3 Lexical variation among mobile speakers

A case study of words for bread in the United Kingdom

George Bailey, Laurel MacKenzie, and Danielle Turton

3.1 Introduction

Lexical variables are typically the most easily acquired type of linguistic feature in cases of dialect contact (Chambers 1992; Siegel 2010). However, studies of dialect contact typically focus on a small number of speakers, meaning that we are missing a large-scale picture of how widespread contact-based lexical assimilation is and what sorts of constraints might govern the process.

This chapter takes a first step in answering these questions by presenting data on lexical variation from a large-scale dialect survey carried out among residents of the United Kingdom. As we intend this chapter as a proof of concept of our method, we focus on one lexical variable in particular, which happens to be simultaneously highly variable, well known among members of the public, and understudied by dialectologists: words for a bread roll in British English. Drawing on data from over 20,000 individuals, our chapter answers the following research questions:

1 Are speakers who were raised in multiple locations more likely to report using multiple lexical variants than those who weren't?
2 If so, is there a baseline amount of time a speaker needs to have spent in a new dialect region before picking up the local variant?
3 Are all lexical variants equally acquirable, or are certain lexical variants more susceptible to acquisition than others?

The first question addresses lexical variation in sheer quantitative terms, and our data confirm that survey respondents who self-report having grown up in multiple locations are also more likely to self-report using more than one term for a bread roll than individuals who self-report having grown up in only one location. We approach the second and third questions by comparing mobile individuals who lived in a given region

DOI: 10.4324/9781003510581-4

for the majority of their childhood with mobile individuals who lived in the same region but for a non-majority of their childhood. This tells us whether a major amount of time during childhood is required in a given dialect area in order for a speaker to pick up that area's variant. We find this to differ across variants in non-random ways.

All in all, we demonstrate that, even given the admittedly coarse measures of mobility that we are limited to by a questionnaire with tens of thousands of respondents, meaningful patterns still emerge. As such, we not only document new findings concerning the patterning of this variable in particular but also pave the way for follow-up work, both on additional lexical variables and on phonological and morphosyntactic variation among the same mobile individuals.

3.2 Background

In this paper, we are asking whether mobility influences linguistic repertoire, with a focus on the lexical domain. As we will discuss further in Section 3.3, our data on mobility are restricted specifically to mobility between the ages of 4 and 13. This is because our dialect survey asked respondents to detail their history of residence during this period only. Thus, throughout this paper, when we discuss 'mobile' and 'non-mobile' individuals, we are really talking about those individuals who were (non-) mobile between the ages of 4 and 13.

Our interest in individuals' residence history between these specific ages follows other dialect surveys, such as that of Labov et al. (2006), who based their sample on respondents who were non-mobile during their youth. Rather than asking about a specific age range, Labov et al.'s survey asked respondents if they 'grew up' in the location where they were currently living; those who responded in the negative were deemed 'non-local' and excluded from the final sample of speakers (Labov et al. 2006: 26). Labov et al. do not explicitly justify their decision to restrict their sample to individuals who grew up in a place and continued to live there, but presumably the implication is that those speakers who were living somewhere they didn't grow up were not likely to speak with the local vernacular of the place in which they were residing. It is to this point that we now turn.

As Siegel (2010: 96) summarises, 'there are certain age periods in which different aspects of a [second dialect] appear to be easier to acquire, and native-like attainment more likely.' Specifically, the takeaway from a large body of second dialect acquisition research is that 'those who began acquisition at a younger age, especially 13 years or younger, had the highest averages of percentage of use of [second dialect] variants, and [constituted]

the greatest proportions of individuals who reached native-like usage overall or in particular variables' (Siegel 2010: 84). What this means for our study is that respondents to our survey who were mobile before the age of 13 will have the highest likelihood of allowing multiple lexical variants, because they are the ones who will have lived in multiple dialect areas during the critical age for second dialect acquisition. This suggests that the answer to our first research question – are speakers who were raised in multiple locations more likely to report using multiple lexical variants than those who weren't? – should be 'yes.'

Additionally supporting this idea is the finding that lexical variants are easier to acquire than phonological ones. For instance, Chambers (1992), in a study of six Canadian youths who moved to southern England, found that every subject showed greater assimilation to local lexical than pronunciation variants. Because we are looking specifically at lexical variation in this chapter, we should expect high rates of uptake of regional variants among our mobile respondents.

All this being said, there are some nuances to the observation that those who were mobile in their youth are likely to reflect this linguistically. For starters, even among those who relocated before the age of 13, the age at which a speaker moved to a new dialect area still influences their acquisition of local features. As Chambers (1992: 689) puts it, 'a person 7 or under will almost certainly acquire a new dialect perfectly, and a person 14 or over almost certainly will not. In between those ages, people will vary.' Additionally, uptake of a second dialect, even in the lexical domain, isn't necessarily categorical: in Chambers' study, even the most lexically assimilated child used British English variants at a rate of only 71%.[1] All told, then, mobility in youth suggests, but does not guarantee, concomitant linguistic malleability.

Another complication is that people can acquire new lexical variants well into adulthood. Kerswill (1996: 200) lists 'new lexical forms of old words' and new vocabulary items as the features that are easiest to acquire in dialect contact situations, with acquisition of these items continuing throughout the lifespan. Supporting this, Foreman (2003: 170), studying 34 North Americans who moved to Australia (all but two of them after the age of 13), found that 'most of the subjects stated that they had substituted some [Australian English] words for [American English] or [Canadian English] words.' Because our survey asked only about mobility between the ages of 4 and 13 and not mobility in later life, we are likely classifying as 'non-mobile' many individuals who were actually mobile after the age of 13. This wouldn't matter if we were studying variables that are difficult to acquire in adulthood, but lexical variables are not this type. We recognise this as a drawback of our method; at the

same time, the coherent differences between our 'mobile' and 'non-mobile' populations that we show in Section 3.4 suggest that this isn't skewing the results too much. A particularly interesting direction for future work will be to compare our 'non-mobile' individuals' performance on lexical variables to their performance on phonological variables like mergers and splits, which are much harder to acquire in later life (Payne 1980; Kerswill 1996).

Our second research question asks whether there is a baseline amount of time a speaker needs to have spent in a new dialect region before picking up the local variant. Chambers (1992: 681) shows that lexical replacement occurs quickly; he concludes that 'dialect acquirers make most of the lexical replacements they will make in the first two years' and in fact presents data on a group of British immigrants to Toronto who were already using North American lexical items within three months of arrival. Still, the same study found that another group of British immigrants, who had been in Toronto longer (one to two years), were using North American lexical items at higher rates. This suggests that acquisition of new lexical variants does progressively increase after arrival. Our dialect survey didn't elicit specific lengths of time in residence, but we did ask our mobile respondents to indicate, among the locations they were raised in, which one they lived in the longest. This then allows us to compare the lexical choices of mobile respondents who lived in Place A for the majority of their mobile childhood to those of mobile respondents who lived in Place A for a non-majority of their mobile childhood, to look for differences in lexical acquisition.

Finally, our third research question asks whether all lexical variants are equally acquirable or whether certain lexical variants are more susceptible to acquisition than others. Chambers (1992: 679) effectively concludes that there is no predictable pattern to which lexical variants have been acquired among his mobile speakers; still, he notes that polysemy may play a role, observing that '*trousers* has replaced the [Canadian English] word *pants* for five of the six subjects, presumably because *pants* in England means 'underpants' and is often a source of confusion and embarrassment for North Americans who use it in conversation.' However, usage of *pants* can be more nuanced even within the UK: our own data from northern England suggest that around 50% of speakers in the North West also say *pants* to mean 'trousers,' underscoring the complexity of lexical variation across regions. Siegel (2010: 144) also comments on the confusion engendered by lexical items that have a different meaning in the new dialect. Plausibly, this confusion makes these new lexical items more salient, perhaps leading to greater uptake. The localised bread roll variant *muffin* is one such item; we explore whether this variant is differently acquired from the other, non-polysemous variants, in Section 3.4.

3.3 Methodology

The data we analyse here come from the long-running *Our Dialects* survey, which started in 2013 as a paper survey distributed by undergraduate students of the authors but is now an online survey hosted on Qualtrics with continuous, open-ended data collection. The survey has led to the creation of an online dialect atlas for the general public (https://www.ourdialects.uk/) as well as research output (see MacKenzie et al. 2022) that identifies examples of language change through comparisons with earlier dialect surveys such as the 1950s *Survey of English Dialects* (Orton 1962).

The survey collects some basic demographic data from participants, such as their age, gender identity, occupation, and geographical background (detailed below), and it now contains over 50 questions targeting various dialect features (10 lexical, 21 phonological, and 23 morphosyntactic). For more details on the survey design and the demographic makeup of our respondents, please see MacKenzie et al. (2022).

The analysis we present here is based on all data collected up to July 2024, totalling 20,501 responses, and as outlined earlier we specifically analyse the variants of *bread roll*, specifically the ten most frequent terms reported by our participants (excluding the term *roll* itself, which is regionally unmarked): *bap*, *barm*, *batch*, *bun*, *cob*, *flourcake*, *muffin*, *softie*, *stottie*, and *teacake*. As with many of the lexical questions on the survey, participants were presented with a picture (see Figure 3.1) and asked to

Figure 3.1 The image of a *bread roll* participants were presented with.

tick the box(es) corresponding to the term(s) they would use. By allowing participants to select multiple responses, we are able to address the question of whether or not mobile speakers are more likely to demonstrate intra-speaker variation and report use of more than one variant (e.g. a local dialect form alongside the supralocal *roll*, or even multiple regional variants). In the analysis, we distinguish speakers who use a variant exclusively (i.e. reporting *only that* variant) and those who use it variably (i.e. report using it alongside other forms), in order to address the first of our research questions as outlined in Section 3.1.

We collect two types of geographical metadata in this survey: participants are asked for the postcode district they lived longest between the ages of 4 and 13, and then in a separate question, they are prompted to enter the full list of towns/cities/villages they lived in within that same age window. In earlier analyses, we plot dialect maps using only the former to geolocate speakers, but here we present the first analysis using information from the latter question. This required going through all responses and identifying those who report multiple locations, which we then input into a geocoder using the tidygeocoder package in R (Cambon et al. 2021) and the default 'OSM' (OpenStreetMap) service to convert addresses (e.g. 'Bury, Greater Manchester') into latitude and longitude coordinates.

Having generated a set of coordinates for all locations reported by each speaker, an initial classification of mobility was based on whether or not speakers moved (in increasing magnitude of mobility) between postcode districts, postcode areas, or regions of the UK. However, this operationalisation is limited given that the dialectal variants for *bread roll* show localised patterns to different degrees, some covering whole regions of the UK while others are restricted to specific cities. In other words, it's entirely possible that there are cases of speakers moving somewhat locally between adjacent postcode areas and encountering different variants and speakers moving between totally different regions of the UK and still finding themselves within the same variant isogloss (*bun*, for example, is used across the North East but also found in Yorkshire).

With that in mind, we instead identified geographic mobility for each speaker relative to each variant isogloss. That is, for each of the ten *bread roll* variants, we identify its regional distribution based on the responses from non-mobile speakers, creating isoglosses at the level of individual postcode areas. The isogloss for *batch*, for example, covers the CV (Coventry) and CH (Chester) postcode areas: this form is used by 42% and 38% of speakers in these locations, respectively, compared with the 0%–8% rates of occurrence in all other postcode areas. We then classified each speaker based on whether, between the ages of 4 and 13, they:

- **always** lived inside that variant's isogloss (whether mobile or not)
- **mostly** lived inside that variant's isogloss (these are mobile speakers who spent longest living in a postcode district inside the isogloss but also lived in other locations outside that isogloss)
- **partially** lived inside that variant's isogloss (these are mobile speakers who report living in at least one location inside the isogloss but spent most of their childhood living *outside* it)
- **never** lived inside that variant's isogloss (whether mobile or not)

The analysis involved calculating rates of use of each *bread roll* variant across the different mobility categories and distinguishing variable use from categorical use at the intra-speaker level. In total, there are 2,186 speakers in the dataset (out of 20,501) who were geographically mobile when growing up between the ages of 4 and 13. Almost 90% (1,951) of these lived in two to three locations, but some report living in as many as seven or ten distinct locations. All dialect maps were created using the sf package in R (Pebesma & Bivand 2023), and mixed-effects regression models were fit using lme4 (Bates et al. 2015).

3.4 Results

In Figure 3.2, we plot individual dialect maps for the ten main variants of *bread roll*, with postcode areas colour-coded by the percentage use of the variant and the solid black outline corresponding to the isogloss we propose for each variant. Alongside these, we present bar charts in Figure 3.3, which plot the percentage use of each variant across the four mobility groups outlined in Section 3.3. Below, we address the regional distribution and usage patterns of each variant in turn.

3.4.1 Batch

We start with a variant which is highly localised, but to *two* quite disparate locations: users of *batch* are clustered around the metropolitan borough of the Wirral in Merseyside, North West England, and around Coventry in the West Midlands.

Among those who spent their whole childhood inside the isogloss, hereafter the ALWAYS group, *batch* is used 39.3% (N = 497) of the time with most of those reporting exclusive use of this variant. There is only a slight drop-off in frequency to 30.5% (N = 36) for those mobile speakers who lived mostly inside the isogloss but who spent some of their childhood outside it (hereafter the MOSTLY group) and a much steeper drop-off to 6% (N = 53) for those who spent a minority of their childhood living inside the isogloss (hereafter the PARTIALLY group).

Figure 3.2 Dialect maps illustrating the regional variation of ten variants of *bread roll*, divided by postcode area boundaries; solid black lines indicate a variant's isogloss, used for mobility classification.

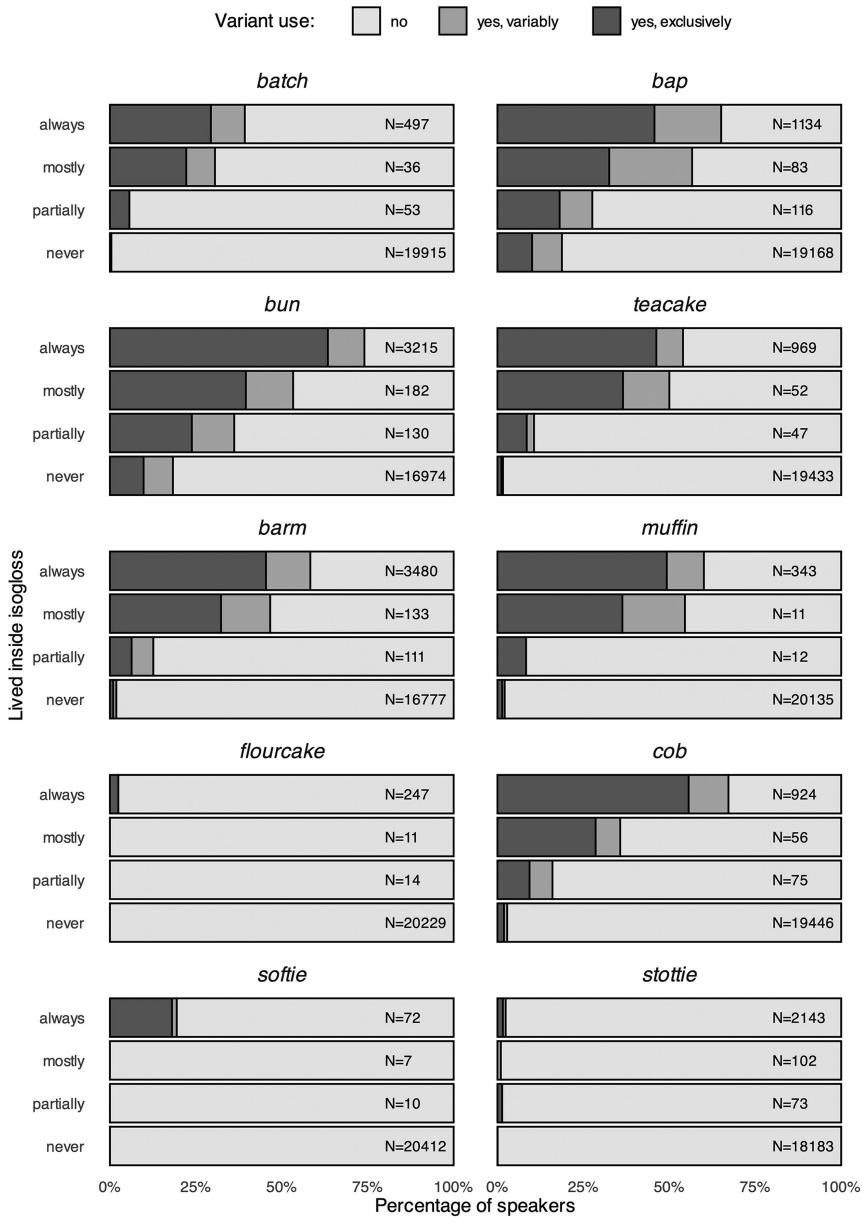

Figure 3.3 Percentage of variant use by mobility category, distinguishing variable from categorical/exclusive use, for each of the ten *bread roll* forms.

3.4.2 Bap

Bap is much more geographically widespread with a centre of gravity towards the western parts of the UK: the highest rates are found in Northern Ireland and the Isle of Man, followed by most of Wales, and parts of the West Midlands and North West of England (such as Walsall, Stoke-on-Trent, and Crewe).

Bap shows a more steady drop-off in exclusive use of the variant from the ALWAYS (45.7%, *N* = 1,134) to the MOSTLY (32.5%, *N* = 83) and to the PARTIALLY (18.1%, *N* = 116) groups. Variable use of this form is also much higher compared to *batch*, indicating that it's much more likely to co-exist in speakers' grammars alongside other variants (most commonly *roll* and *bun*). It's also notable that it is the MOSTLY group in particular that reports the highest *variable* use of this form.

3.4.3 Bun

Bun patterns in much the opposite way relative to *bap*, with an isogloss not only covering the eastern parts of England instead, the North East in particular (including the likes of Newcastle, Middlesbrough, Sunderland, and County Durham), but also extending to parts of Yorkshire such as Doncaster and York.

Overall, rates are very high for those who live(d) inside the isogloss, with 74% (*N* = 3,215) of the ALWAYS group reporting the use of this form. This variant also shows a regular and consistent decrease in usage across the mobility groups but, compared to *bap*, the ratio of categorical to variable users is higher, indicating that if a speaker uses *bun*, they are more likely to only use *bun*.

3.4.4 Teacake

Teacake is used in an area that covers not only various parts of Yorkshire, including Bradford, Halifax, Huddersfield, and Wakefield, but also parts of east Lancashire such as Burnley, Blackburn, and Accrington. It is also used to a lesser extent in Carlisle (31.2%), but we exclude this from the isogloss given how geographically clustered the other locations are and Carlisle's lack of contiguity with them.

For *teacake*, the chart in Figure 3.3 shows very similar rates of usage between the ALWAYS and MOSTLY groups (53.9% and 50%, respectively), but then much like we see for *batch*, there is a significant drop-off in use for the PARTIALLY group (10.6%, *N* = 47), which again suggests that the time spent inside a dialect area during childhood might play an important role in aspects of (second) dialect acquisition.

Note that in other parts of the country, *teacake* refers to a different variation that contains dried fruit (usually currants or sultanas).

3.4.5 Barm

Barm is specific to a relatively small part of the North West of England, with an isogloss encompassing most of Greater Manchester (including the M, BL, SK, and WN postcode areas) and extending to Warrington, Preston, Blackpool, and Liverpool. It is particularly dominant in Wigan, where 91.4% ($N = 233$) of respondents report the use of this form.

In terms of usage patterns across the mobility groups, it stands as yet another example of the ALWAYS and MOSTLY groups being fairly similar (58.3% and 46.6%, respectively) before a significant drop-off in use for the PARTIALLY group (12.6%, $N = 111$).

3.4.6 Muffin

Like *barm*, *muffin* is associated with the North West of England, but our data suggest its regional spread is even more restricted, its strongest concentration being in the OL postcode area (60.1%, $N = 341$), covering locales in eastern Greater Manchester such as Oldham, Ashton-under-Lyne, and Rochdale. Although *muffin* appears in other Greater Manchester postcodes (BL, M, SK) and is dominant in more specific postcode districts in areas like East Manchester and parts of Bolton and Stockport, because overall rates within the wider postcode area are lower (15.1%–21.4%), it falls below the threshold we used to define a distinct isogloss.

The use of *muffin* among our mobile respondents, like many other variants already discussed here, decreases significantly as a function of time spent in the isogloss. Its frequency is very similar between the ALWAYS (60.1%, $N = 343$) and MOSTLY (54.6%, $N = 11$) groups but then decreases dramatically in the PARTIALLY group (8.3%, $N = 12$). These results should be taken with more caution than usual, though, given that we are working with a relatively small number of mobile participants here by virtue of the more geographically restricted isogloss. Note that, similarly to *teacake*, *muffin* has additional meanings in other parts of the country (and this part). One is a more specific type of flat bread roll, and the other is a sweetened cake.

3.4.7 Flourcake

Flourcake is yet another dialectal variant restricted to a very specific part of the North West of England, namely the Greater Manchester boroughs of Bolton and Bury, but even here it is very rare with only six speakers

reporting the use of it (2.5%, *N* = 245), plus one speaker in nearby Wigan. Surprisingly, all of these speakers report using this variant exclusively, despite its apparent rarity; the fact that five of these speakers were born earlier than 1956 suggests that this might be a traditional dialect form that has fallen almost entirely out of use.

3.4.8 Cob

Our results indicate that *cob* is very much associated with the East Midlands of England, with an isogloss encompassing the Derby (67.9%, *N* = 221), Nottingham (71.6%, *N* = 405), and Leicester (60.6%, *N* = 264) postcode areas.

This variant shows a quite different pattern of use compared to many others discussed thus far. Considering how frequent the use of *cob* is among the ALWAYS group (67.2%, *N* = 924), there is quite a large drop-off to the MOSTLY group (35.7%, *N* = 56), and with very low rates in the PARTIALLY group (16%, *N* = 75) the results suggest that this could be a variant easily lost (or not easily acquired) for mobile speakers who moved in/out of this area.

3.4.9 Softie

Softie is the only variant of the ten investigated here which is localised to Scotland, specifically the Aberdeen (19.7%, *N* = 66) and Shetland (25%, *N* = 4) postcode areas (though the severe lack of data from the Shetland Islands means that this corresponds to a single speaker reporting its use, so this should be interpreted with caution).

It clearly still exists only as a minority variant even among those who grew up entirely inside the isogloss. This is also possibly the reason why it is completely absent among both mobile speaker groups.

3.4.10 Stottie

Like *bun*, *stottie* is another variant localised to the North East of England, but in this case showing an even more geographically restricted distribution (namely the DH, NE, SR, and TS postcode areas), and is clearly a minority variant here with rates as low as 1%–3% even among the ALWAYS group. As such, the rates are so low that no clear pattern is evident across the mobility categories, unlike almost all of the other variants discussed here.

3.4.11 Summary

Having established the regional and usage pattern of each variant in turn, here we present an overall summary resulting from mixed-effects logistic regression models that combine responses to all ten variants together.

A model was fit to the 205,010 data points (20,501 respondents × 10 dummy-coded variants) with a fixed effect of *mobility*, using the same classification employed throughout the results section, and random intercepts of *participant* and *variant*. The dependent variable was modelled in a binary way, modelling any kind of use of a variant (whether variable or categorical/exclusive). The results, presented in Table 3.1, indicate that the likelihood of using (and therefore, having acquired) a given dialectal variant decreases significantly as a function of time spent inside the variant's isogloss during childhood. Estimated marginal means were calculated using the `emmeans` package in R in order to conduct pairwise comparisons between mobility classes, and the results highlight how this decrease in log-likelihood is also significant between each pair of categories, decreasing by 0.875 from ALWAYS to MOSTLY, by 1.480 from MOSTLY to PARTIALLY, and by 1.098 from PARTIALLY to NEVER (all significant at $p < 0.001$).

A separate model was also fit to just those speakers who do report the use of a variant (i.e. only the positive responses) and with a dependent variable that instead captures the distinction between variable and categorical use of a given dialectal form, in order to address the possible link between mobility and increased intra-speaker variation. The model summary is presented in Table 3.2, which highlights how localised dialectal forms are significantly more likely to be used variably, alongside other forms, by speakers who were geographically mobile during childhood ($\beta = 0.594$ for the MOSTLY group and $\beta = 0.782$ for the PARTIALLY group, $p < 0.001$ for both).

Table 3.1 Results of mixed-effects logistic regression predicting the use of a dialectal form (binary 0/1) based on mobility relative to its isogloss ($N = 205{,}010$)

	Estimate	*St. error*	*z-value*	*p-value*	
Intercept (*always*)	–1.027	0.173	–5.951	<0.001	***
Mostly	–0.875	0.096	–9.156	<0.001	***
Partially	–2.355	0.101	–23.391	<0.001	***
Never	–3.453	0.028	–125.440	<0.001	***

Table 3.2 Results of mixed-effects logistic regression predicting the variable (rather than categorical/exclusive) use of a dialectal form based on mobility relative to its isogloss ($N = 7{,}148$)

	Estimate	*St. error*	*z-value*	*p-value*	
Intercept (*always*)	–1.454	0.123	–11.292	<0.001	***
Mostly	0.594	0.142	4.190	<0.001	***
Partially	0.782	0.208	3.763	<0.001	***

We can also delve deeper into these variable speakers and ask to what extent does their variability simply represent the use of a local variant alongside what might be deemed as the unmarked supralocal form, i.e. *bread roll*? Have these speakers actually acquired multiple localised forms? The results paint a somewhat mixed picture: of the 3,218 speakers who report using multiple variants for the *bread roll* variable, 47.5% report using one local regional form alongside the supralocal *roll*, which means a slight majority (52.5%) do indeed report using two or more regional variants.

3.5 Discussion and conclusion

At the outset of this chapter, we identified three research questions that motivated our study:

1 Are speakers who were raised in multiple locations more likely to report using multiple lexical variants than those who weren't?
2 If so, is there a baseline amount of time a speaker needs to have spent in a new dialect region before picking up the local variant?
3 Are all lexical variants equally acquirable, or are certain lexical variants more susceptible to acquisition than others?

The answer to the first two questions is a resounding 'yes': intra-speaker variation is significantly higher for the geographically mobile speakers in our population sample, relative to those who never moved outside of a variant's isogloss during childhood (between the ages of 4 and 13). More specifically, we have shown that intra-speaker variation increases as an inverse function of the amount of time spent inside a variant's isogloss during childhood. That is, speakers who lived inside a variant's isogloss for most (but not all) of their childhood are more likely to use other forms alongside the one associated with that region, and this effect is even stronger for those who spent a minority of their childhood in the isogloss. For the latter group, their shorter residential history in the area means that we can assume that they are even more likely to have lived in multiple other locations (and therefore even more likely to have acquired other variants and use them variably).

These findings support the broader phenomenon of dialect levelling, where increased geographical and social mobility leads to the erosion of localised forms in favour of supralocal or majority variants (Trudgill 1986; Britain 2013). This levelling process, which is argued to be driven by increasingly mobile societies, not only reduces linguistic diversity but also results in speakers encountering and potentially adopting features from new dialect regions. Future work could explore whether these patterns are reflected in phonological and morphosyntactic variables, which

are generally less susceptible to mobility-induced levelling (Payne 1980; Chambers 1992; Kerswill 1996).

As discussed at the end of Section 3.4, roughly half of these variable speakers report using just two variants, with one of them being the supralocal *bread roll*, which is predominant in Standard Southern British English (SSBE) and across large swathes of the South West, South East, and East of England, and also much of Wales and Scotland. Although a full analysis of *bread roll* was not presented in Section 3.4, it should be pointed out that 25% of people who have never lived inside the isogloss for that form still report using it. While this could have a contact-based explanation, it might also reflect a process of speakers adopting the supralocal, nationwide-majority variant as they grow older and enter adulthood (in a well-documented process of age-grading, e.g. Wagner 2012). Tracking the use of *roll* in apparent time across the date of birth of our population sample, the results of which are plotted in Figure 3.4, we actually find a fairly consistent and linear increase in the self-reported use of *roll* over time. This is in fact more likely to reflect a gradual process of dedialectalisation with the encroachment of the supralocal variant and the levelling out of more traditional, localised forms (Trudgill 1999).

An important potential confound to consider in our results is the role of social class in shaping the relationship between childhood mobility and language use. Although our study focuses on mobility between the ages

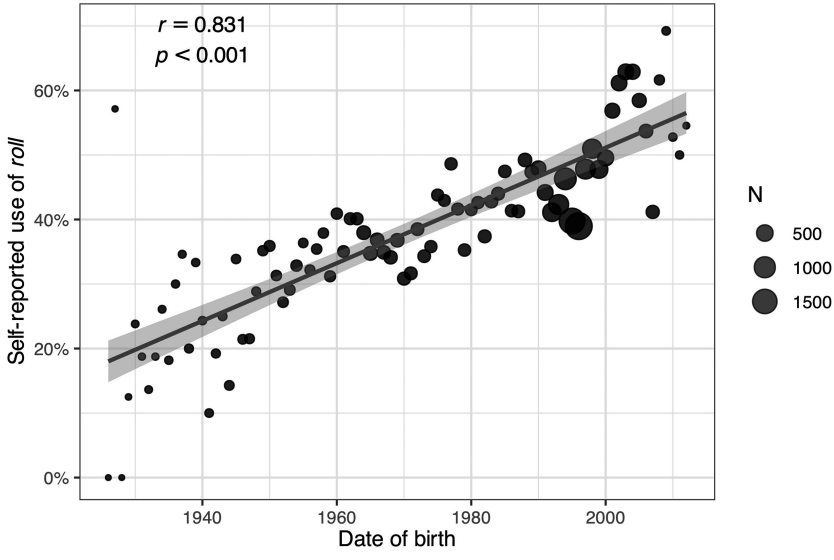

Figure 3.4 Rate of self-reported use of *roll* by date of birth (with linear regression trend line).

of 4 and 13, it is well documented that middle-class families are generally more likely to relocate for employment or other opportunities, resulting in higher geographic mobility than working-class families (Trudgill 1974, 2000; Britain 2011, 2013). Middle-class speakers (and their children) also exhibit more standardised or supralocal language use compared to their working-class counterparts. This raises the possibility that the observed effects of childhood mobility in our data could be partly attributable to social class differences: if mobile participants disproportionately come from middle-class backgrounds, their lower use of local forms might reflect broader social pressures favouring non-local variants, irrespective of their childhood movement. This kind of observation is born out from a phonological perspective by the reports of General Northern English (GNE; Strycharczuk et al. 2020) that educated young Northerners from various locations sound similar to one another.

At the same time, it could also be argued that mobility itself is a driving factor in the linguistic standardisation observed among middle-class speakers, given that frequent exposure to different linguistic norms may reduce allegiance to any single local dialect. However, it is unlikely to be the sole explanation, as not all middle-class individuals experience significant geographic mobility. This intersection of social class and mobility warrants further investigation, as our data do not currently include cleaned socio-economic background information on participants, although we do seek to address this in future studies.

A further point to consider is the accuracy of self-reported usage when participants complete this kind of dialect survey. The question wording clearly prompts respondents to select only the variants they actively use, rather than those they are simply familiar with. However, it is still possible that speakers over- or under-report their use of certain variants as a function of their salience and the speaker's own local pride, e.g. over-reporting their use of a variant they feel they *should* use as a member of the local speech community (see Smith & Durham 2011 and Jamieson 2020 on this kind of 'perceptual hyperdialectalism' in the Shetlands). This could be most prominent for variants that have a high degree of symbolic value as a marker of the local area, with these perceptual factors likely interacting in interesting ways. While we can acknowledge this as a possibility here, recent iterations of the survey do include attitudinal questions about pride in the local dialect/community, which will allow us to explore these questions more closely in future work.

We now turn to our third research question, relating to how lexically specific these patterns of acquisition might be. While we do overall see a remarkably consistent pattern of self-reported usage decreasing steadily and expectedly across the ALWAYS, MOSTLY, PARTIALLY, and NEVER groups, there *are* some variants that appear to pattern in subtly different ways

from each other. For example, the likes of *muffin*, *teacake*, *barm*, and *batch* show a much steeper drop-off in frequency from the MOSTLY to the PARTIALLY group. While it is difficult to pinpoint a single explanation for this, it could reflect the polysemous nature of many of these variants (Chambers 1992: 679): *batch* has a more general meaning, and *muffin* in particular is more commonly used throughout the UK to refer to a completely different food item. As a result, speakers might be more likely to stop using these variants (or not acquire them in the first place), and the effects of mobility might be registered more sharply, as we have seen.

The patterns displayed by *cob* are particularly interesting, as it stands as the only variant to show a large drop-off in frequency from the ALWAYS to the MOSTLY group, compared to many others where the ALWAYS and MOSTLY groups are much closer and, as just discussed, the dramatic drop-off is seen between the MOSTLY and PARTIALLY groups instead. This might suggest that *cob* is a variant more easily lost (or harder to acquire) for mobile speakers moving in/out of the area. In addition to the aforementioned polysemy consideration, a relevant factor might be the level of (perceived) awareness of these forms and enregisterment of the dialects to which they belong (Agha 2003). Northern dialects such as Scouse/Liverpudlian (spoken in Liverpool), Mancunian (spoken in Manchester), and Geordie (spoken in Newcastle) are heavily stereotyped in the media and firmly in the public consciousness (see Beal 2009; Montgomery 2012; Snell 2017 on the cultural and perceptual prominence of Geordie). This contrasts with perceptions of the East Midlands as a recognisable dialect region: Braber (2020: 75) points out how Nottingham, in the East Midlands, lacks this 'strong popular cultural notoriety or salience,' and in earlier perceptual work, it is shown that even people local to the area fail to recognise the accent and that the East Midlands more generally 'do not form an important region in the mental maps of people outside the area' (Braber 2015: 17). As a result, this relative lack of awareness and cultural prominence of East Midlands dialects could result in speakers being less likely to maintain the use of variants local to the region, such as *cob*, in cases of geographic mobility.

As outlined earlier in this chapter, there are some unavoidable limitations with the methods and data employed here, most notably the somewhat coarse operationalisation of mobility with the absence of detailed information on (a) how long speakers spent in each of the locations they report living between the ages of 4 and 13, (b) the order in which they moved between them, and (c) their residential history during adulthood. From a linguistic perspective, we also lack longitudinal self-report data that would allow us to pinpoint which variant(s) speakers were using at different points across their lifespans and connect that up with their residential history.

However, what detail the dataset might lack at the level of individual speakers is arguably overshadowed by the wide-ranging analysis it allows us to conduct. This pilot study has successfully shown how we can address questions about the acquisition of dialectal variants resulting from geographical mobility and dialect contact in a large-scale manner, covering ten regional variants of the same lexical variable and encompassing mobility and migration patterns of a large and diverse sample of speakers across the whole of the UK.

Note

1 It's worth noting that these children were interviewed by Chambers, himself a Canadian, and that perhaps their rate of British English variants might have been higher with a British interviewer. At the same time, it's also worth noting that the non-mobile British children used as controls in Chambers' study were also non-categorical users of local British variants, with occasional usage of Americanisms, which Chambers attributes to media and advertising. Still, these control subjects' usage rates of British variants exceeded those of the Canadian migrants across the board.

References

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23: 231–273. doi:10.1016/S0271-5309(03)00012-0.

Bates, Douglas; Mächler, Martin; Bolker, Ben and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48. doi:10.18637/jss.v067.i01.

Beal, Joan C. 2009. Enregisterment, commodification, and historical context: "Geordie" versus "Sheffieldish". *American Speech* 84(2): 138–156. doi:10.1215/00031283-2009-012.

Braber, Natalie. 2015. Language perception in the East Midlands in England: Investigating East Midlands adolescents' perception of language variation in the UK. *English Today* 31(1): 16–26. doi:10.1017/S0266078414000509.

Braber, Natalie. 2020. Nottingham: City of Literature – Dialect literature and literary Dialect. In: Patrick Honeybone and Warren Maguire (eds.), *Dialect writing and the North of England*, pp75–102. Edinburgh: Edinburgh University Press.

Britain, David. 2011. The heterogeneous homogenisation of dialects in England. *Taal en tongval* 63(1): 43–60.

Britain, David. 2013. Space, diffusion and mobility. In: J. K. Chambers and Natalie Schilling (eds.), *The handbook of language variation and change*, 2nd edition, pp471–500. Malden, MA: Blackwell. doi:10.1002/9781118335598.ch22.

Cambon, Jesse; Hernangómez, Diego; Belanger, Christopher and Daniel Possenriede. 2021. tidygeocoder: An R package for geocoding. *Journal of Open Source Software* 6(65): 3544. doi:10.21105/joss.03544.

Chambers, J. K. 1992. Dialect acquisition. *Language* 68(4): 673–705.

Foreman, Annik. 2003. *Pretending to Be Someone You're Not: A Study of Second Dialect Acquisition in Australia*. Doctoral dissertation, Monash University.

Jamieson, E. 2020. Viewing dialect change through acceptability judgments: A case study in Shetland dialect. *Glossa: A Journal of General Linguistics* 5(1): 19.

Kerswill, Paul. 1996. Children, adolescents, and language change. *Language Variation and Change* 8(2): 177–202.

Labov, William; Ash, Sharon and Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology, and sound change*. Berlin: Mouton de Gruyter. doi:10.1515/9783110167467.

MacKenzie, Laurel; Bailey, George and Danielle Turton. 2022. Towards an updated dialect atlas of British English. *Journal of Linguistic Geography* 10(1): 46–66.

Montgomery, Chris. 2012. The effect of proximity in perceptual dialectology. *Journal of Sociolinguistics* 16(5): 638–668.

Orton, Harold. 1962. *A survey of English Dialects: Introduction*. London and New York: Routledge.

Payne, Arvilla. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In: William Labov (ed.), *Locating language in time and space*, pp143–178. New York: Academic Press.

Pebesma, Edzer and Roger Bivand. 2023. *Spatial data science: With applications in R*. New York: Chapman and Hall/CRC. doi:10.1201/9780429459016.

Siegel, Jeff. 2010. *Second dialect acquisition*. Cambridge: Cambridge University Press.

Smith, Jennifer and Mercedes Durham. 2011. A tipping point in dialect obsolescence? Change across the generations in Lerwick, Shetland. *Journal of Sociolinguistics* 15(2): 197–225.

Snell, Julia. 2017. Enregisterment, indexicality and the social meaning of 'Howay': Dialect and identity in North-East England. In: Chris Montgomery and Emma Moore (eds.), *Language and a sense of place: Studies in language and region*, pp301–324. Cambridge: Cambridge University Press.

Strycharczuk, Patrycja; López-Ibáñez, Manuel; Brown, Georgina and Adrian Leemann. 2020. General Northern English. Exploring regional variation in the North of England with machine learning. *Frontiers in Artificial Intelligence* 3(48): 1–18. doi:10.3389/frai.2020.00048.

Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.

Trudgill, Peter. 1999. New-dialect formation and dedialectalization: Embryonic and vestigial variants. *Journal of English Linguistics* 27(4): 319–327. doi:10.1177/00754249922004741.

Trudgill, Peter. 2000. *The dialects of England*. Oxford: Wiley-Blackwell.

Wagner, Suzanne Evans. 2012. Age grading in sociolinguistic theory. *Language and Linguistics Compass* 6(6): 371–382.