

¹ Towards an updated dialect atlas of British English

² Short title: An updated dialect atlas of British English

³ Laurel MacKenzie¹, George Bailey² and Danielle Turton³

⁴ *New York University¹, University of York², Lancaster University³*

⁵ *Contact details of lead author:*

⁶ Laurel MacKenzie

⁷ Department of Linguistics

⁸ 10 Washington Place

⁹ New York, NY 10003

¹⁰ phone: (212) 998-7958

¹¹ email: laurel.mackenzie@nyu.edu

¹² **Towards an updated dialect atlas of British English***

¹³ **Abstract**

¹⁴ This paper presents the results of a survey of phonological, lexical, and morphosyntactic
¹⁵ variation in British English, based on over 14,000 responses. We map twelve variables using
¹⁶ geospatial ‘hotspot’ analysis. One of our aims is to document the patterning of under- and
¹⁷ unstudied variables. A second aim is to track changes in real time, which we do by
¹⁸ comparing our findings to those of the 1950s-era Survey of English Dialects (SED, Orton
¹⁹ 1962). We improve upon previous dialectological work by paying careful attention to the
²⁰ phonemic status of mergers and splits: in our contemporary data, we do this by asking
²¹ subjects if they have a phonemic contrast; in the SED data, we do this by superimposing the
²² isoglosses for individual phones. We find evidence for both stability and change; we
²³ document previously unverified patterns. Perhaps most importantly, we identify a number
²⁴ of directions for future research.

²⁵

Abstract word count: 146

*The authors would like to thank Marije van Hattum and all of our student data collectors at the University of Manchester, Newcastle University, the University of York, and Lancaster University; audiences at Methods in Dialectology XV, Manchester Forum in Linguistics 2021, and the Lancaster University Phonetics Lab; and an anonymous reviewer. Parts of this research were funded by Multilingual Manchester, The University of Manchester’s Social Responsibility in the Curriculum competition, and The University of Manchester’s Learning through Research small project funding scheme.

26 1 Introduction and background

27 The regional dialects of England and the British Isles present the most fruitful location for
28 studying regional variation in English given that “geographical differentiation of local accents
29 is densest in those places which have long been settled by English-speaking populations”
30 (Wells 1982:10). This paper presents the results of a British English dialect survey, with data
31 taken from a study spanning seven years and comprising 14,438 respondents and 37 linguistic
32 variables, demonstrating the status of regional dialect variation in Great Britain today.

33 Our study follows a long line of tradition by surveying respondents on their use of a range
34 of lexical, phonological and morphosyntactic variables (Cheshire, Edwards & Whittle 1993;
35 Maguire 2012; Orton 1962; Wieling, Upton & Thompson 2014). The most famous of our pre-
36 cursors is the Survey of English Dialects (henceforth SED; Orton 1962), in which fieldworkers
37 collected questionnaire data for 1,300+ items from non-mobile older rural males in 313 lo-
38 calities across England in the 1950s. As will be explained further in Section 3, we regularly
39 compare our findings to this traditional dataset collected 70 years ago. Other surveys have
40 been conducted since the SED, with a range of focuses from morphosyntactic (Cheshire et al.
41 1993) to lexical (Vaux & Jøhndal 2009) and phonological (Maguire 2009). This long history of
42 dialectology studies in England and the British Isles has more recently been updated by map-
43 ping studies which use modern technology, such as mobile phone apps (Britain, Blaxter &
44 Leemann 2020; Jansen, Robinson, Cahill, Leemann, Blaxter & Britain 2020; Kirkham, Turton
45 & Leemann 2020; Leemann, Kolly & Britain 2018), Twitter (Grieve, Montgomery, Nini, Mu-
46 rakami & Guo 2019) and machine-learning techniques (Strycharczuk, López-Ibáñez, Brown &
47 Leemann 2020).

48 The survey presented in the current paper elicits responses from throughout the linguis-
49 tic grammar: phonological variation (e.g. “Do foot and cut rhyme for you?”), morphosyntac-
50 tic variation (e.g. “Would people from your area use the sentence, ‘You was outside when it
51 happened?’”) and lexical variation (e.g. “What do you call the evening meal?”). In doing so,
52 we make a number of contributions to the literature on British English dialectology. First,
53 we draw connections between the patterns shown by variables at different levels of gram-
54 mar. Second, where phonology is concerned, our elicitation strategy differs from the methods
55 found in Orton (1962) and Leemann et al. (2018) by directly testing the phonemic status of a
56 particular pair of vowels, as opposed to a broad phonetic realisation (see Section 2 below for
57 further details). This gives us the benefit of being able to map where certain vowel distinctions
58 or mergers exist directly from an informant’s judged perception, rather than concluding such
59 from comparing phonetic transcriptions across different words. As Wells (1978) points out in
60 his somewhat critical review of the *Linguistic Atlas of England* (based on the Survey of English
61 Dialects data; Orton, Sanderson & Widdowson 1978), the SED’s original survey data took no

account of structuralist phonemics as we attempt to do here. That is, the SED fieldworkers did not obtain informant minimal pair judgments on whether, e.g., two words such as *foot* and *cut* rhymed. The more recent English Dialects App (Leemann et al. 2018) follows the SED in this sense. Thus, our study circumnavigates the issue raised by Wells (1978) through our use of minimal-pair-like tests.

Additionally, throughout the paper, we compare our results to those of the SED where possible, by superimposing isoglosses from the *Linguistic Atlas of England* (henceforth *LAE*) onto our maps. In some cases, such as the presence of the construction *give it me*, this is straightforward. However, when attempting to map the areas which, for example, showed a NURSE-SQUARE merger (a sound change in England which means some areas pronounce *her* and *hair* the same; see Section 3.1.2), we can only develop an isogloss based on the SED data by comparing phonetic transcriptions across two *LAE* maps. Thus, we locate likely merged and distinct areas in the *LAE* by superimposing maps from different lexical sets on one another, revealing a possible isogloss for the 1950s data. As a result, this paper not only provides novel data; it also gives a brand new perspective on some old findings.

The aim of the present paper is to provide an initial exploration of the data we have been collecting over the past seven years, providing maps, discussion and cross-referencing across variables and a descriptive account of the current state of lexical, phonological and morphosyntactic variation in the UK. However, we are aware of the problem raised by Britain (2013:475) of “the portrayal, the display — sophisticated and eyecatching, admittedly — of data, rather than an explanation of the patterns found” (see also Trudgill 1974 on the issue of focussing on the results rather than the process). To avoid this, we also raise issues of theoretical interest alongside these descriptions, such as the mechanisms behind geographical diffusion of innovations, and related issues such as sociolinguistic factors, mobility and hierarchical effects (e.g. urban before rural). In future work, we intend to further build on this from a theoretical and social perspective, and narrow the focus to some of the areas of interest raised throughout this paper.

2 Methodology

2.1 The Our Dialects survey

The data for this study come from a survey of speakers of British English administered by undergraduate students of the authors between 2013 and 2019. The survey was modelled on those of Vaux & Golder (2003) and Labov, Ash & Boberg (2006). It targeted lexical, phonological, and grammatical variables, and led to the creation of an online dialect atlas for the general public (MacKenzie, Bailey & Turton 2016). In the early years of the survey, students collected the data by hand and inputted it electronically for analysis; beginning in 2015, survey

97 respondents were directed to an online form where they could submit their answers directly
98 (see MacKenzie 2018 for details). Respondents were targeted over social media and through
99 students' personal networks.

100 At the time of initial data collection, the survey was the largest and most recent survey of
101 phonological, lexical, and grammatical variation in British English. The data discussed here
102 comes from 14,438 respondents; over the course of the survey, 37 different variables were
103 targeted.¹ Most of the variables that were targeted were those that had been demonstrated to
104 display regional variation in earlier work, such as the FOOT-STRUT split (Hughes, Trudgill &
105 Watt 2012), the choice of *tea* or *dinner* for the evening meal (González 1993), and the pronom-
106 inal theme–goal ditransitive, as in *She gave it me* (Hughes et al. 2012). However, variables
107 were also included that students hypothesized might show regional variation despite a lack
108 of previous research.

109 An online survey naturally carries some disadvantages. We are forced to rely on speakers' intuitions, which may not be accurate (Labov, Karen & Miller 1991; Labov 1996). In the case of mergers-in-progress, for example, judgements are usually ahead of actual production (Herold 1990: 97; Labov 1994: 320, 355). Similarly, with near-mergers, speakers may have a difference in their production but not in their judgements (Labov 1994:359). Apart from issues related to speaker intuition, we cannot confirm that all respondents are native speakers of British English (though we explicitly asked this question on the survey, allowing us to filter out any respondents who reported that they weren't). And there is the possibility that we may receive spurious responses. However, the regional patterns we find for well-studied variables broadly match those found by other dialectological research (Cheshire et al. 1993; Hughes et al. 2012; Orton et al. 1978). This lends confidence to our method, and is consistent with other research in British English dialectology, which has found a high degree of comparability between different types of data, such as surveys and social media (Grieve et al. 2019).

122 An online survey additionally cannot capture information on speaker demographics in a particularly detailed way. Among other demographic information, our survey elicited respondents' sex (operationalized as a binary choice between "female" and "male" with a third "prefer not to answer" response), age, and regional background. This latter question asked speakers for the first half of the postcode of the place where they lived for the longest time growing up (between the ages of 4 and 13). Under the UK postcode system, the first half of a postcode comprises a speaker's postcode area (of which there are 121 in use in the UK) and their district within the area (of which there are typically around 20 per area). As we will detail in Section 2.2, we have mapped responses to our survey using this postcode district information.

¹All 14,438 respondents did not respond to all 37 variables; variables were added and removed from the survey over time, though a few core variables did remain on the survey for all seven years of its run. Precise numbers of respondents for each variable are given below in Table 2 and reiterated throughout the paper.

Responses to our survey skewed female (58% of responses) and young (ages range from 10–96 with median 22). Additionally, 39% of our sample (5,579 respondents) were self-declared students, according to a question on the survey about respondents' occupation. Though most survey runs did not ask respondents about current place of residence (only about where they grew up), we know that a large number of British university students choose to leave home to study: Whyte (2019) puts the figure at just over 80% for the academic year 2017–18. This means that a large proportion of our sample is likely to have experienced some mobility. And mobility is well-known to influence linguistic patterns: for some specific examples in the university context, see Evans & Iverson (2007), Prichard & Tamminga (2012), and Wagner (2012), or see Nycz (2015) for a recent review. An upshot of this is that our sample has a considerably different social profile than that of the SED, meaning that differences between our findings and theirs are almost inevitable. Though we cannot directly investigate the effects of geographical mobility in our current data set, this is a direction for future work. Asking explicitly about place of current residence in addition to place of childhood residence will allow us to compare respondents of matched geographical origins who differ in mobility. This can speak to whether some variables are more likely than others to change in situations of dialect contact (e.g. Chambers 1992).

While respondents to our survey covered a wide range of the UK, they disproportionately gave Northern England — where the authors were located during the survey's run — as the place where they grew up. A full breakdown of response numbers by region is given in Table 1; these regions are the official NUTS 1 (Nomenclature of Territorial Units for Statistics) subdivisions of the UK, including 9 statistical regions of England and the countries of Wales and Scotland. As we call upon these labels throughout the results section when describing the patterns of variation, Figure 1 maps these regional and national subdivisions for reference.

A full list of variables included in this paper is provided in Table 2 alongside the full wording of the survey questions and their possible answers, and the number of responses for each variable. For phonological variables, the questions ask about either homophony or rhyming between pairs of words; for lexical variables, participants were given a pre-determined set of possible variants and asked to choose the one they use 'most often'. Multiple selections were possible for participants who use more than one variant in equal measure, enabling respondents who use, say, a regional form and a standard form in different contexts to select both. By providing pre-determined responses, we depart from many earlier dialect surveys, including the SED, which asked open-ended or fill-in-the-blank questions. While offering a set of responses may artificially steer respondents toward an answer they wouldn't have otherwise provided (Tillery & Bailey 2003), we determined that this potential drawback was preferable to the labor that would have been required to process fill-in-the-blank answers. The survey also provided an "Other" box where respondents could write in additional options, and we re-

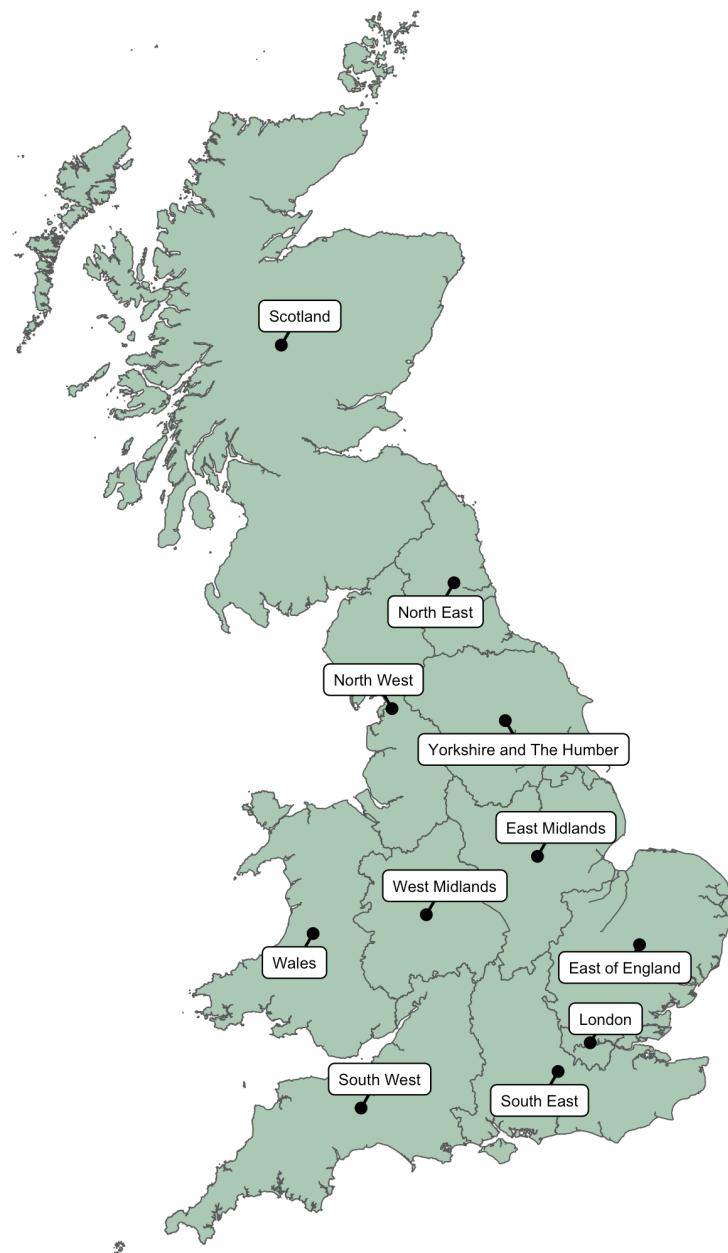


Figure 1: The official regions of England, alongside Scotland and Wales.

region	N	population	proportion sampled
North East	2098	2669941	0.079%
North West	4162	7341196	0.057%
Yorkshire and the Humber	1944	5502967	0.035%
East Midlands	1084	4835928	0.022%
West Midlands	791	5934037	0.013%
East of England	850	6236072	0.014%
London	956	8961989	0.011%
South East	1159	9180135	0.013%
South West	700	5624696	0.012%
Wales	314	3152879	0.010%
Scotland	380	5463300	0.007%

Table 1: Number of survey responses by region and country. Population data taken from the Office for National Statistics (2020).

168 fer to commonly-provided “other” variants throughout the paper where relevant (e.g. Section
 169 3.2.1).

170 For most grammatical variables, participants were asked to rate a given construction on
 171 the following five-point scale: (a) *I'd say this myself*; (b) *I wouldn't use it, but some people from*
 172 *my area do*; (c) *I've heard some people use this form*; (d) *A speaker of English might say this, but I*
 173 *haven't really heard it*; (e) *No native speaker of English would say this*. This phrasing is adapted
 174 from that used for grammatical variables in the Telsur survey that formed the data base for
 175 the *Atlas of North American English* (Labov et al. 2006:29). The potential for mismatch between
 176 grammaticality judgments like these and actual use is well known: see, e.g. Labov (1996), or
 177 Cornips & Poletto (2005) for a discussion in the specific context of dialectology. A known
 178 concern is the interference of the standard language on participants' judgments: respondents
 179 may be likely to call a non-standard sentence ungrammatical when it is in fact grammatical
 180 in their variety but socially dispreferred. Our inclusion of option (b) helps to circumvent this,
 181 by giving respondents the opportunity to pinpoint a form to their region without having
 182 to admit to using it. That said, acceptability judgments can be unreliable for other reasons,
 183 too: they may reflect estimated frequency of usage rather than grammaticality, or they may
 184 reflect speakers' difficulty judging a sentence without plausible pragmatic context (Cornips &
 185 Poletto 2005). For these reasons, we encourage future researchers to triangulate our judgment-
 186 based results with findings from large bodies of spontaneously-produced speech/writing (e.g.
 187 Twitter). Where possible, we do this throughout the paper, and see it as a useful direction for

section	type	variable	N	wording	options
4.1.1	phon	FOOT-STRUT split	14438	Do the words <i>foot</i> and <i>cut</i> rhyme for you?	yes, no
4.1.2	phon	NURSE-SQUARE merger	14438	Do the words <i>fur</i> and <i>bear</i> rhyme for you?	yes, no
4.1.3	phon	<i>book</i> as GOOSE or FOOT	14438	Do the words <i>book</i> and <i>spook</i> rhyme for you?	yes, no
4.1.4	phon	velar nasal plus	14438	Do the words <i>singer</i> and <i>finger</i> rhyme for you?	yes, no
4.1.5	phon	NORTH-FORCE merger	14438	Do the words <i>for</i> and <i>more</i> rhyme for you?	yes, no
4.1.6	phon	CURE-FORCE merger	14438	Do the words <i>poor</i> and <i>pour</i> sound the same to you?	same, different
4.2.1	lex	bread roll	14438	What would you call the soft, round bread pictured below?	<i>barm(cake), tea cake, muffin, cob, batch, bap, bun, roll</i>
4.2.2	lex	frozen treat	1738	What would you call the frozen treat pictured below?	<i>ice lolly, lolly ice</i>
4.2.3	lex	evening meal	14438	What do you call the evening meal?	<i>dinner, supper, tea</i>
4.3.1	gram	2nd person pl. <i>yous(e)</i>	8916	How would you address a group of two or more people?	<i>you, you guys, yous(e), you lot</i>
4.3.2	gram	dative alternation	14438	Would people in your area use the sentence: <i>Robin said, "give it me"?</i>	five-point scale (see text)
				Would people in your area use the sentences: (i) <i>Sam said, "you was outside having a smoke"</i>	five-point scale (see text)
4.3.3	gram	<i>was</i> -levelling	11846	(ii) <i>George said, "and the beaches was superb"</i>	five-point scale (see text)
			5708	(iii) <i>Rose said, "they was all in competition with each other"</i>	five-point scale (see text)
			5708	(iv) <i>Joe said, "all of a sudden we was getting our payslips"</i>	five-point scale (see text)

Table 2: Details of the variables analysed in this paper, with survey question wording and number of responses.

188 the future as spontaneously-produced data sets continue to grow in size, making it easier to
 189 elicit low-frequency lexical and grammatical variables.

190 As indicated in Table 2, this paper presents the results of 12 variables from the complete
 191 set of 37. This selection was made based on three primary criteria:

- 192 1. variables that are under-reported (e.g. the NORTH-FORCE and CURE-FORCE mergers)
- 193 2. variables that are widely-reported but for which no robust sociolinguistic or dialectological
 194 data currently exists (e.g. terms used for a *bread roll* and the *evening meal*)
- 195 3. variables that appear to show different regional patterns in comparisons between this con-
 196 temporary data and earlier dialect surveys, indicating potential language change (e.g. the
 197 FOOT-STRUT split and velar nasal plus)

198 2.2 Mapping & quantitative analysis

199 As mentioned in the preceding section, participants were asked for their postcode district (e.g.
 200 *M45, BB3* etc.), which allows us to map responses on a particularly fine-grained level: there
 201 are over 2,800 postcode districts across England, Scotland and Wales, and on average each
 202 district covers an area of just over 30 square miles. Geospatial analysis was conducted in R
 203 using the *sf* (Pebesma 2018) and *rgdal* (Bivand, Keitt & Rowlingson 2019) packages, and maps
 204 were generated in R using the *ggplot2* package (Wickham 2016). Below, we provide a brief
 205 description of the workflow involved in identifying statistical ‘hotspots’ from the raw survey
 206 data and producing the smoothed dialect maps that appear in Section 3. Similar methods of

hotspot analysis have been used in earlier studies of regional patterns of phonetic (Grieve et al. 2013), morphological (Tammainga 2013), lexical (Grieve, Speelman & Geeraerts 2011), and syntactic (Bart, Glaser, Sibler & Weibel 2013; Wood 2019) variation.

For each question on the survey, we start by calculating the proportion of respondents in each postcode district who use a particular variant. For lexical variables, this includes respondents who use a form exclusively (e.g. those who just select *barm* for the ‘bread’ question) but also those who use it alongside other variants (e.g. those who select *barm* and *bap*). For grammatical variables, we include respondents who report either of the top two acceptability judgements (i.e. speakers who either directly report use of the form in question, or attest its use in their local area). From these raw values, we then perform hotspot detection — specifically Getis-Ord Gi^* local spatial autocorrelation (Ord & Getis 1995) — to identify clusters of locations in which a variant is particularly favoured or disfavoured. The advantage of such methods is that isolated instances of the use of a variable are smoothed over and underlying regional patterns are more easily identified. The end result is a z-score for each location, which quantifies the extent to which that location is surrounded by other locations with similar values: a positive z-score indicates an area in which the linguistic form is favoured, whereas a negative value indicates an area in which the form is disfavoured, and the further this value is from 0 the stronger this pattern is.

The number of neighbouring locations that are taken into account forms the basis of the k-nearest neighbours (k-NN) algorithm, where the value of k is decided upon by the researcher (Getis 2009). We generated maps using 5-NN, 10-NN, and 25-NN. In this paper we report the results of the latter: from our manual comparisons between the raw and smoothed maps, it became clear that an analysis involving fewer nearest neighbours was prone to erroneous hotspots in areas with very few responses, while higher values of k resulted in over-smoothing and the loss of fine-grained spatial patterns for more locally-restricted forms (see Grieve 2017 for a discussion of considerations in setting the nearest neighbour parameter). These neighbours are assigned weights, equal to the reciprocal distance between the geographical centroid of itself and the location in question. As a result, a location’s smoothed value is more strongly influenced by the neighbouring locations that are closest.

In addition to the postcode district datum on which the smoothing was calculated, each survey response is also tagged with higher-level geographic information such as county and region (see Table 1). We recognise that postcodes and local authorities are not socially meaningful units and can both span and divide relevant linguistic areas; nevertheless, we still make reference to these different levels of geographic sub-divisions in our descriptions of regional patterns (Section 3), though we return to this point in Section 4. Additionally, to better help readers localize the patterns we describe, the regions enumerated in Table 1 are indicated

²⁴³ on each map in faint gray outline and locations of particular interest have been labelled as
²⁴⁴ appropriate for each variable.

²⁴⁵ Finally, where available, we have superimposed isoglosses from the *Linguistic Atlas of Eng-*
²⁴⁶ *land* (Orton et al. 1978) onto our maps. In some cases, such as for the NURSE-SQUARE merger
²⁴⁷ (Section 3.1.2), this has required us to overlay isoglosses from two different LAE maps (one for
²⁴⁸ the NURSE vowel and one for the SQUARE vowel), to create a single set of isoglosses represent-
²⁴⁹ ing the presence or lack of the phonemic contrast. We explain the details of these procedures
²⁵⁰ in the appropriate sections.

²⁵¹ 3 Results

²⁵² In this section, we present the findings of our phonological (Section 3.1), lexical (Section 3.2),
²⁵³ and morphosyntactic (Section 3.3) maps. Where applicable, we compare our findings to those
²⁵⁴ of the SED, to look for the possibility of real-time change. We also flag up shared patterns
²⁵⁵ across different variables within our own data, allowing for the identification of isogloss bun-
²⁵⁶ dles.

²⁵⁷ Two common themes run through the results presented here. The first is the departure of
²⁵⁸ our findings from those of the SED. Though patterns are broadly similar between the two data
²⁵⁹ sources, the edges of many dialect regions have clearly shifted since that research was carried
²⁶⁰ out in the first half of the twentieth century (see, for instance, Section 3.1.1 on the FOOT-STRUT
²⁶¹ split, or Section 3.1.4 on velar nasal plus). The second theme in our results concerns the bound-
²⁶² aries between regions: for instance, how far westward do features associated with Manchester
²⁶³ extend; or, conversely, how far east do we find characteristic features of Liverpool? How do
²⁶⁴ the Midlands pattern with respect to variables that show a strong North/South divide? We
²⁶⁵ answer these questions throughout, and elaborate on the directions that they raise for future
²⁶⁶ research in Section 4.

²⁶⁷ While considering these results, we encourage the reader to bear in mind the specific
²⁶⁸ nature of our sample (skewing young and female, with a large proportion of Northerners
²⁶⁹ and students) and our questions (potentially biasing respondents toward local forms via the
²⁷⁰ power of suggestion, at least for lexical and grammatical variables). To some extent, these two
²⁷¹ biases may balance each other out: young, female, mobile respondents may be more likely to
²⁷² avoid local forms, but a questionnaire that explicitly offers local forms as choices may make
²⁷³ respondents more likely to choose them than they would have been otherwise. Another, more
²⁷⁴ concrete consequence of our sample skew is that certain regions of the country are often
²⁷⁵ represented by very little data, which can lead to the appearance of spurious hotspots in the
²⁷⁶ geospatial analysis. We make an effort to flag these up where they arise.

277 **3.1 Phonological variables**

278 **3.1.1 FOOT-STRUT split**

279 Around the middle of the 17th century, a phonemic split occurred that saw an unrounded
 280 /ʌ/ variant emerge primarily from Middle English short /u/. Although the split also involved
 281 a number of intermediate stages and sounds which complicates this simple description, the
 282 consequence is that today speakers in certain regions of the country produce different vowels
 283 in words such as *foot* [fʊt] and *cut* [kʌt] (Beal 2008; Wells 1982). This change – commonly
 284 referred to as the FOOT-STRUT split – never occurred in the North of England, which means
 285 that for northern speakers these words rhyme with each other.

286 Earlier dialectological studies established an isogloss for this variable that runs from the
 287 Severn estuary in the west of England to the Wash in the east, essentially dividing England into
 288 two halves (Orton et al. 1978; Upton & Widdowson 1996; Wells 1982). Aside from the regions
 289 around Herefordshire and Berwick-on-Tweed, where northerners exceptionally exhibit this
 290 phonemic split, all dialects of England north of this Wash–Severn line are said to have a five-
 291 term short vowel system in which FOOT and STRUT are produced with the same quality.

292 It should of course be noted, however, that the placement of isoglosses can oversimplify
 293 what is actually a relatively complex and interesting pattern of regional variation. This is
 294 most notable in the Midlands, which has been described as a transition zone with dialects
 295 that demonstrate variation between the two forms and an intermediate realisation of STRUT
 296 that approximates [ɛ] (Chambers & Trudgill 1998); this was noticed over a century ago by
 297 Ellis (1889) and was more recently explored by Britain (1991, 2001) in the Fens. There have
 298 also been reports in Cannock, Staffordshire of lexically-specific variation in which speakers
 299 have [ʊ] in *rubber* but [ʌ] in *butter* (Heath 1980), and hypercorrect use of /ʌ/ in FOOT words
 300 elsewhere in the Midlands (see Map Ph143 of *foot* in Orton et al. 1978).

301 While the nature of our data makes it impossible to investigate the exact phonetic realisa-
 302 tions of STRUT, there are some advantages to the methodologies employed here: as discussed
 303 earlier in Section 1, the targeted questioning of our survey provides a more reliable indicator
 304 of the presence/absence of a phonemic split and the phonological status of this FOOT-STRUT
 305 contrast relative to other surveys such as the SED and the English Dialects App (Leemann,
 306 Britain & Blaxter 2017; Leemann et al. 2018), which target only isolated phonemes and in do-
 307 ing so potentially overestimate the extent of the split, particularly in areas of the Midlands
 308 that are known to exhibit centralisation of these vowels (see e.g. Jansen & Braber 2020) and
 309 the afore-mentioned patterns of hypercorrection. That said, the isoglosses between the two
 310 present-day studies are very similar, but with Leemann et al. (2017) erring on the side of a
 311 distinction. We return to this point later in Section 4. In Figure 2 we map the distribution of

312 responses to the question “do the words *foot* and *cut* rhyme for you?”, where an affirmative
 313 response would indicate the absence of a phonemic split.

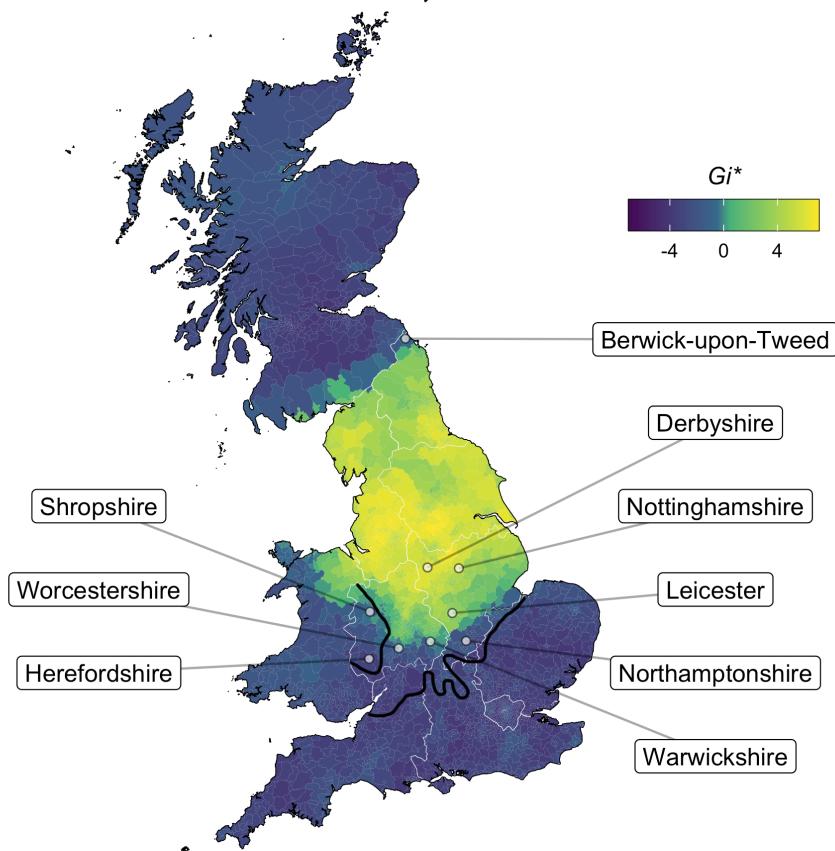


Figure 2: Do *foot* and *cut* rhyme for you? Light yellow areas represent the absence of a phonemic split. Black LAE isoglosses from Orton et al. (1978:Ph50) for the word *butter*.

314 The North–South divide in England is immediately apparent, with 79% ($N=8204$) of speak-
 315 ers across the North West, North East and Yorkshire reporting the same vowel for the FOOT
 316 and STRUT lexical sets, relative to just 5% ($N=2815$) in the south of the country. The ‘transi-
 317 tional’ midland zone is also reflected here, with rates of 63% ($N=1084$) in the East Midlands
 318 and 47% ($N=791$) in the West Midlands.

319 While the data in Scotland is too sparse to look for fine-grained regional patterns, it is
 320 unsurprising to find that an overwhelming 97% ($N=380$) of Scottish speakers exhibit a FOOT–
 321 STRUT split. This is also the case for 92% ($N=25$) of responses from Berwick-upon-Tweed,
 322 which – despite its position south of the Anglo-Scottish border – is known to be linguisti-
 323 cally aligned with Scotland in many regards (see e.g. Pichler 2008, 2010; Watt & Ingham 2000;
 324 Watt, Llamas & Johnson 2014). Wales is somewhat less homogenous with 78% ($N=314$) of re-
 325 spondents reporting a distinction, but this is largely due to a concentration of speakers in

326 North Wales who have resisted the split, possibly due to their proximity to Cheshire in the
327 North West of England.

328 The status of STRUT is arguably most interesting in the Midlands, with our map suggesting
329 a more northern boundary placement relative to the *LAE* isogloss, despite claims that this
330 traditional Severn–Wash boundary is “remarkably stable” (Wales 2006:104). Although most
331 parts of the East Midlands still demonstrate no obvious phonemic split (the words rhyme
332 for 79% of respondents in Derbyshire and 76% in Nottinghamshire, comparable with more
333 northern rates), some of the more southerly locales show very different behaviour: only 43%
334 (N=56) of speakers rhyme these words in the city of Leicester, and this drops even further to
335 just 7% (N=116) in Northamptonshire.

336 Moving on to the West Midlands, the exceptional behaviour of speakers in Hereford-
337 shire and the southern part of Shropshire — as noted before — is still evident. However, this
338 more contemporary data suggests that other parts of the West Midlands also show a strong
339 FOOT-STRUT distinction, contrary to the traditional boundaries put forward by the *LAE*: just
340 24% (N=75) of Warwickshire speakers and 31% (N=99) of Worcestershire speakers report the
341 same vowel in these words, and these are largely concentrated in the more northern parts of
342 the counties. Further research should shed light on this possible change, including both an
343 apparent-time analysis of this survey data as well as independent community-level studies in
344 the Midlands.

345 Setting aside this transitional zone, there is an interesting disparity between the two
346 ‘halves’ of the country when we consider those speakers who go against the regional pattern:
347 the South of England is incredibly homogenous with just 5% reporting a FOOT-STRUT rhyme,
348 whereas 21% of northern speakers are exceptional in reporting a phonemic split. This appar-
349 ent disparity may be partially explained with reference to social class and mobility. There are
350 claims in the literature that it becomes increasingly likely to find northerners with a FOOT-
351 STRUT split further up the social scale (Drummond 2012; Wells 1982); this also finds sup-
352 port from a recent large-scale quantitative study by Turton & Baranowski (2020), who report
353 widespread phonetic lowering of STRUT, and indeed evidence of complete phonological splits,
354 among many upper-middle class speakers in Manchester. Strycharczuk et al. (2019, 2020) also
355 find evidence of speakers in the North of England producing different vowels in these two sets,
356 and partly attribute this to highly-mobile speakers adopting a pan-regional ‘General Northern
357 English’. These changes in population and sampling dynamics may go some way to explaining
358 the apparent change observed here, particularly given the highly conservative nature of the
359 SED with its focus on non-mobile, older rural male speakers (NORMs), and how this contrasts
360 with the largely student-dominated responses collected here.

361 **3.1.2 NURSE-SQUARE merger**

362 The merger of the NURSE and SQUARE lexical sets results in homophony between words such as
 363 *fur* and *fair*, *burr* and *bear*. It is sometimes called the *fur-bear* merger or the *her-hair* merger,
 364 and is typically associated with accents in Merseyside (Knowles 1973; Wells 1982:361; West
 365 2015; Watson & Clark 2013) and in various locations in Greater Manchester and Lancashire
 366 such as Bolton and Blackburn (Turton 2015). Although we cannot consider the phonetic qual-
 367 ity of the merged vowel with our survey methods, it is commonly noted that present-day Liv-
 368 erpool speakers merge to a fronter SQUARE-like [ɛ:] pronunciation, whereas Lancashire has a
 369 more NURSE-like [ɜ:] or [ə:] vowel (Barras 2006, 2015; Knowles 1978:84; Shorrocks 1999:205;
 370 West 2015). It is likely that this difference is connected to rhoticity: the Lancashire areas have
 371 rhoticity or residual rhoticity which may have a centralising effect on the choice of vowel.
 372 This is reported for other vowels in parts of Lancashire due to the retroflex residual rhotic /r/
 373 (Shorrocks 1990).

374 Less commonly, the NURSE-SQUARE merger is reported for various northern varieties on
 375 the east coast of England. This includes Hull (Suddaby 2017; Williams & Kerswill 1999:146)
 376 and further north in Middlesbrough (Llamas 2001), but ‘not north of the Teesside conurbation’
 377 (Beal 2008:125). The reason for the merger being less typically associated with these north-
 378 east areas could be because the phonetic realisation is intermediate between [ɛ:] and [ɜ:] and
 379 therefore less striking than what we find in the North West. The alternative explanation is
 380 simply that it is less common in the speech of locals, or is a more recent merger compared to
 381 the North West.

382 The results from the present investigation, which asked respondents “Do *fur* and *bear*
 383 rhyme for you?”, reveal that 11% of respondents overall exhibit the NURSE-SQUARE merger.
 384 This is mapped in Figure 3, where an affirmative response (mapped as light yellow) indicates
 385 that the speaker has the merger. The vast majority of merged speakers are in the North West
 386 region (28% merged, N=4162), followed by the Yorkshire and Humber region (8% merged,
 387 N=1944), and then the North East (5% merged, N=2098). However, these larger regions are
 388 not particularly useful in diagnosing the geographical centres of this merger. When we break
 389 the regions down into local authorities, we see a clearer picture: although the main effect is
 390 carried by Merseyside in the west (61% merged, N=477), the eastern towns are catching up
 391 (Hartlepool: 54% merged, N=44; both Hull and North East Lincolnshire: 46%, Ns are 44 and
 392 59 respectively; East Riding: 38%, N=171). At the smaller level of postcode area, the Wigan
 393 postcode area in the North West (which also includes St Helens and Skelmersdale) shows
 394 high rates of the merger (63% merged, N=205).

395 Figure 3 also includes a newly developed 1950s isogloss of the merger, created from the
 396 available LAE map data. This is based on the phonetic transcriptions of the words *mare* (Orton
 397 et al. 1978: Ph83) and *third* (Orton et al. 1978: Ph30), selecting out areas where the two words

398 are transcribed with the same phone. The North West area in our data maps very closely to the
 399 *LAE* isogloss. An additional area emerging from the *LAE* isogloss but which does not feature
 400 in our merged responses can be found in the East Midlands, edging slightly into the West
 401 Midlands, including areas in Lincolnshire and Leicester. Wells (1982:361) does mention some of
 402 these areas with respect to this merger, stating his impression that speakers in Leicestershire,
 403 the West Midlands and Lincolnshire may be variably merged. In our data, only the North
 404 East of Lincolnshire persists in merging (as noted earlier). Thus, it seems that, potentially, a
 405 once-variable merger has been stamped out in favour of the standard.

406 The main inconsistency between our findings and those of the *LAE* can be seen along the
 407 east coast. This area shows no evidence of a merger in the 1950s data, but as shown above,
 408 has some of the highest rates of the merger in our dataset. Earlier we hypothesised that the
 409 association of the *NURSE-SQUARE* merger with the North West may be an issue of salience
 410 rather than frequency: perhaps the central [ɜ:] vowel in *SQUARE* stands out more than the
 411 fronted [ɛ:] variant in *NURSE*. However, this somewhat dramatic emergence of the merger on
 412 the east coast over the past 60 years suggests instead that it is a newer sound change in these
 413 areas. Further support for this comes from the fact that the rates in the West are higher, as well
 414 as evidence that the *NURSE-SQUARE* merger is a change in progress led by young women in
 415 eastern areas such as Hull (Suddaby 2017). As Beal (2008) notes, this requires more research
 416 from both a sociolinguistic and dialectological perspective in these eastern towns to draw
 417 such comparisons with areas in and around Merseyside and the North West, which are well
 418 documented with respect to this merger (Barras 2006; Knowles 1973; Watson & Clark 2013;
 419 Wells 1982; West 2015).

420 3.1.3 ‘book’ as *GOOSE* or *FOOT*

421 The lexical incidence of -ook words is regionally variable in British English, with some regions
 422 retaining the historical long vowel [u] – which persists in *spook* – in words such as *book*, *cook*,
 423 *look*. This means a word like *book* would be pronounced [bu:k] and not [bʊk]. Thus, -ook words
 424 are in the *GOOSE* set for these speakers, not the *FOOT* set. This is said to still be the case in areas
 425 such as Tyneside, Stoke-on-Trent and Liverpool (Barras 2015:265; Beal 2008:122; Newbrook
 426 1999; Wells 1982:373). For some time, it has been described a “recessive” feature of Northern
 427 Englishes (Wells 1982:373), restricted to the speech of older informants in areas where it is
 428 now the minority variant, such as Derby (Docherty & Foulkes 1999) and Manchester (Turton
 429 & Baranowski 2020), whilst showing both social class and age effects in the Wirral (Newbrook
 430 1999). Scotland retains the traditional realisation, having no difference between *FOOT* and
 431 *GOOSE*. For the purpose of this investigation, it means we would expect speakers from these
 432 areas to answer “yes” to our particular survey question, “Do *book* and *spook* rhyme for you?”.

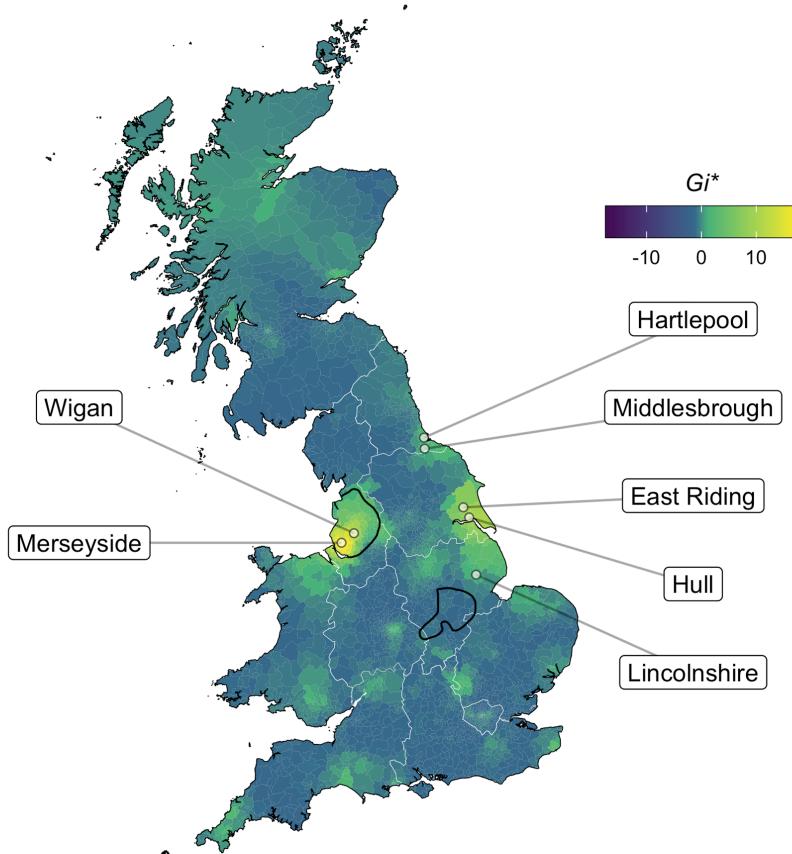


Figure 3: Do *fur* and *bear* rhyme for you? Light yellow areas represent merged responses. Black LAE isoglosses reflect areas with the same phone in *mare* (Orton et al. 1978:Ph83) and *third* (Orton et al. 1978:Ph30).

Figure 4 confirms that the areas listed above (the North East, Stoke-on-Trent and Liverpool pool) are still the representative heartlands of this traditional form, but the situation is much more stable in the North East when compared to areas in the west such as Merseyside and Stoke. The region of Tyne and Wear has the highest rates of the traditional realisation (85%, N=1200), followed by Northumberland (83%, N=206) and Stoke (77%, N=30). Compare this to Merseyside which is now just 25% (N=480). Some areas of Cumbria also pattern with the North East, showing a preference for -ook words being in the GOOSE set, although in most places the rates are more similar to present-day Merseyside (Cumbria overall is 20% 'yes', N=260). Various areas of Wales also report some of the highest rates of rhyme in *book* and *spook*, although overall numbers of responses are small. These include Anglesey and Gwynedd in the north, and Bridgend in the south.

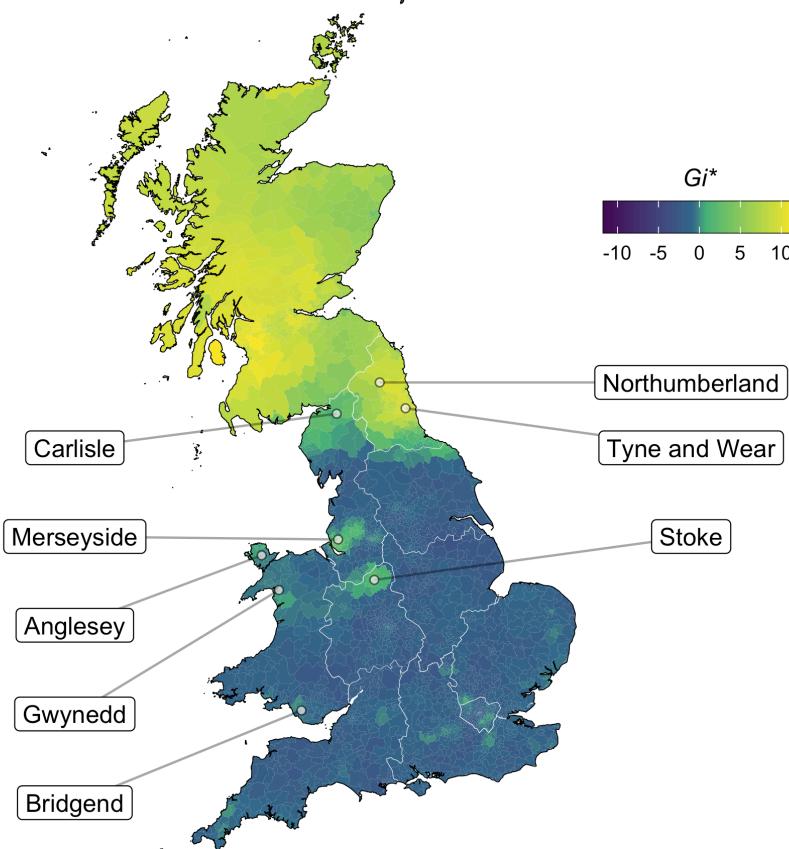


Figure 4: Do *book* and *spook* rhyme for you? Light yellow areas represent affirmative responses.

444 In Tyneside, the lexical incidence of this subset may be slightly different, with many speak-
 445 ers having FOOT in *book*, but GOOSE in other -ook words.² The incidence of the -ook words as
 446 [u:] seems to be productive, with reports of one Tynesider connected with an undergraduate at
 447 Newcastle University pronouncing *Brooklyn* as [b.ru:klin], although this report was not agreed
 448 on from all local speakers, demonstrating lexically specific realisations that vary within the
 449 speech community (see also Newbrook 1999:97).

450 In areas where *book* and *spook* rhyming is more variable than in, say, the stable North East,
 451 the traditional realisation functions as somewhat of a shibboleth. It is likely to be levelled in
 452 the coming years: evidence for this comes from the low rates in Merseyside today, but also
 453 Lancashire which has just 11% of reported rhyming of these words in our data. Stoke-on-Trent
 454 is the place to watch in the coming decades in order to observe the mechanisms by which this
 455 variable may change in future: Stoke has high rates of the traditional form, whilst also being
 456 geographically isolated in terms of [bu:k]-pronouncers.

457 3.1.4 *Velar nasal plus*

458 The *singer-finger* near-minimal-pair reflects a difference in *ng*-coalescence, specifically the
 459 variable presence of [g] following a velar nasal word-medially as in *singer* /sɪŋ(g)ə/ and word-
 460 finally as in *tongue* /tɒŋ(g)/. At a much earlier point in the history of English a [g] was in-
 461 variably present in these words regardless of the regional variety spoken, but around the start
 462 of the 17th century speakers began to simplify the nasal+stop cluster by dropping the [g]
 463 when it occurred either word-finally or word-medially before a morpheme boundary (Wells
 464 1982:188). However, there are many varieties of British English spoken largely in the North
 465 West of England in which this change never took place and speakers exhibit synchronic vari-
 466 ation between [ŋ]~[ŋg] to this day (attested in Heath 1980; Hughes et al. 2012; Knowles 1973;
 467 Schleef, Flynn & Ramsammy 2015; Wakelin 1984; Watts 2005; and explored in detail by Bailey
 468 2018). For these speakers, the words *finger* and *singer* may rhyme because the post-nasal /g/
 469 in *singer* is only variably deleted.

470 The exact geographical boundary of this [g]-retaining area has been described as “most of
 471 the western half of the midlands and middle north, including Birmingham, Coventry, Stoke-

²Whilst running vowel production and perception experiments on Tyneside vowels at Newcastle University, a number of students who had brought their mothers in to take part in an experiment reported a style-shifting effect in the opposite direction to what we might expect. These students reported that their mothers said *book* as [bu:k] in the experimental context, when in normal day-to-day life they would say [bʊk]. This is surprising because we would usually expect style-shifting in the direction of the standard, but often local speakers may not be aware of the direction of formality of a particular variant. The perception of the students was that their mothers were trying to sound “posh”. A similar effect was also found with intrusive-r in Tyneside by Foulkes (1997).

472 on-Trent, Manchester and Liverpool" (Wells 1982:365), as well as most of Derbyshire, the
 473 northern-most parts of Shropshire, Worcestershire and Warwickshire in the West Midlands,
 474 and the western-most part of Leicestershire in the East Midlands. It also creeps slightly into
 475 South Yorkshire, specifically Sheffield, and was attested in a very small part of the South East
 476 around Kent in the 1950s Survey of English Dialects, which until this point remained the most
 477 recent widespread study of this form's regional distribution.

478 Figure 5 maps the responses to the question "do the words *finger* and *singer* rhyme for
 479 you?", with the 1950s LAE isogloss superimposed over this new contemporary data. For the
 480 most part the regional spread of this form has remained relatively stable since the 1950s. The
 481 [g]-retaining areas are clearly centered around the North West (70% rhyme, N=4162) and the
 482 West Midlands (61%, N=791), and many of the aforementioned counties that lie on the border
 483 of the LAE isogloss still show relatively high rates of *singer-finger* rhyming today, such as
 484 Shropshire (62%, N=78) and Worcestershire (70%, N=108).

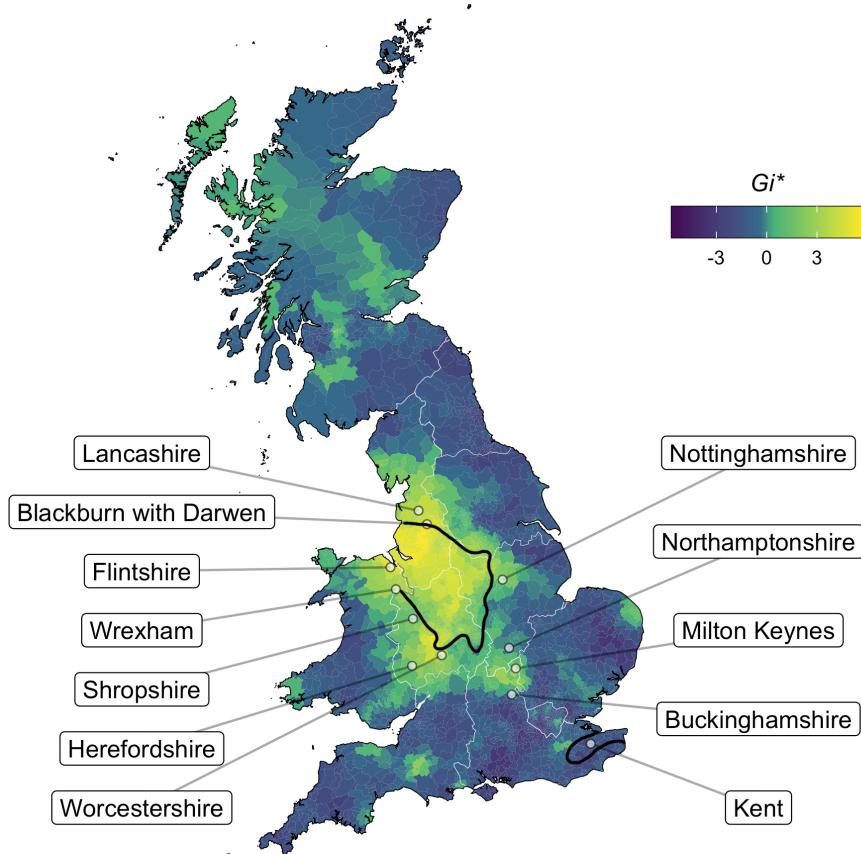


Figure 5: Do *singer* and *finger* rhyme for you? Light yellow areas represent the retention of post-nasal [g]. Black LAE isogloss from Orton et al. (1978:Ph242) for the word *tongue*.

485 There is even a suggestion that [ŋ] has spread beyond the southerly and northerly limits
 486 of the boundaries indicated in the LAE. The map seemingly illustrates a new hotspot appearing
 487 in the northern part of Buckinghamshire, and parts of Northamptonshire and Milton Keynes,
 488 but careful study of the raw data suggests that this is simply an artefact of the low response
 489 rate around this area. However, Herefordshire does seem to be a genuine (albeit weak) hotspot
 490 (50%, N=24) despite it lying completely outside of the older LAE isogloss.

491 Turning to the northern limit of this boundary, our contemporary dialect data indicates
 492 that post-nasal [g]-retention is prevalent throughout the county of Lancashire (68%, N=779),
 493 spreading further northward than the LAE isogloss with evidence of [ŋ] in Preston and the
 494 Ribble Valley (though does not progress as far north as Cumbria³, where the green-shaded
 495 regions simply reflect a handful of postcode districts in this area having a 100% rhyming rate
 496 based on a sole respondent). We also find evidence of a more eastern spread with [ŋ] attested
 497 in parts of Nottinghamshire, where 74% of those from the NG23–25 postcode districts report
 498 a rhyme (N=23).

499 It is also interesting to note that we find evidence of [g]-presence in North Wales (also
 500 noted by Wells 1982:390) although there is unfortunately no SED data with which we can
 501 draw comparisons. Though these patterns should be interpreted with caution due to a scarcity
 502 of data for large parts of Wales, a closer look at the raw data confirms the presence of [ŋ]
 503 in the Welsh counties of Flintshire (78%, N=27) and Wrexham (68%, N=28), adjacent to the
 504 Wales–England border and the English county of Cheshire.

505 The only evidence we find of retrenchment is in the South East of England, where the
 506 pocket of [ŋ]-users reported in the SED has all but vanished: only 26% of respondents from
 507 Kent now report a rhyme (N=182). While 26% may still seem somewhat high, there is likely a
 508 high false-positive rate in the responses to this question with survey participants incorrectly
 509 reporting a rhyme due to the subtle nature of this alternation between [ŋ]~[ŋg] and its con-
 510 tribution to the perception of rhyme in *singer-finger*⁴. For comparison, the rates of reported

³Note that Cumbria, along with Northumberland across to the North East, had /ŋ/ in morpheme-internal onset position in the LAE for the word *finger* (map Ph240; see also map Ph241 for *hungry*). This seems to have almost disappeared today but may remain in some lexical items. Macfadzean (2017) in his study of males in Caldbeck, Cumbria did indeed find that some older males retained the dialectal form /ŋ/ form in morpheme-internal onset position, but this was almost entirely restricted to the word *finger*, occurring 60% of the time. In the younger cohort, this had all but disappeared, with just one token of the traditional form arising. Thus, it is unlikely that we will have many speakers who operate in the opposite direction of what we have described in the rest of this section (i.e., who pronounce *singer* and *finger* to rhyme with the bare velar nasal) but it is something to be aware of in areas like Cumbria.

⁴There is independent evidence to suggest that there is a very low level of sociolinguistic awareness of this feature, at least among northerners (Bailey 2019a). It is of course possible that some of these responses are also from speakers who *do* genuinely rhyme these words but who actually have /ŋ/ in both rather than /ŋg/.

511 rhyming are similar in other regions where we have no reason to believe speakers retain [g]
 512 and where no obvious hotspot emerges, e.g. East of England (31%, N=850) and the North East
 513 (26%, N=2098).

514 It is interesting that these results point more towards [ŋg] spreading rather than retreat-
 515 ing, at least when compared with earlier survey data. As discussed elsewhere in this paper,
 516 these comparisons should be interpreted with some degree of caution due to the differences
 517 in population sample demographics. However, this finding does complement the results from
 518 independent work conducted in Greater Manchester and Lancashire, where the rate of post-
 519 nasal [g]-presence is in fact increasing in apparent time (Bailey 2019b). This might suggest that
 520 the [ŋg] pronunciation is becoming more widespread both probabilistically and spatially, but
 521 further work needs to be conducted targeting these peripheral communities to assess the ex-
 522 tent to which these survey results indicate diachronic change in the regional distribution of
 523 this form.

524 **3.1.5 NORTH-FORCE merger**

525 The seldom reported NORTH-FORCE distinction is a residual distinction left in very few parts
 526 of the English-speaking world, resulting in a difference between pairs such as *for, four; war,*
 527 *wore* and near-pairs such as *sort, sport*. The merger completed in Received Pronunciation in
 528 the 20th century after previously diphthongal FORCE shifted from [ɔə] to [ɔ:] (Wells 1982:235).
 529 A similar process is now happening to CURE (see Section 3.1.6). As noted by Labov (1994:316),
 530 the NORTH-FORCE distinction, where it remains, is not easily deduced from the spelling and
 531 thus likely must be learned in acquisition. For speakers who maintain a distinction, FORCE has
 532 the vowel [ɔ:], and NORTH is lower than FORCE, approximating a low-back [ɒ:].

533 Although Scottish English is said to have not undergone this merger (Wells 1982:408),
 534 there are few reports of exactly where this distinction remains in England. Labov (1994:315)
 535 reports that it remains in r-pronouncing dialects in the North of England. This is not true for
 536 Blackburn in Lancashire, but may refer to areas like Rochdale which is claimed to have had
 537 older rhotic speakers until relatively recently (Wells 1982). The merger is nearly complete in
 538 North America, although this seems to be fairly recent in some regions. Kurath & McDavid
 539 (1961:121) discuss the “extensive preservation” of the distinction in the Eastern states, but
 540 note that the distinction has disappeared in New York as it has done in London (but not the
 541 “folk dialects of England”). By the 1990s, however, the distinction seems to have rapidly all
 542 but disappeared, with Labov et al. (2006) reporting the distinction only among a few speakers

However, as described in the previous footnote, it is exceedingly rare for speakers to have /ŋ/ in *finger* and other morpheme-internal onsets now.

543 in Eastern New England, Southern Illinois, Indiana and the Gulf States. The two phonemes
 544 are still distinct in many areas of Ireland (Wells 1982:421).

545 Figure 6, which maps responses to the question “Do *for* and *more* rhyme for you?”, re-
 546 veals that there are areas of Britain today which retain a robust distinction. Manchester is
 547 one of them (and note that Manchester is not an r-pronouncing area). This distinction has
 548 been studied sociolinguistically in Manchester by Baranowski (2015), who notes that it is
 549 more common in working class speech, and also shows a rare “part of town” effect in that
 550 speakers from North and Central Manchester are more likely to have it than speakers from
 551 South Manchester. Our findings confirm this. Although overall 36% (N=1989) of people from
 552 the larger Greater Manchester metropolitan county are distinct, this effect is much stronger
 553 in North and East Manchester, and in the satellite towns to the North and East of the city:
 554 areas like Ashton, Bolton, Oldham, Rochdale and Wigan tend to have higher rates of distinct
 555 speakers. Around 16% of Warrington (which lies between Manchester and Liverpool, N=125)
 556 is distinct, which is the second highest area after Greater Manchester. This may be a good vari-
 557 able for delineating the Manchester-Liverpool divide, although we do have 10% of Merseyside
 558 speakers reporting a distinction (N=500).

559 In addition to these areas, the highest proportion of distinct responses in terms of post-
 560 code area in England is found in the TD area, which spans both England and Scotland on the
 561 eastern border (58%, N=26). Because our results also show a lack of merger in Scotland, this
 562 is potentially another feature in which the bordering areas of the North East patterns with
 563 Scotland (see also FOOT-STRUT, Section 3.1.1), although the NORTH-FORCE distinction is com-
 564 paratively more restricted, which may be expected given its disappearance in the rest of the
 565 English-speaking world.⁵

566 Figure 6 also includes a newly developed 1950s isogloss of the merger, created from the
 567 available LAE map data. This is based on the phonetic transcriptions of the words *forks* (rep-
 568 resenting NORTH) and *ford* (representing FORCE), selecting out areas where these words are
 569 transcribed with different phones. The LAE findings map fairly closely to the northernmost
 570 limit of our data, but the distinct area to the south of Manchester encompassed in the 1950s
 571 isogloss has since disappeared. There are two additional areas encompassed in our newly cre-
 572 ated 1950s isogloss: a section of the West Midlands on the Welsh border and an area running
 573 from the West Midlands to the north of Oxfordshire. Although there is some evidence that
 574 older speakers in the Black Country had a distinction fairly recently (Clark 2008:153), we find

⁵There are some lighter blue areas of Scotland in Figure 6 where a merger seems more likely. This may be due to our choice of words containing labials, as Wells (1982) reports that a merger can occur post-labially, i.e. *short*, *sport* do not rhyme, but *morning*, *mourning* may. Whether “labial” here includes labio-dental, as in *for*, is not clear.

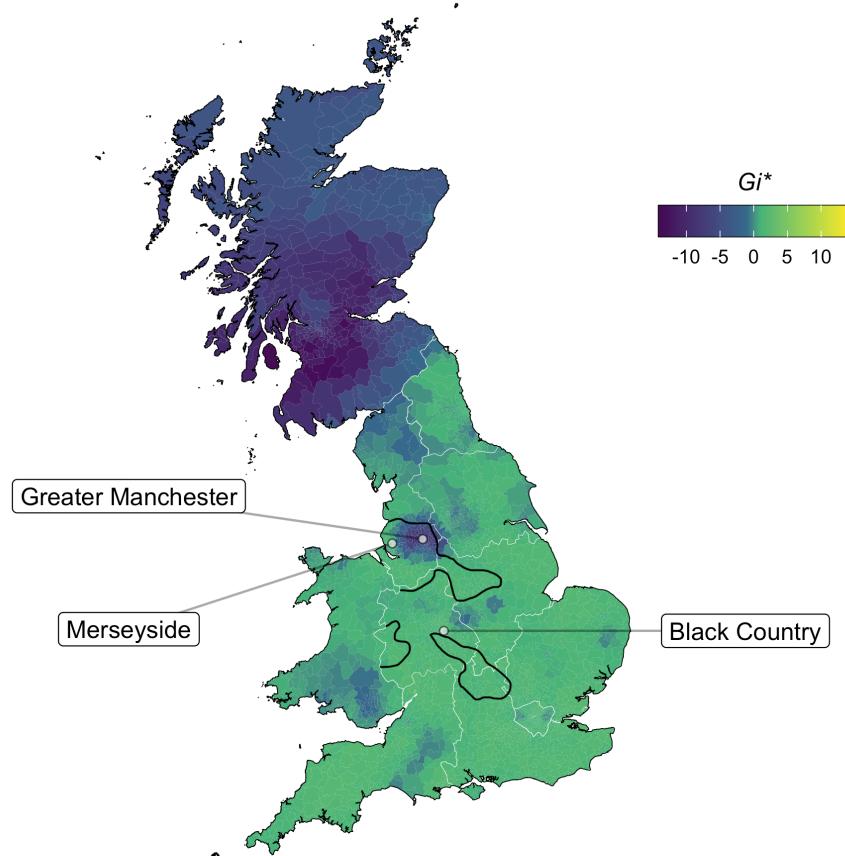


Figure 6: Do *for* and *more* rhyme for you? Dark blue areas represent negative responses, i.e. distinct vowels. Black LAE isoglosses reflect areas with different phones in *forks* (Orton et al. 1978:Ph47) and *ford* (Orton et al. 1978:Ph49).

575 little evidence of a remaining distinction in these areas for our speaker set.⁶ The apparent
 576 expansion of the distinction beyond the 1950s boundary slightly to the east of Manchester is
 577 likely merely because the SED did not survey anyone from this area. If Oldham, for example,
 578 had been included in the SED, the isogloss would be slightly further east.

579 Thus, it seems the progression of the NORTH-FORCE merger is well on its way in England,
 580 in line with Herzog's corollary to Garde's Principle: that mergers expand geographically at the
 581 expense of distinctions (Herzog 1965; Labov 1994, 2007). Further investigation of our data with
 582 reference to age patterns is a potential future avenue for research on this variable, although
 583 data from older speakers in key areas will be vital to assessing the death of this distinction.
 584 Sociolinguistic analyses, such as Baranowski (2015), are the key to understanding how such
 585 distinctions are lost within a speech community.

586 3.1.6 **FORCE-CURE merger**

587 The collapse of the FORCE-CURE distinction, labelled the second FORCE merger by Wells (1982)
 588 (the first being the collapse of NORTH-FORCE; see Section 3.1.5), is an ongoing sound change
 589 in present-day English which likely involves a merger by transfer (Labov 1994:321): mem-
 590 bers of the CURE set (some of which involve a preceding yod), which would traditionally be
 591 pronounced with [ʊə], move to the FORCE set and are pronounced with monophthongal [ɔ:].
 592 For many accents today, the loss of this final schwa offglide in diphthong CURE is complete,
 593 meaning that *poor*, traditionally [pʊə], is now realised in the same way as *pour* i.e. [pɔ:]. Thus
 594 these lexical sets are no longer distinct for many speakers (Hughes et al. 2012:50).

595 The CURE vowel exists in a relatively small number of words (e.g. *cure*, *tour*, *poor*) for
 596 speakers in areas which retain it, and it is variable in Received Pronunciation today (Lindsey
 597 2019). For some northern varieties, the lexical incidence of various words may be different
 598 from Received Pronunciation: e.g. in parts of Lancashire and Yorkshire, *door* can be heard
 599 as CURE rather than FORCE; see also Stoddart, Upton & Widdowson (1999:73). On the whole,
 600 it seems as though younger speakers have mostly lost this distinction, at least in England,
 601 although there are some regional exceptions such as the North East.

602 Figure 7 maps responses to the question "Do *pour* and *poor* sound the same to you?",
 603 where darker blue areas reflect the persistence of the FORCE-CURE distinction.⁷ Overall, our

⁶Although some areas do show darker colours indicative of a distinction, on closer inspection these numbers are small. We wonder whether our choice of words for this question was optimal, as some informants may have interpreted *for* as being realised with a reduced vowel (e.g. [ɒ] or [ə]), particularly when placed next to *more*.

⁷An anonymous reviewer suggests that *poor* is the most likely of the CURE words to use FORCE and thus our results may show an exaggerated effect of the merger. The reviewer also acknowledges that this may not be an accurate summary of all regions. In our experience, as linguists more familiar with Northern varieties, *poor* is one of the most robustly held CURE items. This mismatch of intuition between authors and reviewer could be

604 data show a 23% rate of retention of the distinction. The regional preferences for a distinction
 605 are found throughout the North East, which shows an overall figure of 77% distinct (N=2098),
 606 the highest region of all; we find as much as 94% distinct in some areas of Sunderland, Tee-
 607 side and Durham. Carlisle and the surrounding areas of Cumbria pattern with the North East,
 608 but are categorised as North West geographically, demonstrating that dialect contact and dif-
 609 fusion does not obey county lines. Yorkshire and the Humber is the second highest region
 610 retaining the distinction, but with a steep drop to 23% (N=1944), matching the overall average
 611 of our dataset. Major cities such as Leeds and Sheffield seem to be merged, but smaller places
 612 in between these larger urban areas retain a distinction, demonstrating that cities can show
 613 the effects of sound change first (Britain 2002b; Trudgill 1974). This is further demonstrated
 614 in Table 3, which shows the rates in Leeds and Sheffield alongside smaller towns in between:
 615 Rotherham and Barnsley are much further behind nearby Sheffield in terms of merging the
 616 sets. Leeds, the bigger city, is ahead with a mere 8% distinct, but this is also matched by nearby
 617 smaller Wakefield. Bradford, close to Leeds, shows a similar result of 8% distinct. This merger
 618 is ripe for further analysis of the demographic factors affecting networks in these areas, in-
 619 cluding population movement, transport routes, commuting, and sociolinguistic factors: why
 620 are the areas close to Leeds matching the big city's rates, but the areas which are a part of the
 621 Sheffield postcode area, Barnsley and Rotherham, remaining relatively stable? It is likely that
 622 the merger will show an effect of age, with younger speakers being more likely to be merged.
 623 That said, it is important to note that in areas such as the North East, younger speakers remain
 624 firmly distinct.

625 3.2 Lexical variables

626 3.2.1 Bread roll

627 The diversity of words for a small round bread in British English has been a topic of popular
 628 discussion since well before our survey. We elicited words for this item using a picture-naming
 629 task; the picture we asked respondents to name can be seen in the Appendix. Our survey gave
 630 respondents eight items to choose from: *barm*, *bap*, *batch*, *bun*, *cob*, *muffin*, *roll*, and *tea cake*,
 631 in addition to a write-in option.⁸

reflective of a North/South divide, where potentially Northern CURE is more strongly associated with stigmatised rural areas but Southern CURE with the more prestigious Conservative RP. Nevertheless, it highlights the role of the mechanisms behind a merger by transfer such as this, where all lexical items may not necessarily be affected at once (see also 3.1.3 on *book* as GOOSE or FOOT) and that our results might have turned out differently with the selection of a different lexical item.

⁸Commonly written-in responses include *barm cake* (grouped with *barm* for analysis) and *stottie*. In the North East, *stottie* refers to a specific type of (large, flat) bread item, different from the one pictured in our survey, so we omit it from our maps.

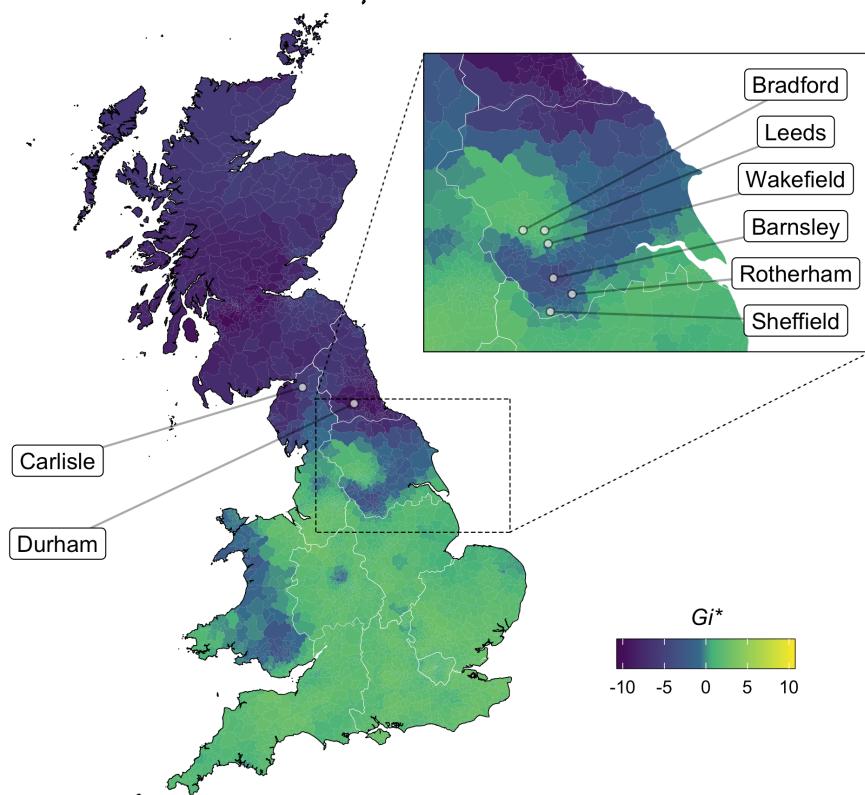


Figure 7: Do *pour* and *pour* sound the same to you? Dark blue areas represent negative responses, i.e. distinct vowels.

area	percent distinct	N
Leeds	8	319
Wakefield	9	23
Barnsley	63	57
Rotherham	38	37
Sheffield	19	212

Table 3: Major cities of Leeds and Sheffield with in-between towns showing the FORCE-CURE distinction (from north to south). Leeds postcodes were taken as LS1–20 and LS25–27; Wakefield, WF1–4; Barnsley, S70–75; Rotherham, S60–63, and Sheffield, S1–17, S20–26, S35–36, S94–99.

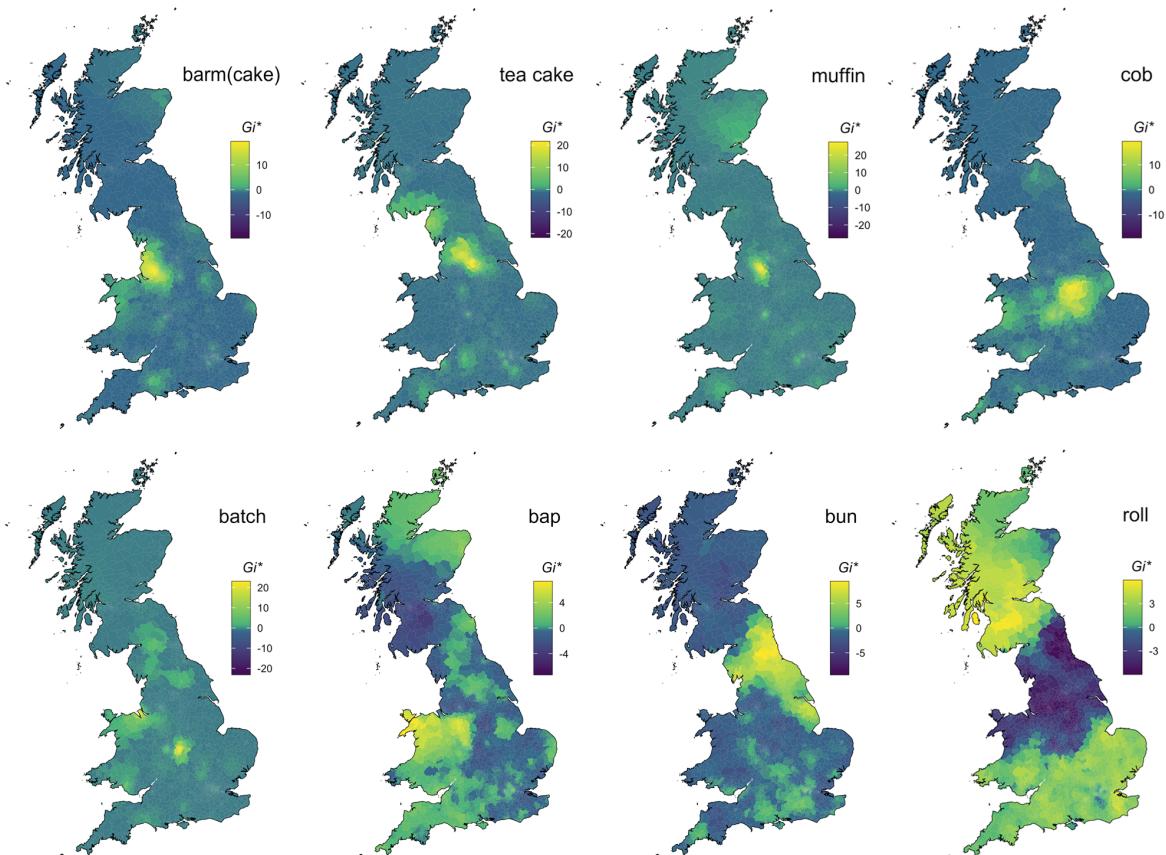


Figure 8: What is your word for a small round bread? Light yellow areas represent respondents who selected the indicated variant.

Figure 8 maps the eight variants provided by the survey. It shows that the terms for *bread roll* divide the country into a number of finely-demarcated divisions. *Barm* is confined to the North West, comprising an area that runs from Manchester westward to Liverpool and northward into the western half of Lancashire (from Blackpool to Preston). *Tea cake* spans the eastern half of Lancashire (Blackburn, Burnley) and the western half of West Yorkshire (Bradford and areas around Leeds). *Muffin* is perhaps the most geographically localized, confined to East Manchester and areas such as Oldham and Rochdale. *Cob* is largely concentrated in the Midlands around Nottinghamshire. *Batch* is used in two very small areas: Liverpool, in the North West, and Coventry, in the West Midlands. *Bap* is fairly widespread, but is most concentrated in Staffordshire, the West Midlands (Stoke-on-Trent, Birmingham), and North Wales. *Bun*, similarly, is fairly widespread, but maintains a stronghold in a broad area of the North East, extending from north of Newcastle down to northern Lincolnshire, tracing a diagonal line north of Leeds over to Cumbria. Finally, *roll* is apparently the normative choice, the most chosen variant and the one with the widest spread across the country, predominating in the South and in Scotland. The general picture is of considerable lexical diversity in the North and Midlands, and much more homogeneity in the South.

Some communities are fairly homogeneous in their choice of response. For instance, in the Nottingham postcode area, 268 out of 309 respondents (87%) selected a single option. Moreover, 174 of those 268 single-choice responses were *cob* (65%). By contrast, Birmingham shows a similar percentage of respondents choosing a single option (81% out of 258), but no variant shows a majority, with 41% *roll*, 20% *bap*, 17% *cob*, 14% *bun*, and the remaining 8% reflecting minority variants such as *batch* and *bread cake*. A fruitful direction for future research is to determine whether these responses pattern among social or geographical lines within the metropolitan area, or whether they might reflect the effects of mobility and dialect contact on our respondent population, given the ease with which new lexical variants can be acquired (Chambers 1992).

The regional divisions we find in the names for *bread roll* do not neatly align with the regions demarcated by phonological variants. For instance, while there is some overlap in the regions with the NURSE-SQUARE merger (Section 3.1.2) and the regions that say *barm*, the *barm* area goes farther east, to Manchester, while the NURSE-SQUARE merger definitively stops short of that city. Similarly, there is an area of overlap in East Manchester between speakers with the NORTH-FORCE distinction (Section 3.1.5) and those who say *muffin*, but the NORTH-FORCE distinction extends farther west, to Warrington, where 19% of respondents report a NORTH-FORCE distinction, but only 1% report using *muffin* (N=422). This kind of mismatching between phonological and lexical variants has been noted in other dialect surveys that consider variables at different levels of grammar (Labov et al. 2006).

668 **3.2.2 Ice lolly**

669 As is the case with the names for a small round bread (Section 3.2.1), the variation in whether
 670 a frozen confection on a stick (also known in American English by the generic trademark *pop-*
 671 *sicle*) is called an *ice lolly* or a *lolly ice* is the subject of considerable interest among laypeople
 672 but little attention by dialectologists. Lay discussions of the variation pinpoint the *lolly ice*
 673 variant to Liverpool (e.g. Anonymous 2012). As with the bread variable, we elicited words for
 674 this item using a picture-naming task with a set of pre-determined choices; the picture we
 675 asked respondents to name can be seen in the Appendix.

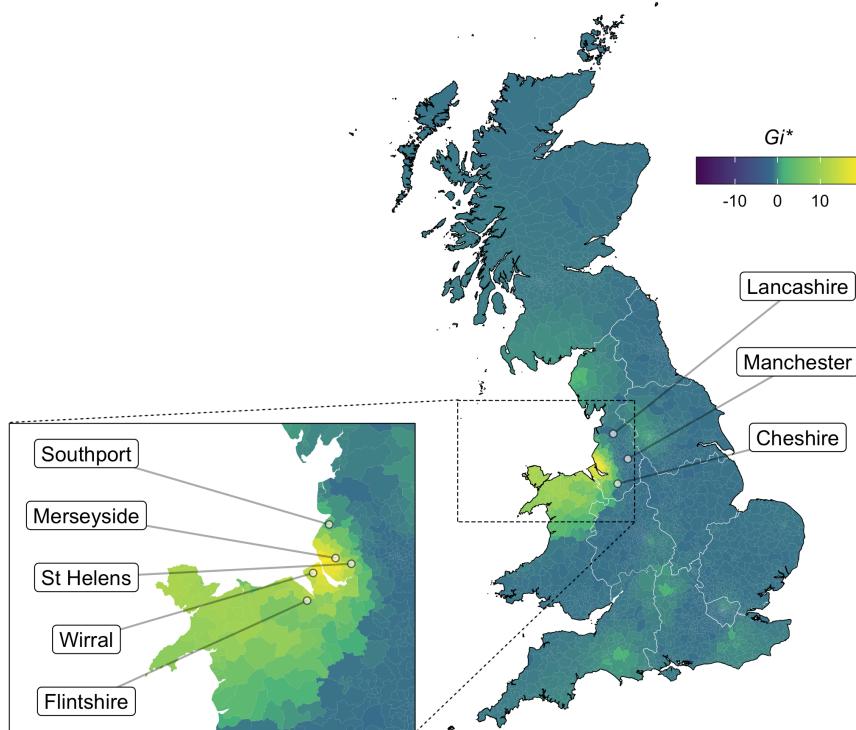


Figure 9: What is your word for a frozen treat on a stick? Light yellow areas represent respondents who selected the term *lolly ice*.

676 The first thing to note about Figure 9 is that it should be taken with caution, as this ques-
 677 tion was only added to our survey toward the end of our data collection period, so the response
 678 rate is much lower than for other variables ($N=1738$), and responses are not distributed evenly
 679 across the country. That said, we have a decent number of responses from Merseyside ($N=72$)
 680 and elsewhere in Northwest England (e.g. Lancashire, $N=144$), so we can draw some conclu-
 681 sions about that part of the country.

682 The lay perception that *lolly ice* is a Liverpool variant is entirely accurate. The form is re-
 683 markably localized to the Liverpool area. It extends along the Wirral peninsula to the south,

and eastward to St. Helens, but stops short of Greater Manchester. Its northern border is roughly Southport, still in Merseyside. Of our 72 respondents from Merseyside, 33 of them (46%) responded that they would use the term *lolly ice*; moreover, only two of those also identified *ice lolly* as a possible variant. This rate of *lolly ice* usage starkly contrasts with that of nearby regions in the North West: 10% *lolly ice* in Cheshire (N=63), 1% in Lancashire (N=144), and less than 1% in Manchester (N=394). *Lolly ice* clearly is a variant that is used only in the Liverpool area, and is used nearly exclusively among those who do use it.

We additionally find evidence for the use of *lolly ice* in North East Wales. The CH7 and CH8 postcode areas (both in the county of Flintshire, Wales) each show presence of *lolly ice* (CH8: 3 out of 4 respondents; CH7: 3 out of 7 respondents). Although the token counts are very low, instances of *lolly ice* usage in the rest of the country are so rare that they suggest this form to be a genuine variant in this county.⁹ In this respect, the western edge of *lolly ice* accords with that of velar nasal plus (Section 3.1.4), where a feature of North West England was also found to extend to Flintshire. This observation is generally consistent with sociophonetic studies of North East Wales (Morris 2013, 2017), and also studies of perceptual dialectology where non-linguists have labelled this area of Wales as ‘Scouse’ and sounding like Liverpool (Williams, Garrett & Coupland 1996).

3.2.3 Names for the evening meal

The terminology used in referring to the midday and evening meals, and the time at which the ‘main’ meal was eaten, was once strongly divided along socioeconomic lines: in the 18th and 19th centuries the wealthy upper classes ate their largest meal later in the evening, calling it *dinner* (or *supper* if the meal was more informal), and would have a lighter meal called *lunch(eon)* during the day. The working classes, on the other hand, would have *dinner* during the day and *high tea* in the evening as a source of sustenance after returning home from a long day of work (Bender 2009; Ayto 2012). Although class divisions had arguably weakened by the mid-twentieth century, Ross (1954:43) does list this variable when discussing British ‘sociolects’ and describes the use of *dinner* for the evening meal as a feature of ‘U-English’ (i.e. the variety spoken by the upper class).

More recently, these class divisions have further diminished and this variable has become a marker of regional varieties: the use of *tea* rather than *dinner* in referring to the evening meal is now considered a chiefly northern form (though this still may interact with social class, with middle-class northerners preferring *dinner* over the regional form), but the exact geographic perimeter of this difference is not yet known. This variable is particularly interesting, being a

⁹ Apparently high rates of *lolly ice* acceptance in western Wales should be disregarded; we have no data from this part of the country, and it is only coloured the way it is due to its proximity to Flintshire.

case of lexical variation in which confusion can arise due to cross-region polysemy: the same word (*dinner*) is used to mean different things depending on the variety of English spoken. In Figure 10 we map the distribution of respondents who indicated that they refer to the evening meal as *tea*.

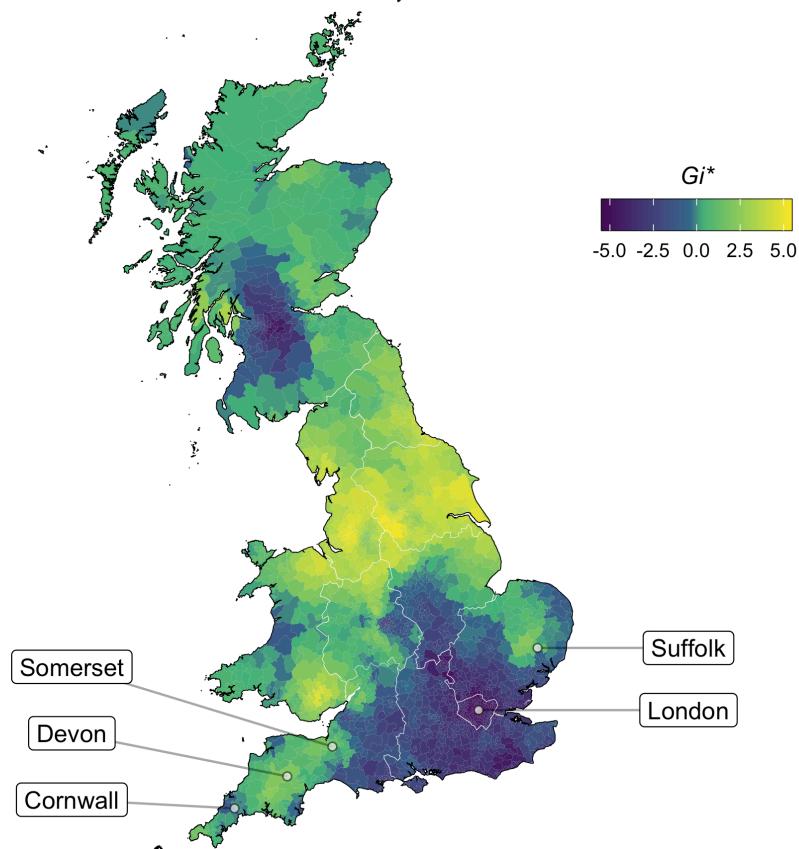


Figure 10: What is your word for the evening meal? Light yellow areas represent respondents who selected the term *tea*.

While a very clear pattern emerges between the North and South of England, this variable does not neatly divide the country into two halves in the same way that, for example, the FOOT-STRUT split does (as described in Section 3.1.1). Although *dinner* is still the preferred term throughout the South, there are areas where its use is far from categorical and where more localised hotspots emerge in which the use of *tea* is surprisingly high, e.g. Cornwall (where 45% select *tea* in their response, N=62), Devon (47%, N=75) and Somerset (47%, N=64) all in the South West, and Suffolk (43%, N=89) in East Anglia. As pointed out by an anonymous reviewer of this paper, it is interesting to note that the western parts of Norfolk and Suffolk, where use of *tea* is relatively high for the wider region, are also the ones least affected by counterurbanisation and rural gentrification. Coupled with the observation that there are dif-

731 ferences in this region between the centre of Cambridge and the northern edge of the wider
732 Cambridgeshire county, which are not connected with strong transport links, this points to
733 the importance of interpreting these results in the context of population dynamics and the
734 rural vs. urban distinction (a point made earlier in Section 3.1.6, on the FORCE–CURE merger).

735 It appears from Figure 10 that the most obvious contrast lies between the northern re-
736 gions (i.e. the North West, North East, and Yorkshire) and the South East, where the former
737 are *tea* strongholds and the latter *dinner*. However, it is of note that there is a much greater
738 level of homogeneity in the South East, where 84% (N=1159) use the favoured variant *dinner*,
739 and particularly in London, where that proportion rises to a near-categorical 95% (N=956).
740 Contrast this with the northern regions, where the dominant form *tea* is still only used by
741 67% of respondents in the North West (N=4161) and North East (N=2098), and by 69% of those
742 in Yorkshire (N=1944). The fact that more variation is found in the North may reflect some
743 residual class effect with northerners of higher socioeconomic status resisting the regional
744 form, similar to what we suggest for FOOT-STRUT in Section 3.1.1. There is in fact interesting
745 evidence of co-variation between these variables: of the northerners without a FOOT-STRUT
746 distinction, 25% (N=6462) report use of *dinner*, but this increases to 43% for northerners who
747 report a phonemic split in FOOT-STRUT (N=1742). Additionally, some survey participants re-
748 port using both forms, and provide qualitative comments revealing that the choice depends
749 on the size and type of meal, e.g. normally *tea*, but *dinner* if eaten in a restaurant.

750 3.3 Grammatical variables

751 3.3.1 Second person plural *yous(e)*

752 Standard English lacks a second person plural form, but many variants exist to fill that paradigm-
753 atic gap across regional and vernacular varieties (Wales 2004). Of these variants, our survey
754 investigated *yous* (also spelled *youse*). This second person plural form is found throughout
755 the English-speaking world, attested in American, Canadian, British, Irish, New Zealand, and
756 Australian Englishes (Bauer 2002; Clarke 2004; Hundt, Hay & Gordon 2004; Pawley 2004;
757 Quinn 2009; Wales 2004). Its considerable spread has been traced to a source in Irish English
758 (possibly calqued from Gaelic); it is localized to areas that experienced high volumes of Irish
759 immigration in the 19th century (Beal 2004; Filppula 2004; Wales 2004).

760 Within England, commonly-cited areas of *yous(e)* use are Liverpool and the North East
761 (Newcastle, Tyneside) (Beal 2004; Filppula 2004; Wales 2004); Beal additionally includes
762 “inner-city” Manchester on this list (p. 114). This latter inclusion accords with the dialect
763 survey results presented in Cheshire et al. (1993). Cheshire et al. find that all four survey sites
764 in the core of the Manchester metropolitan area reported the local occurrence of *yous(e)*, but
765 that only one out of nine sites in the rest of the Manchester metropolitan area, and one out of

766 four sites in the rest of the North West, reported use of the form. (The absence of Liverpool
 767 from Cheshire et al.'s survey sites likely explains the surprisingly low rate of *yous(e)* use in
 768 the general North West.) All of the British urban areas where *yous(e)* has been reported ex-
 769 perienced substantial Irish settlement in the nineteenth century; see Honeybone (2007) and
 770 references cited therein (fn. 2).

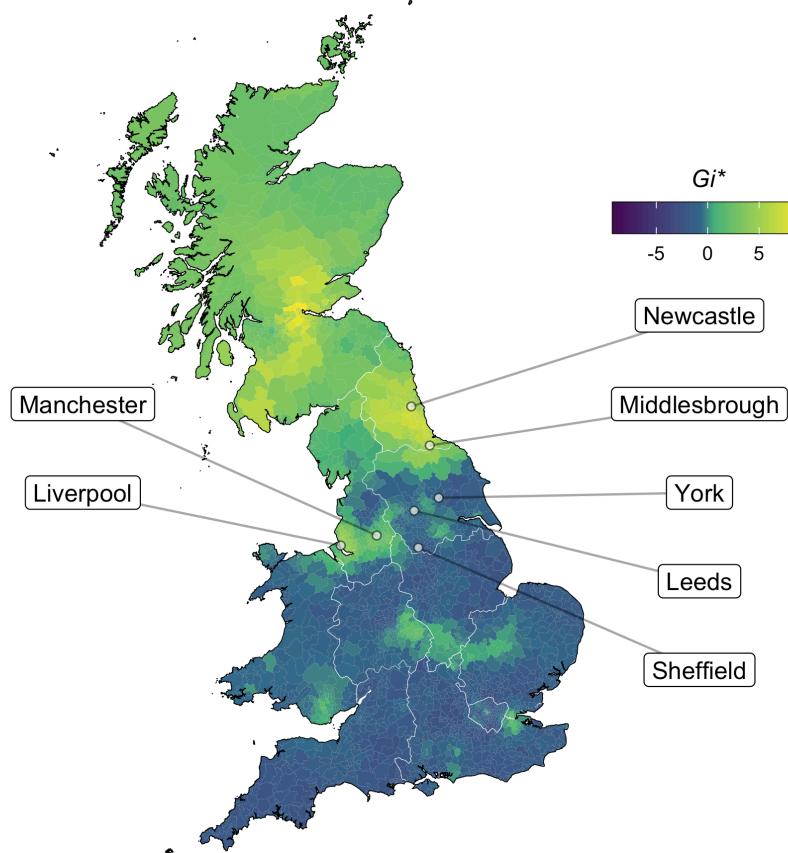


Figure 11: How would you address a group of two or more people? Light yellow areas represent respondents who selected the pronoun *yous*.

771 As shown in Figure 11, our survey results confirm a high rate of use of *yous(e)* in the
 772 North East. In the NE (Newcastle) postcode area, 51% of 1,105 respondents selected *yous* as
 773 an option, with usage continuing southward through Middlesbrough (TS postcode area: 44%,
 774 N=203). Rates are much lower in other Northern urban areas, indicating that *yous(e)* is not
 775 simply a pan-Northern phenomenon: compare York (14%, N=185), Leeds (11%, N=294), and
 776 Sheffield (7%, N=272).¹⁰ We additionally find a relatively high rate of *yous(e)* acceptance in
 777 Scotland (33%, N=263), consistent with previous findings (e.g. Filppula 2004).

¹⁰All counts are based on the postcode area for the respective city, namely YO, LS, and S.

778 Compared to the concentration of the form seen in Newcastle, *yous(e)* is weaker, but still
 779 prevalent, in a corridor of the North West extending from Liverpool (L postcode: 34%, N=164)
 780 to Manchester (M postcode: 25%, N=448). Here, however, *yous(e)* competes more strongly with
 781 alternatives such as *you guys* and *you lot*. The general picture is that when *yous(e)* is used in
 782 England, its utterer is almost certainly from either the North East or the North West, but that
 783 speakers from the North West use *yous(e)* less exclusively than those from the North East do.
 784 Still, our findings agree with those of previous research in that *yous(e)* tracks areas of heavy
 785 Irish settlement.¹¹

786 **3.3.2 Give it me**

787 Variability in the English ditransitive, or dative, construction has been the subject of much
 788 interest in the linguistic literature. Variation between what is called the full double-object
 789 construction (with two full noun phrase objects, as in *Dad read the baby a story*) and the full
 790 prepositional dative (with one full noun phrase object and one prepositional phrase, as in
 791 *Dad read a story to the baby*) is widespread across Englishes around the world (Szmrecsanyi,
 792 Grafmiller, Bresnan, Rosenbach, Tagliamonte & Todd 2017). In addition, there are regionally
 793 localized variants.

794 Our interest here is in a particular variant of the ditransitive when both non-subject argu-
 795 ments are pronouns. As in dit transitives with full noun phrases, speakers can allow a preposi-
 796 tional dative construction (e.g. *Dad read it to him*); while double-object constructions where
 797 the goal precedes the theme (as in *Dad read him it*) are uncommon (Szmrecsanyi, Grafmiller,
 798 Heller & Röthlisberger 2016), speakers in parts of Britain can allow an alternative double-
 799 object construction in which the theme precedes the goal (e.g. *Dad read it him*). It is this
 800 third variant that is the subject of our attention here; henceforth, we call it the “alternative
 801 double-object construction,” but it should be understood that we are referring only to that
 802 construction when both objects are pronominal.¹²

803 The Survey of English Dialects found that the alternative double-object construction is
 804 attested across much of the North West and Midlands, with small pockets of use in the extreme
 805 South West and South East as well. Research using present-day spoken corpora confirms its
 806 prevalence in the North West and Midlands (Gerwin 2013; Yáñez-Bouza & Denison 2015),

¹¹An anonymous reviewer also points out an apparent *yous(e)* hotspot east of London. This appears to be driven primarily by respondents from the DA (Dartford) and RM (Romford) postcodes (respective *yous(e)* rates: 14% of 21 respondents and 13% of 17 respondents). These rates do not approach what we see in the north of the country, and mentions of South East England as a *yous(e)* area are rare in the literature (though see Stenström 1997), but the history of heavy Irish settlement in East London (Walter 2010) suggests it as another possible site of transfer.

¹²The alternative double-object construction is also attested with full noun phrase objects – see Haddican (2010) and Biggs (2016) – but our survey didn’t target this.

as does research using Twitter data (Stevenson 2019). Stevenson's Twitter data additionally reveal fine-grained regional differences within the North West and Midlands in the actual rate at which the alternative double-object construction is used relative to the two other variants.

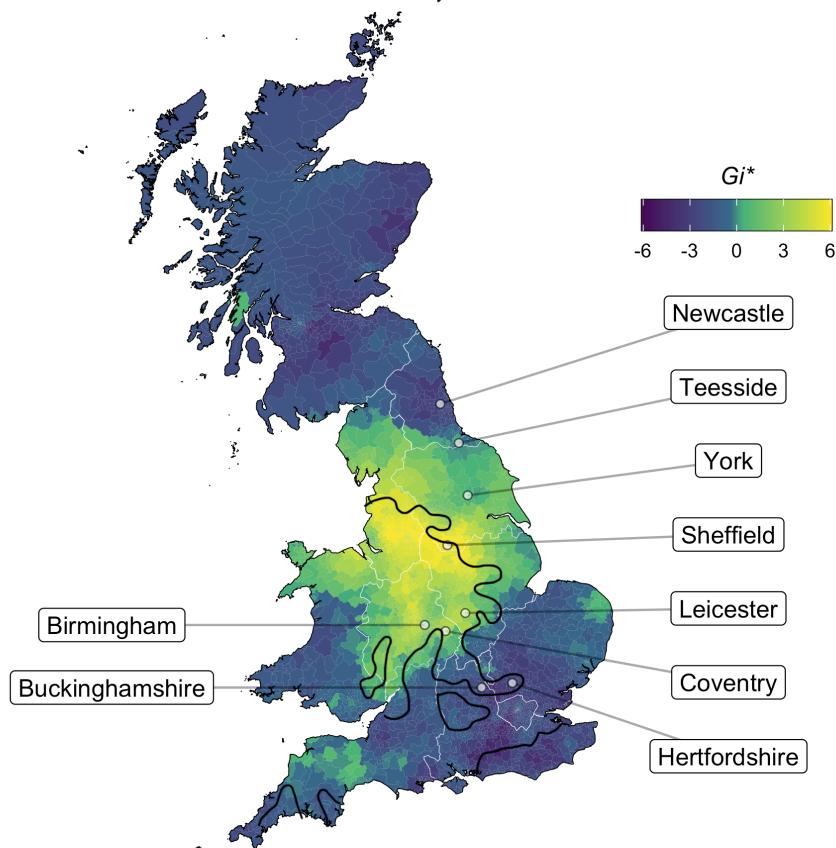


Figure 12: The acceptability of the alternative double-object construction with pronominal arguments. Light yellow areas represent respondents who said that either they or those in their area would use *give it me*. LAE isoglosses from Orton et al. (1978:S1).

The patterns in Figure 12 broadly agree with the *give it me* isoglosses from the *Linguistic Atlas of England* (superimposed in black), as well as the patterns found by Stevenson (2019) on Twitter, demonstrating consistency across different methodological approaches. Acceptance of the form predominates in the North West (80% acceptance, N=4162), continuing down through the West Midlands (70% acceptance, N=791) to the Severn Estuary, and into the East Midlands as well (72% acceptance, N=1084). The farther to the north east we go, the less acceptable *give it me* becomes: hence, we find 87% acceptance in Sheffield (N=497), 56% in York (N=256), 41% in Teesside (N=243), and 25% in Newcastle (N=1218). This is in direct agreement with the LAE, which also found *give it me* to be a North West and West Midlands form, with some spillover into the East Midlands. Though our map does show more acceptance of *give it*

820 *me* in Yorkshire than would be expected from the *LAE* isogloss, it is worth bearing in mind the
 821 methodological differences between the two projects: the SED elicited one preferred dative
 822 construction from each respondent, while our survey asked for acceptability judgments of
 823 the alternative double-object construction in particular. It's thus very possible that those SED
 824 respondents who generally accepted the construction would have shown a wider distribution
 825 than those for whom this construction was their primary variant.

826 That said, we do find a slight departure from the *LAE* in the precise location of the southern
 827 border of the *give it me* stronghold. We find the boundary of *give it me* acceptance to be firmly
 828 in the Midlands, just south of Birmingham (65% acceptance, N=258), Coventry (64%, N=100),
 829 and Leicester (71%, N=200). The *LAE* shows *give it me* usage to extend farther south than this,
 830 through Buckinghamshire into Hertfordshire, but our data find only 19% *give it me* acceptance
 831 in each of these counties (Buckinghamshire N=94, Hertfordshire N=219). The *LAE* additionally
 832 shows pockets of *give it me* use in the extreme South West and South East, which do not surface
 833 in our data.¹³ This suggests that there has been some attrition of *give it me* in the decades since
 834 the SED data was collected, and in this respect the *give it me* pattern is reminiscent of what
 835 we found for FOOT-STRUT, where our data also show that the southern boundary has shifted
 836 north compared to that presented in the *LAE* (Section 3.1.1). In fact, the southern boundary of
 837 *give it me* in our data is nearly identical to that of FOOT-STRUT, suggesting that the two might
 838 covary, and raising the questions of whether they have changed together over the course of
 839 the twentieth century, and whether similar social evaluation underlies each.

840 3.3.3 Was-levelling

841 There is considerable dialectological and sociolinguistic research on variation in the use of *was*
 842 and *were* in non-existential constructions. Rupp & Britain (2019:ch. 4) provide a comprehen-
 843 sive summary and synthesis of over a hundred different studies of this variation. Throughout
 844 the literature, three main patterns of variation arise:

- 845 1. “Was-levelling”, where *was* can be substituted for standard *were* in all contexts: e.g. *We was*
 846 *outside and she was outside. We wasn't inside and she wasn't inside.*
- 847 2. “Were-levelling”, where *were* can be substituted for standard *was* in all contexts: e.g. *We*
 848 *were outside and she were outside. We weren't inside and she weren't inside.*

¹³Jansen et al. (2020) similarly find use of *give it me* in the South East, with 4.9% of Sussex respondents to the English Dialects App claiming they use the form in preference to the two others (N=1254). In fact, we find 20% acceptance of *give it me* in Sussex (N=174). But compared to the very high rates of *give it me* acceptance that we find elsewhere in the country (70% or over in the lightest/yellowest regions on our map), this does not qualify Sussex as a *give it me* hotspot from our perspective.

849 3. A “mixed system”, where *was* is substituted for standard *were* in affirmative clauses, while
850 *weren’t* is substituted for standard *wasn’t* in negative clauses: e.g. *We was outside and she*
851 *was outside. We weren’t inside and she weren’t inside.*

852 (The fourth logical possibility, a mixed system with *were* in affirmative clauses and *wasn’t*
853 in negative clauses, is rare [Rupp & Britain 2019:176].)

854 Within the three main patterns, there are subtleties to the variation: it is sensitive to con-
855 textual factors such as subject type, and it shows social correlates within communities. Ad-
856 ditionally, the different patterns are not all equally attested throughout Britain, which is our
857 interest here.

858 Our survey asked only about the acceptability of sentences with regularized *was* (i.e. *was*
859 in place of standard English *were*) in affirmative clauses. This means we are unable to comment
860 on the regional distribution of *were*-levelling (pattern #2 above), or on whether levelled *was*
861 co-occurs in any region with levelled *weren’t* (pattern #3 above), as opposed to *was* being
862 levelled throughout the system, in negative as well as affirmative clauses (as in pattern #1).
863 However, we can still compare our patterns to those of previous research on the levelled *was*
864 pattern and on the mixed system, both of which regularize past *be* to *was* in affirmatives.

865 Our survey initially asked only about levelled *was* with the second person subject *you*.
866 Later instantiations of the survey contained questions with three more subjects: *we*, *they*, and
867 the plural noun phrase *the beaches*. As Rupp & Britain (2019:ch. 4) discuss at length, *was*-
868 levelling has not been attested in all four contexts equally.

869 Historically, dating back to Middle English, *was*-levelling was found in the North with
870 singular *you*, a pattern which stretched down into the Northern Midlands and has contin-
871 ued diffusing southward, such that *was*-levelling with *you* is now found as far south as Lon-
872 don. Additionally in the North, there has historically been evidence of *was*-levelling with
873 plural non-pronominal subjects (such as *the beaches*). Singular agreement with a plural non-
874 pronominal subject like this is reminiscent of what is known as the Northern Subject Rule, a
875 pattern under which plural non-pronominal subjects take third singular -s verbal marking in
876 the present indicative (e.g. de Haas & van Kemenade 2015). This Northern pattern of *was* with
877 plural non-pronominals has been observed in SED materials among speakers from the Central
878 North; more recently, variationist studies have found it in Buckie, Scotland; Newcastle; Read-
879 ing; and Inner London, demonstrating that it, too, has spread widely. Where *was* is levelled
880 with plural non-pronominals, it tends to be avoided with *they*; this is again a Northern Subject
881 Rule type of effect, by which there is different agreement patterning for non-pronominal ver-
882 sus pronominal plural subjects. Areas in East Anglia, by contrast, show the reverse pattern,
883 with more levelled *was* after *they* than after plural non-pronominals; Rupp & Britain (2019)
884 call this the “East Anglia Subject Rule,” but also suggest that it may be more broadly Southern.

885 Finally, recent studies have shown increased levelling of *was* with *we*, an environment where
 886 it was generally not attested historically.

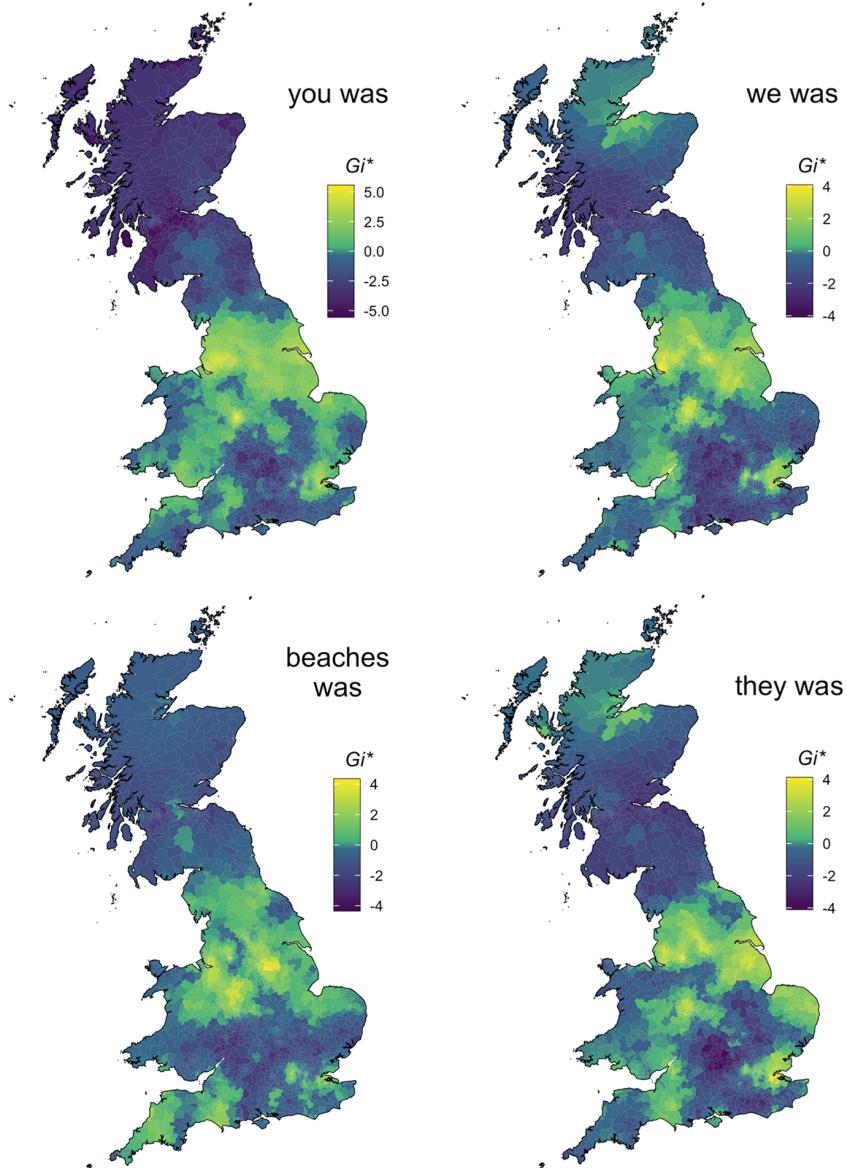


Figure 13: The acceptability of *was*-levelling with different subjects. Light yellow areas represent respondents who said that either they or those in their area would use levelled *was* with the indicated subject.

887 Figure 13 plots acceptance of the four different constructions under study. Before con-
 888 sidering the regional distribution of responses, it is informative to look at the variation in
 889 acceptability rates across the different subjects. Bearing in mind that we have much more
 890 data for *you was* than the other three, this construction does show the highest acceptance
 891 rate: 52% of 11,846 responses. This is consistent with the literature, which has found *you* to
 892 be the most common environment for *was*-regularization both diachronically and synchron-

893 ically. That is, levelling with *you* has been attested since Middle English; *you* is also the most
 894 common environment for levelled *was* in many communities, including York (Tagliamonte
 895 1998:180), Buckie (Smith 2000:66), the Fens (Britain 2002a:32), and London (Cheshire & Fox
 896 2009:21).

897 The remaining two pronouns show comparable amounts of levelling: 38% acceptance
 898 of *they was*, and 36% acceptance of *we was*. Finally, acceptance of levelling with the non-
 899 pronominal subject *the beaches* is at only 26%.¹⁴ Note that we only have 5,708 data points for
 900 these three constructions.

901 There are a few points of interest in Figure 13. First of all, we can see that regions where
 902 levelling is most accepted are the North West, Yorkshire and the Humber, the Midlands (both
 903 West and East), and London. To a lesser extent, we also see acceptance of levelling in the East
 904 of England and in the South West. By contrast, acceptance rates are comparatively low in the
 905 North East, the South East (outside of London), and Scotland. To some extent, this aligns with
 906 previous research: *was*-levelling has been historically attested in the (North) Midlands, and
 907 contemporary sociolinguistic work confirms its presence in London. However, other findings
 908 are surprising. Specifically, the high rates of levelling in the North West (54% acceptance over-
 909 all) and Yorkshire and the Humber (50% acceptance overall) contrast with studies of specific
 910 localities in these regions that uncover low rates of *was*-levelling: for instance, Moore (2011)
 911 finds extremely little evidence of *was*-levelling in Bolton, in the North West, and Tagliamonte
 912 (1998:161) finds only 6% *was*-levelling in York when her data is restricted to the four contexts
 913 we examined.¹⁵ Whether these differences are attributable to differences in methodology (both
 914 of the cited studies made use of conversational speech data) or in participant demographics
 915 (perhaps reflecting contact effects among our mobile student-biased sample) remains to be
 916 seen in future work.

917 To some extent, regional differences in *was*-levelling rates may be attributable to varia-
 918 tion elsewhere in the grammatical system. Most notably, Cheshire et al. (1993:72) suggest that
 919 rates of *was*-levelling with *you* may be low in regions where the second person plural pro-
 920 noun *yous(e)* is attested. This is because *you was* “is thought to have been used to restore the
 921 distinction of number in second person verb forms” — in other words, historically, *you was*
 922 was used when the referent was singular, and *you were* when the referent was plural (see also

¹⁴An anonymous reviewer raises the possibility that the overt plural marking on *beaches* may be lowering the leveling rate, and suggests that a non-overtly marked plural like *people* may lead to more leveling (see Walker (2020) for a recent consideration of this factor within English existentials). We hope to explore this in future work.

¹⁵This 6% rate was calculated based on the figures in Tagliamonte’s Table 3 for affirmative standard *were* contexts with *you*, *we*, *they*, and *NP*.

923 Rupp & Britain 2019:fn. 5). Varieties that had a unique second person plural pronoun did not
924 need to make this grammatical distinction.

925 Though our questionnaire did not specify the intended number of the *you* pronoun in our
926 example sentence, the most likely assumption is that respondents interpreted it as singular.
927 Thus, we can test whether *you was* is less common where *yous(e)* is prevalent. To some ex-
928 tent, this holds up. The two strongest regions of *yous(e)* usage, Newcastle and Teesside (see
929 Section 3.3.1), both show rates of *you was* acceptance that are significantly lower than the na-
930 tional average according to a chi-square test (Newcastle: 32% acceptance, N=1045, $p < 0.001$;
931 Teesside: 42% acceptance, N=177, $p=0.014$). Two weaker regions of *yous(e)* presence, though,
932 do not show the expected negative correlation: Liverpool and Manchester both show 60% *you*
933 *was* acceptance (Liverpool N=89, Manchester N=152), a non-significant difference from the
934 national average. This raises the possibility that *yous(e)* usage needs to have reached a cer-
935 tain threshold to block the emergence of *you was*, though to thoroughly test this theory, we'd
936 need to have access to *yous(e)* rates at the time when the *you was*/*you were* distinction was
937 still operative.

938 Finally, there are some clear generalizations to be drawn concerning subject hierarchies.
939 Out of the eleven regions studied here, ten of them show the highest rate of *was* acceptance
940 with *you*, and every region shows its lowest rate of *was* acceptance with *the beaches*. The
941 relative ordering of *we* and *they* varies by region, but both pronouns' rates are consistently
942 higher than that for *the beaches*. This means that no region shows more acceptance of *was* with
943 *the beaches* than with *they* – in other words, there is no evidence for the Northern Subject Rule
944 pattern in any region. The disappearance of the Northern Subject Rule pattern in Newcastle
945 has been noted by Beal (2004:122); our data would seem to suggest that it has spread even
946 further.

947 The findings presented here are intriguing in their departure from previous literature; we
948 hope other researchers can follow up on them with a larger set of data, ideally also comparing
949 them to results from conversational speech in particular communities.

950 4 Discussion

951 Throughout Section 3, we have identified a number of directions for future research. Here, we
952 summarize and elaborate on them.

953 First, we have found several apparent cases of change in progress which represent exciting
954 areas for further study. These include:

- 955 • The emergence of a FOOT-STRUT split in the Midlands (Section 3.1.1)
956 • The emergence of a NURSE-SQUARE merger running from North East England to the north
957 of the East Midlands (Section 3.1.2)

- The decrease in *book-spook* rhyming in areas such as Stoke (Section 3.1.3)
- The spread of velar nasal plus beyond the LAE boundaries (Section 3.1.4)
- The diffusion of the CURE-FORCE merger in Yorkshire (Section 3.1.6)
- The loss of a Northern Subject Rule-type pattern of *was*-leveling (Section 3.3.3)

Some of these changes have been confirmed in real time by comparing our results to isoglosses (composite where necessary) from the LAE. Where we do not have real-time data, confirmation that these are indeed cases of change can come from two sources: analysis of the apparent-time patterns in our own survey data (pending further data collection from older speakers) and dedicated follow-up studies of the communities in question. In the case of *was*-leveling, our data offer us a rare opportunity to study a case of constraint change in apparent time (MacKenzie 2019).

Though these particular variables are each interesting in their own right, further study of them as a group presents avenues for better understanding the transmission and diffusion of changes from above and below (Labov 1994:78, Labov 2007). Most of the phonological variables studied here are changes from above. The erosion of the traditional realisation of the *-ook* words, the merger of CURE and FORCE, the hypothesised spread of the FOOT-STRUT distinction and the eradication of the NORTH-FORCE distinction are all changes which are above the level of conscious awareness, originate outside the speech community, show style-shifting and originate in the highest social class (with the exception of CURE and FORCE in conservative Received Pronunciation). By contrast, one potential change in progress that we hypothesise is change from below – i.e., below the level of conscious awareness, from within the speech community, and originating in a centrally-located social class – is the NURSE-SQUARE merger on the east coast (e.g. Hartlepool, Hull; see Section 3.1.2). Consulting our data on respondent occupation (recently argued to still be the optimal measure of social class in the UK, Baranowski & Turton 2018) will shed light on this. It will additionally help to clarify whether the low social awareness of this merger in this community is due to its status as a change from below, or due to the low phonetic salience of the merged vowel. Experiments comparing speaker perceptions of this merger in Hull versus Liverpool will help us better understand the social differences that underlie them.

Another avenue from which to study changes in progress is in the nature of their geographical diffusion. We find tentative evidence that the CURE-FORCE merger is affecting larger towns before smaller ones. Controlling for differences in respondent age and social class between the Yorkshire towns in question will help confirm this pattern. Our data also reveal clear transitional zones between dialects, where towns may display considerable heterogeneity. These are particularly interesting in the Midlands, where several isoglosses coincide, and in the towns between Manchester and Liverpool, two major cities which both display characteristic features that are not found elsewhere (e.g. *lolly ice*, the NORTH-FORCE distinction).

995 These variables should serve as the basis for targeted sociolinguistic research which documents
996 their precise geographical spread, correlates them with communication patterns, and
997 assesses their social meanings.

998 On this subject, it is worth considering more carefully our finding of the northward spread
999 of the FOOT-STRUT isogloss. We conclude that the FOOT-STRUT phonemic split has spread
1000 northwards since the 1950s. This goes against Herzog's Principle that mergers expand geo-
1001 graphically at the expense of distinctions (Herzog 1965; Labov 1994), but the high social value
1002 of the prestigious distinction has previously been cited to explain the ability to overcome this
1003 linguistic tendency (Labov 1994; Turton & Baranowski 2020). Though our finding of the rais-
1004 ing of the FOOT-STRUT isogloss in some ways agrees with the findings from Leemann et al.'s
1005 (2017) dialect survey, we argue that the change may not be as vigorous as their results suggest,
1006 and that their methods overestimate the proportion of split speakers in the North of England,
1007 particularly those in the Midlands close to the isogloss. This is because their method is to ask
1008 a question about pronunciation with a forced-choice response. Speakers with no phonemic
1009 split, but with a schwa-like realisation for FOOT and STRUT, when faced with a choice between
1010 *butter* as pronounced with an RP-like [ʌ] as opposed to a Northern [ʊ], will select the RP-like
1011 option.¹⁶ Crucially, in the same way as the SED, Leemann et al. (2017, 2018) use the selection of
1012 the [ʌ]-like form as a proxy for presence of a distinction, which may result in an overestima-
1013 tion of the spread of the split. Indeed, recent results from Jansen & Braber (2020) in three East
1014 Midlands cities, which show that FOOT and STRUT are becoming more similar in their young
1015 East Midlands speakers, gives us added confidence in this interpretation of the state of the
1016 phonemic distinction today. In summary, these divergent interpretations on the presence of
1017 the FOOT-STRUT split in the Midlands highlight the problems with assuming phonemic status
1018 via survey methods which do not elicit structural properties, something originally highlighted
1019 by Wells (1978).

1020 Still another direction for further study is the covariation of variables that seem to be
1021 changing together in real time. For instance, when we compare our data to the LAE maps,
1022 we find that the area which lacks the FOOT-STRUT distinction and the area that uses *give*
1023 *it me* have both shrunk on their southern ends in very similar ways. This raises questions
1024 of whether similar social evaluation underlies the two variables, and whether they co-vary
1025 within individuals (Tamminga 2019). We do find some evidence of intra-speaker covariation
1026 among Northerners between the presence of the FOOT-STRUT split and the use of *dinner* for the
1027 evening meal, both characteristically southern features. Whether speakers who report both

¹⁶Incidentally, the Northern [ʊ] option in the English Dialects app (Leemann et al. 2017) is very high and rounded, and thus it may be likely that any non-distinct participant with even a remotely centralised FOOT-STRUT vowel would opt for the RP-like [ʌ] form.

1028 of these forms are socially similar (for instance, in their occupation level and/or their contact
1029 with Southern speakers) remains to be determined.

1030 Finally, by analysing which areas affiliate with one another linguistically, we speak to
1031 questions about how people communicate. We find evidence of dialect regions crossing county
1032 and even national boundaries: for instance, *-ook*, FOOT-STRUT, velar nasal plus, and *lolly ice*
1033 all show patterns by which North Wales affiliates with Liverpool or Northwest England more
1034 generally; FOOT-STRUT also shows linguistic alignment between Berwick-upon-Tweed and
1035 Scotland. Simple geographical proximity is not a guarantee of shared linguistic repertoire,
1036 however: Central Wales does not pattern like the West Midlands, its nearest English-speaking
1037 area. These findings suggest a role for commuting and communication patterns in uniting
1038 regions, as well as the influence of local identity (e.g. Duncan 2018; Llamas 2007). They also
1039 suggest that a more nuanced approach to geographically subdividing the country is preferable
1040 to working with postcodes and local authorities, which may span and divide relevant linguistic
1041 areas. A direction for future research is to apply machine learning classification techniques
1042 to our data, to identify which areas are more or less similar and which features play the most
1043 crucial role in dividing them (Strycharczuk et al. 2020).

1044 5 Conclusion

1045 This paper has analyzed the regional patterning of over 14,000 responses to twelve linguistic
1046 variables in England, Scotland, and Wales. We have additionally contributed a novel real-time
1047 perspective by comparing our findings for six variables to those obtained in the first half of
1048 the twentieth century. Bearing in mind the specific nature of our sample and questions, we
1049 find evidence for both stability and change; we document previously unverified patterns; and
1050 we identify a number of directions for future research.

1051 Though the linguistic landscape of Britain has been investigated in several previous large-
1052 scale dialectological studies already, this paper shows that there are still novel observations to
1053 be made. In fact, we see it as a boon for British dialectology that there are now several dialect
1054 mapping projects, because they each contribute a different perspective: for instance, Orton
1055 (1962) provide historical data; Leemann et al. (2018) contribute phonetic data; Grieve et al.
1056 (2019) make use of spontaneous language in context from social media. We hope that future
1057 research will continue to compare findings obtained through different methods, in order to
1058 shed light on all the complex nuances of English as it is spoken throughout Britain.

1059 **Appendix**



Figure 14: Visual prompt for the ‘bread roll’ question on the survey.

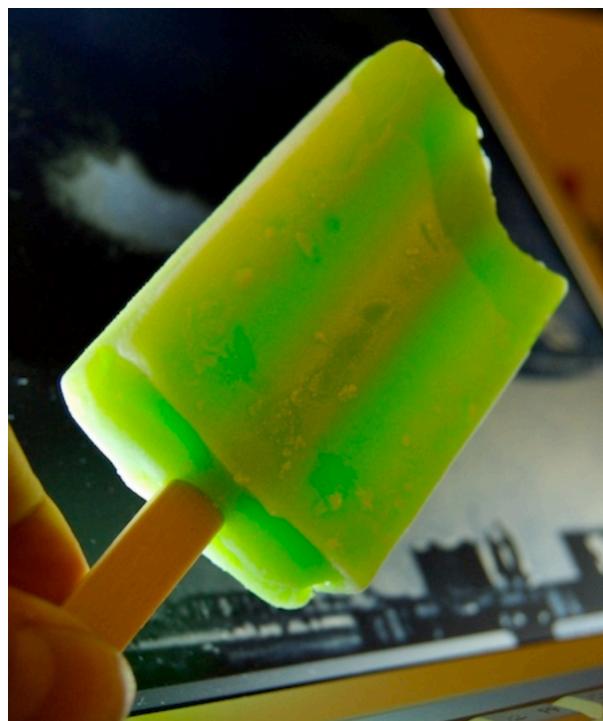


Figure 15: Visual prompt for the ‘frozen treat’ question on the survey.

1060 References

- 1061 Anonymous. 2012. Lolly ice or ice lolly? The great debate. *The Liverpool Echo*, ac-
1062 cessed 10 August 2020. <https://www.liverpoolecho.co.uk/news/liverpool-news/lolly-ice-ice-lolly-great-3337667>.
- 1063
- 1064 Ayto, John. 2012. *The diner's dictionary: Word origins of food and drink*. Oxford: Oxford Uni-
1065 versity Press.
- 1066 Bailey, George. 2018. *Variation and change in northern English velar nasals: Production and*
1067 *perception*: University of Manchester Doctoral dissertation.
- 1068 Bailey, George. 2019a. Emerging from below the social radar: Incipient evaluation in the North
1069 West of England. *Journal of Sociolinguistics* 23(1). 3–28. doi:10.1111/josl.12307.
- 1070 Bailey, George. 2019b. Ki(ng) in the north: Effects of duration, boundary, and pause on post-
1071 nasal [g]-presence. *Laboratory Phonology: Journal of the Association for Laboratory Phonol-*
1072 *ogy* 10(1). 1–26. doi:10.5334/labphon.115.
- 1073 Baranowski, Maciej. 2015. Part of town as an independent factor: The NORTH-FORCE merger
1074 in Manchester English. Paper presented at ICLaVE 8, Leipzig, 27–29 May.
- 1075 Baranowski, Maciej & Danielle Turton. 2018. Locating speakers in the socioeconomic hier-
1076 archy: Towards the optimal indicators of social class. Paper presented at NWA 47, New
1077 York University, 18–21 October.
- 1078 Barras, William. 2015. Lancashire. In Raymond Hickey (ed.), *Researching Northern Englishes*,
1079 271–292. Philadelphia, PA: John Benjamins. doi:10.1075/veaw.g55.12bar.
- 1080 Barras, William Simon. 2006. *The SQUARE-NURSE merger in Greater Manchester: The impact*
1081 *of social and spatial identity on phonological variation*: University of Edinburgh MA thesis.
- 1082 Bart, Gabriela, Elvira Glaser, Pius Sibler & Robert Weibel. 2013. Analysis of Swiss German
1083 syntactic variants using spatial statistics. In Ernestina Carrilho, Catarina Magro & Xosé
1084 Álvarez (eds.), *Current approaches to limits and areas in dialectology*, 143–169. Newcastle:
1085 Cambridge Scholars Publishing.
- 1086 Bauer, Laurie. 2002. *An introduction to international varieties of English*. Edinburgh: Edinburgh
1087 University Press.
- 1088 Beal, Joan. 2004. English dialects in the North of England: Morphology and syntax. In Bernd
1089 Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A*
1090 *handbook of varieties of English: A multimedia reference tool, vol. 2: Morphology and syntax*,
1091 114–141. Berlin: Mouton de Gruyter. doi:10.1515/9783110208399.2.373.

- 1092 Beal, Joan. 2008. English dialects in the north of England: phonology. In Bernd Kortmann &
1093 Clive Upton (eds.), *Varieties of English. vol. 1: The British Isles*, 122–144. Berlin: Mouton de
1094 Gruyter.
- 1095 Bender, David A. 2009. *A dictionary of food and nutrition*. Oxford: Oxford University Press.
1096 doi:10.1093/acref/9780199234875.001.0001.
- 1097 Biggs, Alison. 2016. Locating variation in the dative alternation. *Linguistic Variation* 16(2).
1098 151–182. doi:10.1075/lv.16.2.01big.
- 1099 Bivand, Roger, Tim Keitt & Barry Rowlingson. 2019. *rgdal: Bindings for the ‘geospatial’ data*
1100 *abstraction library*. <https://CRAN.R-project.org/package=rgdal>. R package version
1101 1.4-4.
- 1102 Britain, David. 1991. *Dialect and space: A geolinguistic study of speech variables in the Fens*:
1103 University of Essex Doctoral dissertation.
- 1104 Britain, David. 2001. Welcome to East Anglia! Two major dialect ‘boundaries’ in the Fens. In
1105 Peter Trudgill & Jacek Fisiak (eds.), *East Anglian English*, 217–242. Woodbridge: Boydell &
1106 Brewer.
- 1107 Britain, David. 2002a. Diffusion, levelling, simplification and reallocation in past tense BE in
1108 the English Fens. *Journal of Sociolinguistics* 6(1). 16–43. doi:10.1111/1467-9481.00175.
- 1109 Britain, David. 2002b. Space and spatial diffusion. In J.K. Chambers, Peter Trudgill & Natalie
1110 Schilling-Estes (eds.), *The handbook of language variation and change*, 603–637. Malden, MA:
1111 Blackwell. doi:10.1002/9780470756591.ch24.
- 1112 Britain, David. 2013. Space, diffusion and mobility. In J.K. Chambers & Natalie Schilling (eds.),
1113 *The handbook of language variation and change*, 471–500. Malden, MA: Blackwell 2nd edn.
1114 doi:10.1002/9781118335598.ch22.
- 1115 Britain, David, Tamsin Blaxter & Adrian Leemann. 2020. East Anglian English in the En-
1116 glish Dialects App: Regional variation in East Anglian English based on evidence from a
1117 smartphone-based survey. *English Today* 36(3). 14–30. doi:10.1017/S0266078420000206.
- 1118 Chambers, J. K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.
- 1119 Chambers, J.K. 1992. Dialect acquisition. *Language* 68(4). 673–705.
- 1120 Cheshire, Jenny, Vivian Edwards & Pam Whittle. 1993. Non-standard English and dialect
1121 levelling. In James Milroy & Lesley Milroy (eds.), *Real English: The grammar of English*
1122 *dialects in the British Isles*, 53–96. New York: Routledge.
- 1123 Cheshire, Jenny & Sue Fox. 2009. *Was/were* variation: A perspective from London. *Language*
1124 *Variation and Change* 21(1). 1–38. doi:10.1017/S0954394509000015.

- 1125 Clark, Urszula. 2008. The English West Midlands: Phonology. In Bernd Kortmann & Clive
1126 Upton (eds.), *Varieties of English. vol. 1: The British Isles*, 145–177. Berlin: Mouton de Gruyter.
- 1127 Clarke, Sandra. 2004. Newfoundland English: Morphology and syntax. In Bernd Kortmann,
1128 Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A handbook of
1129 varieties of English: A multimedia reference tool, vol. 2: Morphology and syntax*, 303–318.
1130 Berlin: Mouton de Gruyter. doi:10.1515/9783110208405.2.492.
- 1131 Cornips, Leonie & Cecilia Poletto. 2005. On standardising syntactic elicitation techniques
1132 (part 1). *Lingua* 115(7). 939–957.
- 1133 Docherty, Gerard & Paul Foulkes. 1999. Derby and Newcastle: Instrumental phonetics and
1134 variationist studies. In Paul Foulkes & Gerard Docherty (eds.), *Urban voices: Accent studies
1135 in the British Isles*, 47–71. Arnold London.
- 1136 Drummond, Rob. 2012. The Manchester Polish STRUT: Dialect acquisition in a second lan-
1137 guage. *Journal of English Linguistics* 41(1). 65–93. doi:10.1177/0075424212449172.
- 1138 Duncan, Daniel. 2018. *Language variation and change in the geographies of suburbs*: New York
1139 University dissertation.
- 1140 Ellis, Alexander J. 1889. *On early English pronunciation. Part V: The existing phonology of
1141 English dialects compared with that of West Saxon speech*. London: Trübner and Co.
- 1142 Evans, Bronwen G. & Paul Iverson. 2007. Plasticity in vowel perception and production: A
1143 study of accent change in young adults. *The Journal of the Acoustical Society of America*
1144 121(6). 3814–3826.
- 1145 Filppula, Markku. 2004. Irish English: Morphology and syntax. In Bernd Kortmann, Kate Bur-
1146 ridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A handbook of varieties of
1147 English: A multimedia reference tool, vol. 2: Morphology and syntax*, 73–101. Berlin: Mouton
1148 de Gruyter. doi:10.1515/9783110197181-085.
- 1149 Foulkes, Paul. 1997. Rule inversion in a British English dialect: A sociolinguistic investigation
1150 of [r]-sandhi in Newcastle upon Tyne. In Charles Boberg, Miriam Meyerhoff, Stephanie
1151 Strassel, Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.),
1152 *University of Pennsylvania working papers in linguistics 4.1: Selected papers from NWAVE 25*,
1153 259–270.
- 1154 Gerwin, Johanna. 2013. *Give it me!*: Pronominal ditransitives in English dialects. *English
1155 Language & Linguistics* 17(3). 445–463. doi:10.1017/S1360674313000117.
- 1156 Getis, Arthur. 2009. Spatial weights matrices. *Geographical Analysis* 41(4). 404–410. doi:
1157 10.1111/j.1538-4632.2009.00768.x.
- 1158 González, Félix Rodríguez. 1993. The naming of meals. *English Today* 9(4). 45–52. doi:10.1017/
1159 S026607840000732X.

- 1160 Grieve, Jack. 2017. Assessing smoothing parameters in dialectometry. In Martijn Wieling,
1161 Martin Kroon, Gertjan van Noord & Gosse Bouma (eds.), *From semantics to dialectometry:*
1162 *Festschrift in honor of John Nerbonne*, 119–126. London: College Publications.
- 1163 Grieve, Jack, Chris Montgomery, Andrea Nini, Akira Murakami & Diansheng Guo. 2019. Map-
1164 ping lexical dialect variation in British English using Twitter. *Frontiers in Artificial Intelli-*
1165 *gence* 2(10). doi:10.3389/frai.2019.00011.
- 1166 Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2011. A statistical method for the identification
1167 and aggregation of regional linguistic variation. *Language Variation and Change* 23(2). 193–
1168 221. doi:10.1017/S095439451100007X.
- 1169 Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2013. A multivariate spatial analysis of vowel
1170 formants in American English. *Journal of Linguistic Geography* 1(1). 31–51. doi:10.1017/jlg.
1171 2013.3.
- 1172 de Haas, Nynke & Ans van Kemenade. 2015. The origin of the Northern Subject Rule: Subject
1173 positions and verbal morphosyntax in older English. *English Language & Linguistics* 19(1).
1174 49. doi:10.1017/S1360674314000306.
- 1175 Haddican, William. 2010. Theme–goal ditransitives and theme passivisation in British English
1176 dialects. *Lingua* 120(10). 2424–2443. doi:10.1016/j.lingua.2009.11.003.
- 1177 Heath, Christopher. 1980. *The pronunciation of English in Cannock, Staffordshire*. Oxford:
1178 Blackwell.
- 1179 Herold, Ruth. 1990. *Mechanisms of merger: The implementation and distribution of the low back*
1180 *merger in eastern pennsylvania*: University of Pennsylvania dissertation.
- 1181 Herzog, Marvin. 1965. *The Yiddish language in northern Poland: Its geography and history*.
1182 Bloomington: Indiana University Press.
- 1183 Honeybone, Patrick. 2007. New-dialect formation in nineteenth century Liverpool: A brief
1184 history of Scouse. In Anthony Grant & Clive Grey (eds.), *The Mersey sound: Liverpool's*
1185 *language, people and places*, 106–140. Liverpool: Open House Press.
- 1186 Hughes, Arthur, Peter Trudgill & Dominic Watt. 2012. *English accents and dialects: An in-*
1187 *introduction to social and regional varieties of English in the British Isles*. London: Hodder
1188 Education 5th edn. doi:10.4324/9780203784440.
- 1189 Hundt, Marianne, Jennifer Hay & Elizabeth Gordon. 2004. New Zealand English: Morphosyn-
1190 tax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton
1191 (eds.), *A handbook of varieties of English: A multimedia reference tool, vol. 2: Morphology and*
1192 *syntax*, 560–592. Berlin: Mouton de Gruyter. doi:10.1515/9783110208412.2.305.
- 1193 Jansen, Sandra & Natalie Braber. 2020. FOOT-fronting and FOOT-STRUT splitting: vowel
1194 variation in the East Midlands. *English Language and Linguistics* 1–31.

- 1195 Jansen, Sandra, Justyna A. Robinson, Lynne Cahill, Adrian Leemann, Tamsin Blaxter & David
1196 Britain. 2020. Sussex by the sea: A descriptive analysis of dialect variation in the South East
1197 of England based on English Dialect App data. *English Today* 36(3). 31–39. doi:10.1017/
1198 S0266078420000218.
- 1199 Kirkham, Sam, Danielle Turton & Adrian Leemann. 2020. A typology of laterals in twelve
1200 English dialects. *The Journal of the Acoustical Society of America* 148(1). EL72–EL76. doi:
1201 10.1121/10.0001587.
- 1202 Knowles, Gerald. 1973. *Scouse: The urban dialect of Liverpool*: University of Leeds Doctoral
1203 dissertation.
- 1204 Knowles, Gerry. 1978. *The nature of phonological variables in Scouse*. London: Arnold.
- 1205 Kurath, Hans & Raven I. McDavid, Jr. 1961. *The pronunciation of English in the Atlantic States:
1206 Based upon the collections of the linguistic atlas of the Eastern United States*. University of
1207 Michigan Press.
- 1208 Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford, U.K.: Blackwell.
- 1209 Labov, William. 1996. When intuitions fail. In L. McNair, K. Singer, L. Dolbin & M. Aucon
1210 (eds.), *Papers from the parasession on theory and data in linguistics, Chicago Linguistic Society*
1211 32, 77–106.
- 1212 Labov, William. 2007. Transmission and diffusion. *Language* 83(2). 344–387. doi:10.1353/lan.
1213 2007.0082.
- 1214 Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Pho-
1215 netics, phonology, and sound change*. Berlin: Mouton de Gruyter. doi:10.1515/9783110167467.
- 1216 Labov, William, Mark Karen & Corey Miller. 1991. Near-mergers and the suspension of
1217 phonemic contrast. *Language Variation and Change* 3(1). 33–74. doi:https://doi.org/10.
1218 1017/S0954394500000442.
- 1219 Leemann, Adrian, David Britain & Tam Blaxter. 2017. Evidence of sound change in British En-
1220 glish crowdsourced using the ‘English Dialects App’. Poster presented at 4th International
1221 Workshop on Sound Change, University of Edinburgh.
- 1222 Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English Dialects App: The
1223 creation of a crowdsourced dialect corpus. *Ampersand* 5. 1–17. doi:10.1016/j.amper.2017.
1224 11.001.
- 1225 Lindsey, Geoff. 2019. *English after RP: Standard British pronunciation today*. Cham, Switzer-
1226 land: Palgrave Macmillan. doi:10.1007/978-3-030-04357-5.
- 1227 Llamas, Carmen. 2001. *Language variation and innovation in Teesside English*: University of
1228 Leeds dissertation.

- 1229 Llamas, Carmen. 2007. "A place between places": Language and identities in a border town.
1230 *Language in Society* 36(4). 579–604. doi:10.1017/S0047404507070455.
- 1231 Macfadzean. 2017. Post-nasal velar stop deletion in a word medial position in Cumbrian En-
1232 glish. Undergraduate dissertation, Newcastle University.
- 1233 MacKenzie, Laurel. 2018. Dialect surveys and student-led data collection. In Christine
1234 Mallinson, Becky Childs & Gerard Van Herk (eds.), *Data collection in sociolinguistics: Meth-*
1235 *ods and applications*, 151–153. New York, NY: Routledge 2nd edn.
- 1236 MacKenzie, Laurel. 2019. Perturbing the community grammar: Individual differences and
1237 community-level constraints on sociolinguistic variation. *Glossa: A Journal of General Lin-*
1238 *guistics* 4(1). 28. doi:10.5334/gjgl.622.
- 1239 MacKenzie, Laurel, George Bailey & Danielle Turton. 2016. Our Dialects: Mapping variation
1240 in English in the UK. Accessed 19 April 2021. <https://www.ourdialects.uk/>.
- 1241 Maguire, Warren. 2009. A survey of accents of English in Britain and Ireland. Accessed 16
1242 July 2020. <http://lel.ed.ac.uk/~wmaguire/survey/survey.html>.
- 1243 Maguire, Warren. 2012. Mapping the existing phonology of English dialects. *Dialectologia et*
1244 *Geolinguistica* 20. 84–107. doi:10.1515/dialect-2012-0006.
- 1245 Moore, Emma. 2011. Interaction between social category and social practice: Explaining
1246 *was/were* variation. *Language Variation and Change* 22(4). 347–371. doi:10.3389/frai.2019.
1247 00011.
- 1248 Morris, Jonathan. 2013. *Sociolinguistic variation and regional minority language bilingualism:*
1249 *An investigation of Welsh-English bilinguals in North Wales*: University of Manchester dis-
1250 sertation.
- 1251 Morris, Jonathan. 2017. Sociophonetic variation in a long-term language contact situation:
1252 /l/-darkening in Welsh-English bilingual speech. *Journal of Sociolinguistics* 21(2). 183–207.
- 1253 Newbrook, Mark. 1999. West Wirral: Norms, self reports and usage. In Paul Foulkes & Gerard
1254 Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 90–106. Arnold London.
- 1255 Nycz, Jennifer. 2015. Second dialect acquisition: A sociophonetic perspective. *Language and*
1256 *Linguistics Compass* 9(11). 469–482.
- 1257 Office for National Statistics. 2020. Estimates of the population for the UK, Eng-
1258 land and Wales, Scotland and Northern Ireland (Mid-2019). Accessed 19 Oc-
1259 tober 2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>.

- 1262 Ord, J. Keith & Arthur Getis. 1995. Local spatial autocorrelation statistics: Distributional issues
1263 and an application. *Geographical Analysis* 27(4). 286–306. doi:10.1111/j.1538-4632.1995.
1264 tb00912.x.
- 1265 Orton, H., S. Sanderson & J. Widdowson (eds.). 1978. *The linguistic atlas of England*. London:
1266 Croom Helm.
- 1267 Orton, Harold. 1962. *Survey of English Dialects (A): Introduction*. London and New York:
1268 Routledge.
- 1269 Pawley, Andrew. 2004. Australian Vernacular English: Some grammatical characteristics. In
1270 Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.),
1271 *A handbook of varieties of English: A multimedia reference tool, vol. 2: Morphology and syntax*,
1272 611–642. Berlin: Mouton de Gruyter. doi:10.1515/9783110208412.2.362.
- 1273 Pebesma, Edzer. 2018. Simple features for R: Standardized support for spatial vector data. *The
1274 R Journal* 10(1). 439–446. doi:10.32614/RJ-2018-009.
- 1275 Pichler, Heike. 2008. *A qualitative-quantitative analysis of negative auxiliaries in a northern
1276 English dialect: I don't know and I don't think, innit?*: University of Aberdeen dissertation.
- 1277 Pichler, Heike. 2010. Methods in discourse variation analysis: Reflections on the way forward.
1278 *Journal of Sociolinguistics* 14(5). 581–608.
- 1279 Prichard, Hilary & Meredith Tamminga. 2012. The impact of higher education on Philadelphia
1280 vowels. In Hilary Prichard (ed.), *University of pennsylvania working papers in linguistics 18.2:
1281 Selected papers from nwav 40*, 87–95.
- 1282 Quinn, Heidi. 2009. Pronoun forms. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative
1283 studies in Australian and New Zealand English: Grammar and beyond*, 31–47. Philadelphia,
1284 PA: John Benjamins Publishing. doi:10.1075/veaw.g39.03qui.
- 1285 Ross, Alan S. C. 1954. Linguistic class-indicators in Present-Day English. *Neuphilologische
1286 Mitteilungen* 55(1). 20–56.
- 1287 Rupp, Laura & David Britain. 2019. *Linguistic perspectives on a variable English morpheme:
1288 Let's talk about –s*. London: Palgrave Macmillan. doi:10.1057/978-1-349-72803-9.
- 1289 Schleef, Erik, Nicholas Flynn & Michael Ramsammy. 2015. Production and perception of (ing)
1290 in Manchester English. In Eivind Torgersen, Stian Hårstad, Brit Mæhlum & Unn Røyneland
1291 (eds.), *Selected papers from the seventh International Conference on Language Variation in
1292 Europe (ICLAVE 7)*, 197–210. Amsterdam: John Benjamins. doi:10.1075/silv.17.15sch.
- 1293 Shorrocks, Graham. 1990. Of gardens and gerdins: Mid-central vowels before /r/ in the north
1294 of England. *Journal of the Lancashire Dialect Society* 39. 24–28.
- 1295 Shorrocks, Graham. 1999. *A grammar of the dialect of the Bolton area*. Frankfurt am Main:
1296 Peter Lang.

- 1297 Smith, Jennifer. 2000. *Synchrony and diachrony in the evolution of English: Evidence from Scot-*
1298 *land*: University of York dissertation.
- 1299 Stenström, Anna-Brita. 1997. Can i have a chips please? – just tell me what one you want:
1300 Nonstandard grammatical features in London teenage talk. In Jan Aarts, Inge de Mönnink
1301 & Herman Wekker (eds.), *Studies in english language and teaching: In honour of flor aarts*,
1302 141–151. Atlanta, GA: Rodopi.
- 1303 Stevenson, Jon. 2019. Variation in the pronominal ditransitive in British English Twitter mes-
1304 sages. Talk given at UKLVC10, Queen Mary University London.
- 1305 Stoddart, Jana, Clive Upton & J. Widdowson. 1999. Sheffield dialect in the 1990s: Revisiting the
1306 concept of NORMs. In Paul Foulkes & Gerald Docherty (eds.), *Urban voices: Accent studies*
1307 *in the British Isles*, 72–89. London: Arnold.
- 1308 Strycharczuk, Patrycja, Georgina Brown, Adrian Leemann & David Britain. 2019. Investigat-
1309 ing the FOOT-STRUT distinction in northern Englishes using crowdsourced data. *Proceed-*
1310 *ings of ICPhS 2019* 76.
- 1311 Strycharczuk, Patrycja, Manuel López-Ibáñez, Georgina Brown & Adrian Leemann. 2020. Gen-
1312 eral Northern English. Exploring regional variation in the North of England with machine
1313 learning. *Frontiers in Artificial Intelligence* 3(48). 1–18. doi:10.3389/frai.2020.00048.
- 1314 Suddaby, George. 2017. The NURSE-SQUARE merger in Kingston upon Hull. Poster presented
1315 at Speech and Bubbles undergraduate research conference, Newcastle University. <https://bit.ly/suddaby-nurse-square-hull>.
- 1317 Szmrecsanyi, Benedikt, Jason Grafmiller, Joan Bresnan, Anette Rosenbach, Sali Tagliamonte
1318 & Simon Todd. 2017. Spoken syntax in a comparative perspective: The dative and genitive
1319 alternation in varieties of English. *Glossa: A Journal of General Linguistics* 2(1). doi:10.5334/
1320 gjgl.310.
- 1321 Szmrecsanyi, Benedikt, Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger. 2016.
1322 Around the world in three alternations: Modeling syntactic variation in varieties of En-
1323 glish. *English World-Wide* 37(2). 109–137. doi:10.1075/eww.37.2.01szm.
- 1324 Tagliamonte, Sali. 1998. *Was/were* variation across the generations: View from the city of York.
1325 *Language Variation and Change* 10(2). 153–191. doi:10.1017/S0954394500001277.
- 1326 Tamminga, Meredith. 2013. Phonology and morphology in Dutch indefinite determiner syn-
1327 cretism: Spatial and quantitative perspectives. *Journal of Linguistic Geography* 1. 115–124.
1328 doi:10.1017/jlg.2013.10.
- 1329 Tamminga, Meredith. 2019. Interspeaker covariation in Philadelphia vowel changes. *Language*
1330 *Variation and Change* 31(2). 1–15. doi:10.1017/S0954394519000139.

- 1331 Tillery, Jan & Guy Bailey. 2003. Approaches to real time in dialectology and sociolinguistics.
1332 *World Englishes* 22(4). 351–365.
- 1333 Trudgill, Peter. 1974. Linguistic change and diffusion: Description and explanation in soci-
1334 olinguistic dialect geography. *Language in Society* 3(2). 215–246.
- 1335 Turton, Danielle. 2015. 4,000 /r/s from Blackburn, Lancashire: A (socio)phonological analysis
1336 of rhoticity in Northern England. Plenary talk given at the Manchester Forum in Linguistics,
1337 November 2015.
- 1338 Turton, Danielle & Maciej Baranowski. 2020. Not quite the same: The social stratification
1339 and phonetic conditioning of the FOOT-STRUT vowels in Manchester. *Journal of Linguistics*
1340 1–39. doi:10.1017/S0022226720000122.
- 1341 Upton, Clive & John Widdowson. 1996. *An atlas of English dialects*. Oxford: Oxford University
1342 Press.
- 1343 Vaux, Bert & Scott A. Golder. 2003. Harvard Dialect Survey. Accessed 17 September 2019.
1344 <http://dialect.redlog.net/>.
- 1345 Vaux, Bert & Marius L. Jøhndal. 2009. The Cambridge online survey of world Englishes. Ac-
1346 cessed 5 October 2020. http://www.tekstlab.uio.no/cambridge_survey.
- 1347 Wagner, Suzanne Evans. 2012. Real-time evidence for age grad(ing) in late adolescence. *Lang-
1348 uage Variation and Change* 24(2). 179–202.
- 1349 Wakelin, Martyn. 1984. Rural dialects in England. In Peter Trudgill (ed.), *Language in the
1350 British Isles*, 70–93. Cambridge: Cambridge University Press.
- 1351 Wales, Katie. 2004. Second person pronouns in contemporary English: The end of a story or
1352 just the beginning? *Franco-British Studies* 33(4). 172–185.
- 1353 Wales, Katie. 2006. *Northern English: A social and cultural history*. Cambridge: Cambridge
1354 University Press. doi:10.1017/CBO9780511487071.
- 1355 Walker, James A. 2020. Modelling socio-grammatical variation: Plural existentials in Toronto
1356 English. In Karen V. Beaman, Isabelle Buchstaller, Susan Fox & James A. Walker (eds.),
1357 *Advancing socio-grammatical variation and change: In honour of jenny cheshire*, 165–184.
1358 New York, NY: Routledge.
- 1359 Walter, Bronwen. 2010. Irish/Jewish diasporic intersections in the East End of London:
1360 Paradoxes and shared locations. In Michel Prum (ed.), *La place de l'autre*, 53–67. Paris:
1361 L'Harmattan.
- 1362 Watson, Kevin & Lynn Clark. 2013. How salient is the NURSE~SQUARE merger? *English
1363 Language & Linguistics* 17(2). 297–323. doi:10.1017/S136067431300004X.

- 1364 Watt, Dominic & Catherine Ingham. 2000. Durational evidence of the Scottish Vowel Length
1365 Rule in Berwick English. *Leeds Working Papers in Linguistics* 8. 205–228.
- 1366 Watt, Dominic, Carmen Llamas & Daniel Ezra Johnson. 2014. Sociolinguistic variation on
1367 the Scottish-English border. In Robert Lawson (ed.), *Sociolinguistics in Scotland*, 79–102.
1368 London: Palgrave Macmillan. doi:10.1057/9781137034717_5.
- 1369 Watts, Emma. 2005. *Mobility-induced dialect contact: A sociolinguistic investigation of speech
1370 variation in Wilmslow, Cheshire*: University of Essex Doctoral dissertation.
- 1371 Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- 1372 Wells, John C. 1978. Review of *The Linguistic Atlas of England* by Harold Orton, Stewart
1373 Sanderson and John Widdowson. *The Times Higher Education Supplement*, 1 December
1374 1978. <https://www.phon.ucl.ac.uk/home/estuary/lae-revw.htm>.
- 1375 West, Helen Faye. 2015. Language attitudes and divergence on the Merseyside/Lancashire
1376 border. In Raymond Hickey (ed.), *Researching Northern Englishes*, Amsterdam: John Ben-
1377 jamins. doi:10.1075/veaw.g55.14wes.
- 1378 Whyte, William. 2019. *Somewhere to live: Why British students study away from home – and
1379 why it matters*. Oxford: Higher Education Policy Institute.
- 1380 Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer. doi:
1381 10.1007/978-0-387-98141-3. <https://ggplot2.tidyverse.org>.
- 1382 Wieling, Martijn, Clive Upton & Ann Thompson. 2014. Analyzing the BBC Voices data: Con-
1383 temporary English dialect areas and their characteristic lexical variants. *Literary and Lin-
1384 guistic Computing* 29(1). 107–117. doi:10.1093/lrc/fqt009.
- 1385 Williams, Angie, Peter Garrett & Nikolas Coupland. 1996. Perceptual dialectology, folklinguis-
1386 tics, and regional stereotypes: Teachers' perceptions of variation in Welsh English. *Multi-
1387 lingua* 15(2).
- 1388 Williams, Ann & Paul Kerswill. 1999. Dialect levelling: change and continuity in Milton
1389 Keynes, Reading and Hull. In Paul Foulkes & Gerald Docherty (eds.), *Urban voices: Accent
1390 studies in the British Isles*, 141–162. London: Arnold.
- 1391 Wood, Jim. 2019. Quantifying geographical variation in acceptability judgments in regional
1392 American English dialect syntax. *Linguistics* 57(6). 1367–1402. doi:10.1515/ling-2019-0031.
- 1393 Yáñez-Bouza, Nuria & David Denison. 2015. Which comes first in the double object construc-
1394 tion? *English Language & Linguistics* 19(2). 247–268. doi:10.1017/S136067431500012X.