 38, 526–539.

Flynn, Nicholas Edward John. 2012. A sociophonetic study of Nottingham speakers. Doctoral Dissertation, University of York.

Halfacre, Caitlin. 2019. North-South dividers in privately educated speakers: a sociolinguistic study of Received Pronunciation using the FOOT-STRUT and TRAP-BATH distinctions in the North East and South East of England. In *Proceedings of the International Congress of Phonetic Science 2019*, 5-9 August, Melbourne.

Harrington, Jonathan. 2006. An acoustic analysis of `happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics* 34, 439–457.

Herzog, Marvin I. 1965. The Yiddish language in northern Poland: Its geography and history. Special issue of International Journal of American Linguistics 31.2, part 2.

van Heuven, Walter JB, Pawel Mandera, Emmanuel Keuleers, and Marc Brysbaert. 2014. SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology* **67**, 1176–1190.

Hickey, Raymond. 1999. Dublin English: current changes and their motivation. In *Urban voices: Accent studies in the British Isle*s, eds. Paul Foulkes and Gerard Docherty. London: Arnold. pp. 265-281.

Hughes, Arthur, Peter Trudgill, and Dominic Watt. 2005. English Accents and Dialects: An Introduction to Social and Regional varieties of English in the British Isles, 4th edition. London: Hodder Arnold.

Kenyon, John Samuel and Thomas Albert Knott. 1944. *Pronouncing dictionary of American English*. Merriam-Webster.

Kerswill, Paul. 1996. Children, adolescents, and language change. Language variation and change **8**, 177–202.

Kurath, Hans and Raven Ioor McDavid. 1961. *The pronunciation of English in the Atlantic States: Based upon the collections of the linguistic atlas of the Eastern United States*. University of Michigan Press.

Labov, William, 1972. Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press.

Labov, William. 1984. Field methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Sherzer(eds.), *Language in use: readings in sociolinguistics*. Englewood Cliffs: Prentice Hall. 28–54.

Labov, William. 1994. Principles of linguistic change: Internal factors. Oxford: Blackwell.

Labov, William. 2007. Transmission and diffusion. *Language* **83**, 344–387.

Labov, William. 2011. Plotnik 10.1. URL http://www.ling.upenn.edu/~wlabov/Plotnik.html

Lass, Roger. 2000. Phonology and morphology. In *The Cambridge History of the English Language*, vol. iii 1476-1776 , ed. Norman Blake, 23–155. Cambridge: Cambridge University Press.

Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English Dialects App: The creation of a crowdsourced dialect corpus. *Ampersand* **5**, 1-17.

Lindsey, Geoff. 2019. *English After RP: Standard British Pronunciation Today*. Cham: Springer International Publishing : Imprint: Palgrave Macmillan.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. The *Journal of the Acoustical Society of America* 49: 606–608.

Lonergan, John, 2015. A phonetic analysis of back vowel raising in Dublin English. In ICPhS. In *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow.

MacKenzie, Laurel. 2017. Frequency effects over the lifespan: a case study of Attenborough's r's. *Linguistics Vanguard*, **3**(1).

MacKenzie, Laurel, George Bailey & Danielle Turton. 2016. Our Dialects: Mapping variation in English in the UK [Website]. http://tiny.cc/OurDialects

MacKenzie, Laurel, George Bailey & Danielle Turton. In prep. *Dialectal variation in North-West England: mapping phonological, syntactic and lexical variation across changing borders*. Manuscript in preparation.

Maguire, Warren. 2009. A survey of accents of English in Britain and Ireland [Website]. http://www.lel.ed.ac.uk/~wmaguire/survey/survey.html

Minkova, Donka. 2014. A Historical Phonology of English. Edinburgh: Edinburgh University Press.

Morris, Jonathan. 2013. Sociolinguistic variation and regional minority language bilingualism: an investigation of Welsh-English bilinguals in North Wales. PhD thesis, University of Manchester.

Orton, H., S. Sanderson, and J. Widdowson, eds. 1978. *The linguistic atlas of England*. London: Croom Helm.

Payne, Arvilla, 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.), *Locating language in time and space*. New York: Academic Press, 143-78.

Peterson, Gordon E, and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* **32**: 693–703.

Raumolin-Brunberg, Helena. 2005. Language change in adulthood: Historical letters as evidence. *European Journal of English Studies*, *9*(1), pp.37-51.

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard, and Jiahong Yuan. 2014. FAVE 1.1.3. URL: http://dx.doi.org/10.5281/zenodo.9846

Sankoff, Gillian. 2004. "Adolescents, young adults and the critical period: Two case studies from 'Seven Up'". Sociolinguistic Variation: Critical reflections, ed. Carmen Fought 121–39. Oxford: Oxford University Press.

Sneller, B., 2018. Mechanisms of phonological change. PhD thesis, University of Pennsylvania.

Storme, Benjamin. 2017. Closed-syllable vowel laxing: A contrast enhancement strategy. *The Journal of the Acoustical Society of America*, *141*(5), p.3583.

Strycharczuk, Patrycja, Georgina Brown, Adrian Leemann and David Britain. 2019. Investigating the STRUT- FOOT distinction in Northern Englishes using crowdsourced data. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne.

Thomas, Eric. 2019. Trends from Outside. In Eric Thomas (Ed.), *Mexican American English: Substrate Influence and the Birth of an Ethnolect* (Studies in English Language, pp. 125-148). Cambridge: Cambridge University Press.

Trudgill, Peter. 1986. Dialects in Contact. Oxford: Blackwell.

Upton, Clive. 1995. Mixing and fudging in midland and southern dialects of England: The *cup* and *foot* vowels. In Studies in general and English phonetics. Essays in honour of Professor J.D. O'Connor, ed. J Windsor Lewis, 385–94. London: Routledge.

Wells, J. C. 1982. Accents of English. Cambridge: Cambridge University Press.

Wickham, Hadley. 2009. ggplot2: Elegant graphics for data analysis. Springer-Verlag: NYC

Watt, Dominic and Lesley Milroy. 1999. Variation in three Tyneside vowels: is this dialect levelling?', in: Foulkes, Paul and Docherty, Gerry (Eds.), *Urban Voices: Accent Studies in the British Isles*. Arnold, London, pp. 25–46.

Williams, Anne and Paul Kerswill. 1999. Dialect levelling: change and continuity in Milton Keynes, Reading and Hull, in: Foulkes, Paul and Docherty, Gerry. (Eds.), *Urban Voices: Accent Studies in the British Isles*. Arnold, London, pp. 141–162.

Authors' Address:

Danielle Turton

Department of Linguistics & English Language,

Lancaster University,

Bailrigg, Lancaster, LA1 4YW

United Kingdom

d.m.turton@lancaster.ac.uk

Maciej Baranowski

School of Arts, Languages and Cultures,

University of Manchester,

Manchester, M13 9PR

United Kingdom

maciej.baranowski@manchester.ac.uk

FOOTNOTES

---

[1] This research was funded by the UK Economics and Research Council (ESRC, grant ES/I009426/1, Maciej Baranowski PI). We would like to thank Ricardo Bermúdez-Otero, Dan McCarthy, Laurel MacKenzie, three anonymous reviewers, and audiences at UKLVC 11 Cardiff, NWAV 46 Madison, the 3rd Edinburgh Symposium in Historical Phonology, and the 8th Northern Englishes Workshop in Newcastle for their insightful comments and questions. A special thanks goes to Reviewer 2 who has provided additional examples in footnotes 3, 4 and 6, and suggested the analysis of dialect data in Section 2.2, as well as suggesting some of the articulatory/phonetic factors in Section 4.3.2. All errors are our own.

[2] The use of the vowel phone in cardinal position 14 is conventional rather than indicating that speakers with the split have a backed vowel for STRUT. The position indicated in Figure 1 is typical of speakers in the South of England. Lindsey (2019: 27) describes modern day RP as using a centralised [ɵ] for FOOT.

[3] The [ʊ] of *cushion* could be accounted for by the following lip rounding in [ʃ], or even in in analogy with *push*. An anonymous reviewer reports that they have heard speakers with the distinction do the same with *mush*, as well as unfamiliar /ʌl/ words such as *culvert* and *pulverise.*

[4] Evidence can also be found in variation in the pronunciation of *room* (the OED cites both /rʊm/ and /ruːm/ for RP) also still possible in *roof, tooth* in some regions and could occur with any closed syllable with Middle English /ōː/ e.g. the SED shows *hoof* and even *goose.*

[5] Scotland, as can be seen from Figure 4, does have a distinction between the FOOT-STRUT vowels. STRUT has [ʌ] and there is no distinction between FOOT and GOOSE which have [u] (Wells 1982: 133; Maguire 2009).

[6] It is worth noting, as Upton (1996: 20) explains, these forms may have always been present in some areas of the country. Upton differentiates those words which have a preceding labial and following /l/ (*pull, bull, wolf*) from those which have a preceding labial without following /l/ (*put, butcher, pudding, woman*), and notes that the latter set were always variable in the transition from Middle English [u] to [ʌ]. In the 18th century this became fixed on [ʊ], but Upton provides evidence for [ʌ] in this subset in small areas in the South and South West. He does, however, note that many modern forms of *butcher* and *pudding* with [ʌ] are likely hypercorrections in the South, as they would almost certainly be in the North. Kurath & McDavid (1961: 147) also note [ʌ] for *butcher* in "the folk speech of the southernmost counties of England (Kent to Dorset) as well as Norfolk", based on their observations from the Lowman survey.

[7] Of course, here we mean public school in the British sense, which is traditionally referred to a select number of private boarding schools, although today is refers more generally to mean any fee-paying private school.

[8] The now-famous sociolinguistic interview question, *'Have you ever been in a situation where you were in serious danger of getting killed? When you said to yourself "this is it?"'* is used by sociolinguists to elicit narratives of high emotion, which are known to give an abundance of vernacular features (Labov 1972).

[9] Note that the speakers in Evans and Iverson's informants are from Ashby-de-la-Zouch in the East Midlands, which is close to the reported isogloss of the FOOT-STRUT split. Whether speakers from further North would still show such a shift remains to be seen.

[10] The assignment of informants to either the working or the middle classes reflects the traditional division into blue-collar workers on the one hand and white-collar labour on the other. Unskilled working-class informants are assigned to lower working class, whereas skilled workers, such as plumbers and electricians, represent the upper working class. Occupations such as administrative assistants, secretaries, and small business owners represent the lower middle class; middle middle class includes occupations such as teachers, managers, and higher level administrators; the upper middle class includes accountants, company directors, university professors, etc. It is worth noting that speakers are assigned to a particular occupational level on the basis of their entire occupational history, not just the most recent occupation, and, in the case of children and teenagers, on the basis of their parents' socio-economic status.

[11] We also determined who these speakers were from an additional minimal pair question, "Do *book* and *spook* rhyme for you?"

[12] As explained, the reason we do not include the complex interaction model in the main discussion is because the coefficients of models with interactions reflect baseline categories only, a fact which does not seem to be well-understood in the field. For our purposes, that means that a model with an interaction for vowel, social class and age shows in the coefficients for e.g. preceding segment the results relevant only for FOOT vowels uttered by lower working class speakers of the mean age. Therefore, we find a simple model with no interactions is more relevant for the main discussion, and the interested reader is referred to Appendix A for the more complex model.

[13] We are keen to discover more fine-grained phonetic effects, such as through ultrasound tongue imaging, or perhaps by applying Principal Components Analysis to combinations of neighbouring segments. However, this is beyond the scope of this paper.

[14] It is not entirely clear from the interview that this speaker (a 59-year-old female) meant to say that these words sound different. This highlights the importance of the sociolinguistic interviewer being able to probe

these minimal pair test claims further. In fact, she does confirm that *crooks* and *crux* sound the same in the second minimal pair test.

[15] It is notable that only six tokens of *but* make it into our final dataset because the remainder are unstressed. We have noticed that some individuals, with a definite lack of a distinction or even fudged variants, do produce something fudged in this word only, particularly in stressed conjoining position.

[16] This is Thomas's (2019:133) example but note that *officer* is in the THOUGHT class for many American English speakers. The Oxford English Dictionary online and Kenyon & Knott (1944: 303) list both pronunciations as possible in this word.