Investigating geospatial models of the diffusion of morphosyntactic innovations: The Welsh strong second-person singular pronoun chdi

David Willis
Department of Theoretical and Applied Linguistics
University of Cambridge
dwew2@cam.ac.uk

Abstract
Morphosyntactic dialect variation, once a neglected area of dialect research, has recently witnessed a large growth in interest. Various methods from geospatial data analysis have been applied to morphosyntactic data. To date, the focus has largely been on analysing the distribution of stable patterns of variation. This article extends this work to examine patterns of ongoing change. It uses a body of data from the Syntactic Atlas of Welsh Dialects and the Siarad Corpus of spoken Welsh to examine the innovation and diffusion of a new second-person singular pronoun, chdi, testing the usefulness of Geographically Weighted Regression (GWR) as a method for identifying and modelling patterns of ongoing syntactic change. It is shown that GWR provides plausible models of the diachronic development of changes that are still in progress. Furthermore, it allows us to test whether rates of change are constant across geographical space, allowing us to test whether the Constant Rate Hypothesis (that diffusion of change proceeds at the same rate in different environments) holds between dialects.

1 Introduction
Morphosyntactic dialect variation, once a neglected area of dialect research, has recently witnessed a large growth in interest. Various large-scale projects aimed at mapping the geographical distribution of morphosyntactic variants have been undertaken, among them the Syntactic Atlas of Dutch Dialects (Barbiers et al., 2005, 2008), the Dialect Syntax of Swiss German project (Bucheli & Glaser, 2002; Glaser & Bart, 2011) and regional atlases in Germany, the Syntactic Atlas of Northern Italy (Benincà & Poletto, 2007), the Nordic Dialect Corpus (Johannessen et al., 2009), and the Syntactic Atlas of the Basque Language (Etxepare, 2009). These projects have renewed interest in applying geospatial statistical techniques to dialect data, but the focus has largely been on analysing the distributions of static syntactic features as found in traditional dialects (e.g. establishing dialect areas, correlating linguistic distance with various measures of geographic or perceptual distance etc.). Data on the current distribution of syntactic variants also has the potential to make a significant contribution to our understanding of dynamic variation, the diffusion of innovative syntactic systems, particularly if recent developments in geospatial data analysis can be fully exploited in the study of language change. Investigation of syntactic diffusion is still an underdeveloped area in comparison to the well-established tradition in phonology, and the extent to which models developed for phonology can simply be transposed to syntax has not been properly established. This article aims to contribute to our understanding of the diffusion of morphosyntactic innovations by applying Geographically Weighted Regression (GWR) as a method for identifying and modelling patterns of ongoing syntactic change. Geographical and age-related variation in the use of the innovative second-person singular pronoun chdi in northern varieties of spoken Welsh will be examined. We will develop a model of diffusion based on data from two current projects (the Syntactic Atlas of Welsh Dialects and the Siarad Corpus), corroborated against the background of independent historical data taken from nineteenth-century sources. GWR will be compared to other possible ways of analysing and modelling diffusion. It will be argued that GWR more successfully identifies a wave-like...
pattern of diffusion and provides a useful exploratory tool to guide further investigation, without bringing any significant disadvantages. We will also explore whether rates of change are constant across geographical space, as implied by a natural extension of the Constant Rate Hypothesis (CRH). While the data cannot offer conclusive evidence on this point, allowing rates of change to vary across geographic space provides no improvement in model fit, and thus we cannot disconfirm the Geospatial Constant Rate Hypothesis.

2 The S-curve model and the Constant Rate Hypothesis

We begin by outlining the dominant model of syntactic change as developed mainly in studies of historical rather than contemporary linguistic variation. It has long been recognized that linguistic change typically proceeds in a slow–quick–slow pattern. In studies that measure the frequency of an incoming linguistic variant over time, frequency generally rises only very slowly in the early stages, before a period of rapid frequency rise sets in. However, once the new variant is dominant, it struggles to oust the older variant completely, with a small residue of relic use typically being maintained for some time (G. Bailey et al., 1993:366). The result is the classic S-curve pattern of the diffusion of linguistic innovation (C.-J. N. Bailey, 1973:77–80; Osgood & Sebeok, 1954:155; Weinreich, Labov & Herzog, 1968:113–114). S-curves have been noted for the diffusion of such innovations as the spread of zero genitive plural forms for Russian nouns denoting units of measurement (Altmann et al., 1983), univerbation of the past particle + auxiliary ‘be’ in Polish (Andersen, 1990), the loss of preverbal objects in Middle English (Pintzuk & Taylor, 2009), the spread of new negative-polarity indefinite pronouns in Welsh (Willis, 2012:299–301), the spread of the new progressive passive in English (Denison, 2003:56–57; Hundt, 2004:63–64), among many others. Blythe & Croft (2012:278–80) consider 39 changes documented in the linguistic literature and conclude that 22 follow a full S-curve, 13 represent either the beginning or end of an S-curve, and 3 seem to be interrupted S-curves that go into reverse. Crucially they did not find any changes that seemed to follow a linear or exponential path.

The S-curve has generally been modelled in linguistics, as in population biology, with the logistic function (Altmann et al., 1983; Kroch, 1989):

\[
p = \frac{e^{k+st}}{1 + e^{k+st}} \quad \text{(Kroch, 1989:204)}
\]

where,
- \( p \) = the frequency of the innovation
- \( t \) = time
- \( s \) = the slope of the function
- \( k \) = the y-intercept, the frequency of the innovation at \( t = 0 \)
- \( e \) = Euler’s number (approx. 2.71828)

Items \( s \) and \( k \) are constants for each innovation scenario: \( s \), the slope of the function, is effectively the rate of change, while \( k \), the intercept, is the frequency of the innovation at some arbitrary point in time, \( t \), defined as the origin of the x-axis (\( t = 0 \)). Higher values of \( s \) produce steeper S-curves, while lower values produce gentler ones, as illustrated in Figure 1.
Figure 1. Effect of varying the slope constant \( s \) of an S-curve (for fixed \( k = 0 \)).

Higher values of \( k \) move the entire S-curve to the left, effectively locating the entire innovation earlier in time, while lower values of \( k \) move it to the right, delaying the change in the model. This is illustrated in Figure 2.

Figure 2. Effect of varying the intercept \((k)\) of an S-curve (for fixed \( s = 0.05 \)).

Since an S-curve is not a linear model, it is its overall shape and not the absolute increase in frequency of the innovating variant for any given point in time that is compared across contexts. The properties of a logistic curve can be more easily described and its parameters estimated if it is mapped to a straight line using the logit transformation as a link function:
The logit transforms for the S-curves given in Figure 1 are shown in Figure 3. There, it can more readily be seen that the curves differ in their rate of change.

\[
\ln \frac{p}{1 - p} = k + st
\]

Kroch (1989) investigated the properties of S-curves across different contexts for a number of changes that had been the subject of detailed quantitative studies, specifically, the replacement of *have* by *have got* in British English from 1700 to 1935; the spread of the definite article in possessive noun phrases in European Portuguese from 1750 to 1935; the loss of the verb-second constraint in Middle French from 1400 to 1700; and the spread of *do*-support in English from 1400 to 1700. In all of these cases, he found that the S-curves for the spread of the innovation in different grammatical contexts had slope parameters whose values could not be distinguished statistically. He concluded that these results were confirmation of the following hypothesis, termed the Constant Rate Hypothesis (or Effect):

when one grammatical option replaces another with which it is in competition across a set of linguistic contexts, the rate of replacement, properly measured, is the same in all of them. (Kroch, 1989: 200)

The set of contexts that change together is not defined by the sharing of a surface property, but a shared (abstract) grammatical structure. For instance, in the case of *do*-support, the change is defined by the loss of the verb’s ability to move to the inflectional head
(V-to-I movement), which is manifested in the spread of do-support in a number of different environments. This additionally means that we can use the slope parameter as a source of evidence for whether surface manifestations of language change should be attributed to a single, unified underlying change. Kroch noted that there was a sudden divergence in the behaviour of different contexts (clause types) for the spread of do-support in English. Up until around 1575, do-support spread in all grammatical contexts at the same rate (i.e. with S-curves whose slopes were indistinguishable) despite different absolute frequencies of use of do-support in the different contexts. However, after that date, one context, namely affirmative declarative clauses, diverged from the common development. Not only did the slope of the curve (rate of change) for this context fall; the frequency of do-support actually declined in this context, while all the other contexts continued their S-curves with the same slope as before. He concluded that this sudden reversal of the S-curve could only be explained if a major reorganization of the auxiliary system (specifically, the loss of V-to-I movement leading to the specialization of do-support to contexts other than affirmative declarative clauses) occurred around this time, interrupting the progression of the S-curve for affirmative declarative clauses (in effect applying a new but inverted S-curve reversing the change in that context only). Up until this point, then, change in all contexts was due to a single underlying change in the grammar (identifiable as a single slope value for all contexts), while two distinct changes were at work after this point (identifiable as two distinct slope values for change in different contexts).

3 Geographically Weighted Regression in dialect research

The S-curve model is in widespread use in research in historical syntax and, explicitly or implicitly, in most work on sound change (Labov, 1994), and will be adopted here. However, it has not been widely applied in dialectological work. C.-J. N. Bailey (1973) did indeed propose a model of diffusion in which linguistic innovations spread in wave-like fashion through geographic space and via S-curves at each location. This was presented as a theoretical model, but it has proven difficult to apply it to actual instances of dialect variation and change. This section outlines recent developments in geospatial data analysis that offer ways to apply the S-curve model to dialect data and integrate it with other approaches to diffusion.

There are many situations, within and beyond linguistics, where a process has a different effect at different geographical locations. Fotheringham, Brunsdon & Charlton (2002: 27–64) discuss as an example the impact of various independent variables on residential property prices in London in 1991. Using a global (i.e. city-wide) linear regression model, the impact of each variable can be estimated on a city-wide basis. At this level of analysis, the presence of a garage adds £5,956 to a property’s selling price. However, it is self-evident that the usefulness of a garage depends on the desirability of owning a car and that this may vary significantly between central locations served well by public transport and suburban locations with less extensive provision. There is thus a great deal of variation that is hidden within the global model: it overestimates prices for properties with garages in some areas and underestimates them in others, and thereby misrepresents the underlying process by which prices vary. Under such circumstances, a model that incorporates a geographical element (a ‘local’ model) is likely to perform much better than one which does not (a ‘global’ model).

It is this kind of spatial nonstationarity that Geographically Weighted Regression (GWR) was developed to explore. In a GWR analysis, a regression model is calculated at different geographical points across the region under investigation using data from within a moving window centred around the point under consideration. In this particular case, a geographically weighted regression produces local estimates across the whole city of the
amount added to a property’s price by the presence of a garage, ranging from under £1000 to over £12000. These estimates can themselves be mapped: in this case, high parameter values are found in upmarket central areas, where garages are scarce, and in distant suburbs, where the practical benefits of car ownership are greatest, while low values are found in a semi-circle of moderately affluent areas served well by public transport. Mapping the regression values therefore produces a geographic pattern for which an explanation can be sought. In this particular case, explanations readily present themselves in terms of public transport infrastructure and the practical benefits and difficulties in owning a car.

GWR has not previously been applied in the study of linguistic variation and change. However, dialect variation is precisely the kind of situation that the technique was designed to investigate. Linguistic variation is clearly spatially dependent in the sense that people who acquired a language in the same place tend to have grammars and lexicons that are similar, while people who grew up far apart are likely to have more different ones. The exact nature of the relationship between linguistic and geographic distance has been studied extensively, see, for instance, Séguy (1971) for lexical variation, Heeringa & Nerbonne (2001) for phonology, and Spruit (2008) and Szmrecsanyi (2012) for morphosyntax. These studies approach variation from a synchronic, static standpoint, but GWR offers another way to investigate language change by plotting S-curves across a region for a linguistic feature hypothesized to be undergoing change.

4 Innovation and diffusion in the Welsh pronominal system

Having set out the theoretical context, we now turn to consider the case study that will be the focus of the remainder of this article, namely variation and change in the form of the second person pronoun in northern varieties of Welsh over the last 150 years. We begin in section 4.1 by considering a fairly conservative spoken variety that lacks a number of recent innovations in the agreement system; these innovations, typical of younger speakers in the northwest, are set out in section 4.2; finally, section 4.3 brings us to our main focus, the new second-person singular strong pronoun chdi, setting it in its wider diachronic context.

4.1 The basic system

Full forms of Welsh personal pronouns do not manifest case distinctions. Instead, a central feature of the pronominal system is the distinction between strong and weak pronouns, two sets of pronouns found in distinct syntactic contexts. Strong and weak forms are identical in the plural (ni ‘we/us’, chi ‘you (pl.)’ and nhw ‘they/them’) and in the feminine third-person singular (hi ‘she/her’). However, in the remaining singular forms, there is a distinction. Since it is the second-person singular that is the subject of investigation here, the basic system will be illustrated from the first person and masculine third person. Here there is a distinction between a strong form with initial /v/ and a weak form without it: first-person singular fi /vi/ ‘I/me (strong)’ vs. i /i/ ‘I/me (weak)’, and third-person masculine fo /vo/ ‘he/him (strong)’ vs. o /o/ ‘he/him (weak)’. Forms given throughout are general northern spoken forms, which may differ significantly from spoken southern forms or literary Welsh.

Strong forms are used in contexts not associated with agreement. Most saliently, this includes ‘independent’ uses of the pronouns, for instance, in fragment answers in (3), in clause-initial focus position in (4), after dyma ‘here is’ in (5) and dyna ‘there is’, and after

---

1 Historically (and still in formal written Welsh), clitics distinguished accusative case from genitive, but this distinction has been given up entirely in speech, with accusative clitics now replaced by full pronouns.
other non-inflecting particles such as na ‘than’ in (6).

(3) a fi hefyd and me too ‘and me too.’ (Siarad corpus, fusser22)

(4) a fi sy ’n gorod wneud o and me be.PRS.REL PROG be.obliged.to.INF do.INF it ‘and it’s me that has to do it.’ (Siarad corpus, stammers6)

(5) so dyma fi ’n ffonio hi so here is me PROG phone.INF her ‘so here’s me phoning her / so then I phoned her’ (Siarad corpus, stammers7)

(6) ti licio teithio hyd ’n oed mwy na fi you like.INF travel.INF even more than me ‘You like travelling even more than me.’ (Siarad corpus, fusser14)

Additionally, strong pronouns are used after non-inflecting prepositions, such as efo ‘with’, gyda ‘with’, and â ‘with’:

(7) ac mae nhw ’n hapus i gyfarfod efo fi and be.PRS.3 they PRED happy to meet.INF with me ‘and they’re happy to meet with me.’ (Siarad corpus, davies15)

While this last type could synchronically just be treated as a subset of the non-agreeing contexts, it represents a significant context historically, and, for this reason, the object position of non-inflecting positions will be distinguished from the independent context in the rest of this article.

Agreeing contexts use weak pronouns. Traditionally, Welsh finite verbs agree with their subjects (Borsley 2009); most prepositions (except the non-agreeing ones in (7)) agree with their objects; and, according to the most prominent generative analyses (Awbery, 1976:23; Borsley, Tallerman & Willis, 2007:157–161; Sadler, 1988:78), nouns agree with their possessors. In all of these contexts, the relevant head shows agreement through a suffix (with verbs and prepositions) or through an agreeing proclitic (with nouns), and weak pronouns appear:

(8) es i nôl i mewn go.PST.1SG I back to in ‘I went back in.’ (Siarad corpus, fusser29)

(9) roedd Lana ’n cwyno wrtha i neithiwr be.IPFLV.3SG Lana PROG complain.INF to.1SG me last.night ‘Lana was complaining to me last night.’ (Siarad corpus, fusser 15)

2 Glossing throughout follows the Leipzig Glossing Rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php), except that IMPERS marks the impersonal forms of finite verbs, and PRT marks preverbal particles that appear immediately before verbs when an element is fronted, and in complement and relative clauses. Note also that REL marks the forms of finite verbs that appear in relative clauses, and PRED denotes the predicate marker yn, which appears before noun phrases and adjective phrases acting as predicates.
Two complications should be noted. Even in weak contexts, if the word immediately preceding the pronoun ends in a vowel, an epenthetic /v/ is often added to the front of the pronoun, yielding a result identical to the strong pronoun, compare (11) with (9) above.

(11) dyna ti ’n ddeu’ ’tha fi?
there.is you PROG say.INF to.1SG me
‘Is that what you’re telling me?’ (Siarad corpus, stammers5)

Secondly, a pronominal subject must immediately follow its verb and a pronominal object must immediately follow its preposition. This means that the word division between the two cannot easily be determined. Literary Welsh orthography represents all such combinations in ways that maximize the appearance of distinct agreeing forms. Hence, in its written form, amdanaf i /’am danavi/ ‘about you’ has agreement morphology that is clearly distinct from amdanat ti /’am danati/ ‘about you’. This implies an analysis of these as agreeing preposition + weak pronoun, hence /’am danav/ + /i/ and /’am danat/ + /di/ (with reduction of the geminate /td/ to /t/). However, another analysis is conceivable in which this pair does not have distinct agreeing forms, but a general pronominal form /’am dana/ followed by strong pronouns, hence /’am dana/ + /vi/ and /’am dana/ + /ti/.

4.2 Areas of ongoing change
As noted at the outset, the above description represents a somewhat idealized representation of conservative spoken northern Welsh. It probably accurately represents the only possibility for the distribution of forms across the whole of the north in the nineteenth century. However, in more advanced spoken northern Welsh, there is evidence that the second analysis set out above for agreeing heads is in fact the correct one. This means that the division between agreeing and non-agreeing contexts is more fluid, and some heads characterized as agreeing in the preceding discussion are in fact non-agreeing in some varieties.

The evidence for this comes in two forms. First, there is evidence from cases where the default third-person singular form is substituted in place of a clearly distinct agreeing form. Consider the paradigm for the modal DYLAI ‘should’, given in Table 1. The literary system given there clearly has agreeing form + weak pronoun throughout. In the northwestern dialects under consideration here, the vowel of the ending is levelled to /a/ in all persons by regular sound changes, namely, monophthongization of /ai/ > /e/ in the immediately post-tonic syllable followed by lowering of /e/ to /a/ in the same context; see Thomas (1972:176) for a synchronic formalization of these rules which mirrors the historical developments, and Awbery (2012) for a consideration of the historical evidence for the dating and ordering of these sound changes. The first of these sound changes affected all Welsh dialects, the second only northwestern ones. As we saw above, word division between the head and the pronoun is essentially arbitrary, so this levelling opens the way for reinterpretation of the system as involving a general non-agreeing form dyla /’daːla/ + strong pronoun, with the exception of the first-person singular, where it has agreeing dylwn /’daːlun/ + weak pronoun. The reality of this reanalysis can be confirmed by the appearance of the most innovative northwestern pattern, where even the agreeing first-person singular is replaced by the non-agreeing form dyla /’daːla/ + strong pronoun /vi/, as in the rightmost column of Table 1, or in the following attested example:
(12) **dyla fi** helpu chdi?
    should I help.INF you

‘Should I help you?’ (Siarad corpus, davies16)

<table>
<thead>
<tr>
<th></th>
<th>contemporary literary Welsh</th>
<th>conservative north-western spoken</th>
<th>advanced north-western spoken</th>
</tr>
</thead>
<tbody>
<tr>
<td>first sg.</td>
<td>dylwn i</td>
<td>dylwn i</td>
<td>dyla fi</td>
</tr>
<tr>
<td>second sg.</td>
<td>dylet ti</td>
<td>dyla ti</td>
<td>dyla ti</td>
</tr>
<tr>
<td>third sg.</td>
<td>dyla ti</td>
<td>dyla fo</td>
<td>dyla fi</td>
</tr>
<tr>
<td>third sg. masc.</td>
<td>dyla fi</td>
<td>dyla ti</td>
<td>dyla fi</td>
</tr>
<tr>
<td>third sg. fem.</td>
<td>dyla ti</td>
<td>dyla fo</td>
<td>dyla fi</td>
</tr>
<tr>
<td>first pl.</td>
<td>dyleny n</td>
<td>dyla ni</td>
<td>dyla ni</td>
</tr>
<tr>
<td>second pl.</td>
<td>dylech chi</td>
<td>dyla chi</td>
<td>dyla chi</td>
</tr>
<tr>
<td>third pl.</td>
<td>dyla nhw</td>
<td>dyla nhw</td>
<td>dyla nhw</td>
</tr>
</tbody>
</table>

Table 1. Paradigms of **dyla** ‘should’ in various Welsh varieties.

Such substitutions happen across the verbal and prepositional paradigms. A selection of examples is given in (13)–(16). In (13), **oedd** + strong pronoun *fi* ‘I was’ is substituted for inflected *oeddwn* + weak *i* (cf. third-person *oedd* o ‘he was’); in (14), *bysai* (*bysa* by regular sound change) + *fi* is substituted for inflected *byswn* + weak *i* (cf. third-person *bysa* fo ‘he would be’); in (15), *geitha* is a (presumably non-agreeing) re-formation based on third-person singular *geith* ‘he, she is allowed to’ in place of agreeing *ga* i ‘I am allowed to’; and, in (16), *gyynno* *fi* is a non-agreeing form (cf. *gynno* fo ‘with him’) in place of agreeing *gen* i ‘with me’.

(13) **oedd** fi ’im yn gwybod bod nhw wneud olwynion fel ’na was I NEG PROG know.INF be.INF they do.INF wheels like that ‘I didn’t know they made wheels like that.’ (Siarad corpus, stammers8)

(14) **you know**, **bysai** fi ’n gallu teithio o gwmpas you know be.COND I PROG be.able.INF travel.INF around ‘you know, I could travel around…’ (Siarad corpus, robert6)

(15) …*gofyn* iddo fo … gei … *geitha* fi … mynd arno fo? ask.INF to.3MSG him ca ca can I go.INF on.3MSG it ‘…ask him if I can go on it?’ (Siarad corpus, stammers2)

(16) mae ’r tocynnau ’n barod **gyynno** fi be.PRS.3SG the tickets PRED ready with me ‘I already have the tickets.’ (Siarad corpus, fusser21)

In many other cases, the reanalysis cannot have any easily evident surface morphological effects, because the historically agreeing form is either identical to the default form once the boundary reanalysis has occurred, or else differs only in terms of minor vowel differences that could easily be attributed to vowel reduction in fast speech. This is commonly the case in the second-person singular, where, most of the time, agreeing forms can readily be reanalysed as part of a non-agreeing paradigm, as can be seen with *dylet* ti ‘you should’ > *dyla* ti in Table 1. In one instance, though, the second-person singular does show a morphological reflex of the reanalysis. Many speakers have re-formed the future of *bod* ‘be’ from *byddi* /ˈbaði/ to *bydd* /biːd/, as in (17) (where it appears in its soft-mutated form *fydd*).

(17) ond … gobeithio **fydd** di fewn cyn Dolig but hope.INF be.FUT you in before Christmas ‘but hopefully you’ll be in before Christmas.’ (Siarad corpus, stammers 3)
4.3 The innovation of a new second person singular pronoun

Having looked at the context of the development of the pronominal system generally, we are now in a position to focus on the second-person singular. There is currently dialect variation in the form of this pronoun. Put simply, some dialects use ti as the form of this pronoun, while others use chdi. All southern dialects use ti, and will not be considered further. Chdi is found predominantly in the northwest, but its syntactic distribution is complex. Its use does not appear to be socially marked and it is not stigmatized. However, it is regionally marked and speakers are aware of it as being distinctive of their region. Speakers from outside of the area where it is found frequently misunderstand it, either mishearing it as chi ‘you (pl., formal)’ or making sense of it using the common but erroneous folk etymology that it is a ‘mixture’ of informal ti and formal chi, in both cases interpreting it as more formal than intended.

The question at issue here is the distribution, both geographical and syntactic, of the innovative form chdi, and the process by which it is diffusing through both of these media. The hypothesis that will be put forward is that syntactic variation must be understood against the background of change in the wider pronominal system set out above. This means that the most obvious analysis, namely one in which chdi is diffusing smoothly from one syntactic context to another, has to be rejected in favour of one that recognizes (at least) two different innovations.

This section sets out the historical evidence for the emergence of chdi. It is based on searching the following sources: the Historical Corpus of the Welsh Language (HCWL, http://people.ds.cam.ac.uk/dwew2/hcwl/menu.htm) (Willis & Mittendorf, 2004); Welsh-language materials in Early English Books Online (EEBO, http://eebo.chadwyck.com), Eighteenth Century Collections Online (ECCO, http://quod.lib.umich.edu/e/ecco), and Welsh Newspapers Online (Beta) (http://papuranewyddcymru.llgc.org.uk); along with selected additional plays and novels from the eighteenth and nineteenth centuries, as detailed in the list of primary sources. Collectively, these sources provide reasonable access to spoken registers of northern Welsh over the past 400 years. We will use this evidence to test whether the narrative derived from the GWR analysis of synchronic dialect data is plausible. The data in this section therefore represent real-time evidence derived from historical written sources, while we will turn to apparent-time evidence from present-data synchronic investigations in the next section.

<table>
<thead>
<tr>
<th></th>
<th>independent simple</th>
<th>independent reduplicated</th>
<th>reduced form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg.</td>
<td>mi /mi/</td>
<td>miui, myui /məˈvi/</td>
<td>fi /vi/</td>
</tr>
<tr>
<td>2sg.</td>
<td>ti /ti/</td>
<td>tidi, tydi /təˈdi/</td>
<td>di, ti /ti/</td>
</tr>
<tr>
<td>3sg. masc.</td>
<td>ef /ev/</td>
<td>efo /əˈvo/</td>
<td>fo /vo/</td>
</tr>
<tr>
<td>3sg. fem.</td>
<td>hi /hi/</td>
<td>hihi, hyhi /haˈhi/</td>
<td>hi /hi/</td>
</tr>
<tr>
<td>1pl.</td>
<td>ni /ni/</td>
<td>nini, nyni /nəˈni/</td>
<td>ni /ni/</td>
</tr>
<tr>
<td>2pl.</td>
<td>chwi /χwi/</td>
<td>chw(i)chwi /χwəˈχwi/</td>
<td>chi /χi/</td>
</tr>
<tr>
<td>3pl.</td>
<td>wynt(nt)/ui(nt)/</td>
<td>(h)wyntwy /uinˈtui/</td>
<td>nhw ˈnɨu/</td>
</tr>
</tbody>
</table>

Table 2. Independent (strong) forms of pronouns in Middle and Early Modern Welsh.

The beginning of the emergence of chdi as a new pronoun goes back to the early modern period. The ancestors of the modern strong pronouns are the reduplicated pronouns. While these are still present in very formal literary Welsh, in spoken Welsh they began to undergo phonological reduction from the fourteenth century onwards, ultimately merging with and/or replacing the independent simple series (Willis, 2007:439–450, 2009:136–137). The general pattern of reduction is shown in Table 2. However, after uninflected prepositions (mostly à ‘with’ and gyd à/gyda ‘(together) with’), na ‘than’ and a(g) ‘and’, all of which end
in /a/ or /aː/, a different pattern of reduction occurred in the first- and second-person singular, the difference evidently being due to phonological factors. In this context, the schwa of the first syllable is dropped, but the initial consonant remains, being resyllabified onto the vowel of the previous syllable. In the first-person singular, this produces, for instance, â /aː/ ‘with’ + /mvi/ ‘me’, as in (18).

(18) Onid eich tâd a anghywiodd a ’m fi…
FOC.Q 2PL father PRT cheat.PST.3SG with me
‘Was it not your father who dealt unjustly with me…?’ (1588 Bible, Genesis 31.7)

In the second-person singular, there is a further complication. By chance, these contexts happen all to trigger aspirate mutation, an otherwise relatively uncommon morphosyntactically triggered alternation on initial consonants. Aspirate mutation changes initial voiceless stops into fricatives, in this case changing the initial /t/ of tydi to /θ/. With vowel loss (syncope) and resyllabification, we end up with â /aː/ ‘with’ + /θdi/ ‘you’, as in (19). Full descriptions of the mutation system in Welsh can be found in Ball & Müller (1992) and Borsley, Tallerman & Willis (2007:19–26).

(19) Pe y baei δyn yn gwneuthur bargain a masnach ath di
if PRT be.PRFV.SBJV.3SG person PROG do.INF bargain and trade with.you
yn y llun hyn…
in the way this
‘If someone were to make a bargain and a deal with you in this way…’ (Robert Gruffydd, Y drych christianogawl, p. 21, 1585)

There are two ways to derive this system synchronically. We can either derive it in a way that mirrors diachrony, from underlying forms /məˈvi/ and /təˈdi/ with productive application of the processes just outlined (aspirate mutation, syncope and resyllabification). Alternatively, we can posit a grammar with exceptional strong forms of the first and second-person singular pronouns for use after uninflecting prepositions. The first approach makes sense as long as the full forms myfi and tydi surface in the spoken language, but as these forms themselves reduced in speech and became confined to writing, a spoken grammar of the second type must have emerged.

In the first-person singular, the form /mvi/ was lost in favour of the general reduced form fi /vi/. This is also the fate of the second-person singular form thdi /ðdi/ in some dialects, but, in parts of the north, it continued to develop. First, /ðdi/ spreads to other strong contexts which are not triggers for aspirate mutation, specifically, in (fronted) focus position, in (20), and after dyna ‘there is’, in (21). Once this has happened, it is clear that we are dealing with a distinct pronoun and not simply the result of a series of phonological rules. As regards dating this innovation, it seems that the earliest examples showing clear evidence of innovation are from the interlude plays of Thomas Edwards (Twm o’r Nant) (born 1739).

(20) Mae ’th di a gerir gore
FOC.COMP you PRT love.PRS.IMPERS best
‘…that it’s you who is loved best.’ (Thomas Edwards, Cybydd-dod ac oferedd 135.16, 1874 [1870])

(21) Dyna ’th di yn fîwl ddiddeunyydd
there.is you PRED fool useless
‘There you are, a useless fool.’ (Thomas Edwards, Y farddoneg Fabilonaidd 197.30, 1874 [1768])
A dissimilated form *chdi* /χdi/ is attested from the mid-nineteenth century (Morris-Jones, 1913:272) (presumably from speakers born in the early decades of the nineteenth century), in exactly the same range of exclusively strong contexts:

(22) Bobbach, cerdd i ’r society, da ch di…
Bob little go.IMP.2SG to the society good you
‘Little Bob, go to the society [religious meeting], there’s a good boy…’ (Yr Amserau, 24 February 1848, p. 4)

(23) ond os ch’di geiff y lle
but if you get.PRS.3SG the place
‘…but if it’s you that gets the place…’ (Lewis Lewis, Huw Huws, p. 15, 1860)

Again, these are all strong, non-agreeing contexts. Examples in these contexts are well attested in the second half of the nineteenth century in the dialogue of fiction set in the north and especially the northwest.

However, in the last two decades of the century, two more contexts are added and robustly attested, namely after *i* ‘to, for’, as in (24), and as the subject of nonfinite *bod* ‘be’, as in (25).

(24) …ond cofia y bydd yma gartra i chdi bob amsar
but remember.IMP.2SG PRT be.FUT.3SG here home to you every time
y leici di dwdad yma…
PRT like.PRS.2SG you come.INF here
‘…but remember there’ll be a home for you any time you’d like to come here…’
(Hunangofiant hogyn, Papur Pawb, 14 April 1894, p. 10)

(25) …mi fasa n well gin i weld dy gladdu di na chlywad
be.COND.3SG PRED better with me see.INF 2SG bury.INF you than hear.INF
bod chdi yn ’thrachu hefo nhw.
be.INF you PROG associate.INF with them
‘…I’d prefer to see you buried than to hear that you’re associating with them.’
(Llythyr Meri Jones, Y Werin, 24 November 1888, p. 4)

These are traditionally classified as weak contexts, but there are reasons to believe that they may have been recategorized as strong at this time in the relevant dialects. Recall that strong contexts are those without agreement between a head and a pronoun, while weak contexts are those with such agreement. The preposition *i* ‘to, for’ occupies an intermediate position, being semi-inflected, with inflected forms in the third-person singular and plural, but not elsewhere, as shown in Table 3. The literary forms are plausibly analysed as inflected with an optionally dropped weak pronoun. The first-person singular in literary Welsh is *im(i)*, plausibly inflected *im* + (optional) weak pronoun *i*. However, the spoken forms are more analytic, and, outside the third person, look to consist of an invariant preposition *i* plus a strong pronoun. If so, the spread of *chdi* to this environment is a reflex of its recategorization as a strong context.
|
|---|---|---|---|
| literary Welsh | spoken northern Welsh |
| sing | plur | sing | plur |
| 1sg. | im, imi | in, inni | i fi | i ni |
| 2sg. | itt, itti | iwich, ichwi | i ti, i chdi | i chi |
| 3sg. (masc.) | iddo (ef) | iddynt hwy | iddo fo | iddyn nhw |
| (fem.) | iddi (hi) | | iddi hi | |
| 1sg. | fy mod (i) | ein bod (ni) | bo’ fi | bo’ ni |
| 2sg. | dy fod (ti) | eich bod (chwi) | bo’ ti, bo’ chdi | bo’ chi |
| 3sg. (masc.) | ei fod (ef) | eu bod (hwy) | bod o | bo’ nhw |
| (fem.) | ei bod (hi) | | bod hi | |

Table 3. Paradigm of *i* ‘to, for’ in literary and spoken northern Welsh.

The same can be said of the subject of nonfinite *bod* ‘be’ in (25). Historically, and in literary Welsh, nonfinite *bod* agrees with its subject via agreement proclitics which trigger various mutations on the initial consonant of the verb itself. In spoken Welsh, a non-agreeing option is, however, also available, lacking both the proclitics and their mutation effects, see Table 4. In the first-person singular, the non-agreeing option requires a strong pronoun *fi*, suggesting this change represents a recategorization of the environment as strong. If so, then the spread of *chdi* to this environment in the second-person singular can be explained as a result of this recategorization, rather than as a result of a change in the properties of the pronoun itself.

|
|---|---|---|---|
| literary Welsh | spoken northern Welsh |
| sing | plur | sing | plur |
| 1sg. | fy mod (i) | ein bod (ni) | bo’ fi | bo’ ni |
| 2sg. | dy fod (ti) | eich bod (chwi) | bo’ ti, bo’ chdi | bo’ chi |
| 3sg. (masc.) | ei fod (ef) | eu bod (hwy) | bod o | bo’ nhw |
| (fem.) | ei bod (hi) | | bod hi | |

Table 4. Paradigm of *bod* ‘(nonfinite) be’ in literary and spoken Welsh.

*Chdi* is attested sporadically in certain other syntactic contexts from the start of the twentieth century: after an inflected preposition, in (26), and as the subject of an inflected auxiliary, in (27). These are, in principle, agreeing contexts requiring a weak pronoun, but they are also contexts where reanalysis of the paradigm as non-inflected, as discussed in section 4.2 above, is conceivable.

(26) Bedi y mater *rhyngog chdi* a Neli rwan, Bob?

‘What’s the matter between you and Neli now, Bob?’ (*Papur Pawb*, 10 August 1901, p. 2)

(27) Robet, wyt ti yn meddwl mewn difri *y basa chdi*,

‘Robert, do you seriously think that you, for instance, would happily pay to have me buried?’ (*Darllen y papur*, *Papur Pawb*, 26 October 1901, p. 14)
The overall timetable of innovation attested in different syntactic contexts in the nineteenth century is shown in Table 5. According to the narrative developed in this section, this proceeds via two distinct processes. An initial reanalysis creates a new strong pronoun \( \text{thdi} \) by the second half of the eighteenth century, with a phonological change (dissimilation) creating \( \text{chdi} \) in the mid-nineteenth century. The second process is loss of agreement in various syntactic contexts, underway from the final quarter of the nineteenth century onwards, which recategorizes these contexts as strong, indirectly leading to the spread of \( \text{chdi} \) to these contexts too.

This section has set out a real-time narrative of the development of \( \text{chdi} \) based on written historical sources. We now turn to contemporary, synchronic dialect data to establish whether they tell an apparent-time story that is compatible with the real-time evidence. In doing so, we will be testing GWR as a tool for investigating the diffusion of morphosyntactic change.

### 5 Data sources

Data for the apparent-time analysis come from two sources: the Bangor Siarad corpus of spoken Welsh and the pilot project of the Syntactic Atlas of Welsh Dialects (SAWD). The aim was to establish the contexts in which each speaker’s grammar makes \( \text{chdi} \) available. The contexts chosen and the forms that would be available in an innovating grammar are given in (28). Evidence from the independent real-time analysis presented in section 4.3 suggested that these run broadly in the historical chronological order listed there.

\[
(28)
\begin{align*}
\text{efo chdi} & \quad \text{‘with you’} \\
\text{i chdi} & \quad \text{‘to, for you’} \\
\text{inflected preposition: } & \quad \text{tha chdi} \quad \text{‘to you’}, \text{amdana chdi} \quad \text{‘about you’} \\
\text{dyla chdi} & \quad \text{‘you should’} \\
\text{oedda chdi} & \quad \text{‘you were’} \\
\text{rhai(d) chdi} & \quad \text{‘you must’} \\
\text{(bua)sa chdi} & \quad \text{‘you would’} \\
\text{by(dd) chdi} & \quad \text{‘you will’} \\
\text{gynno chdi} & \quad \text{‘with you’}
\end{align*}
\]

The Bangor Siarad corpus (Deuchar et al., 2014) is a 460,000-word corpus of spoken (largely northern) Welsh, consisting of transcriptions of 69 recordings of conversations in which 151 Welsh speakers participated, collected from 2005 to 2007. Instances of second-person singular pronouns were extracted from the corpus to establish the grammar of each speaker. Of the total 151 speakers, only 139 produced second-person singular pronouns, as a few conversations were conducted using formal pronouns. Given that our interest is in spatial
variation, each speaker was localized to the location where they had lived for at least 10 years before the age of 18. On this basis, 20 speakers could not be localized (having moved too much during childhood) and were removed from the database. Finally, since it is known that chdi is found only in northern dialects, and to facilitate comparison with the SAWD data, the 15 speakers who grew up in south Wales were also removed. For each context, each speaker was categorized as (i) having chdi in this context; (ii) not having chdi in this context (i.e. using only ti); or (iii) not having provided relevant data (i.e. the speaker did not use the context during their recorded speech). A speaker was counted as having chdi if they used it in the relevant context at least once. If they produced the context, but never used chdi, they were counted as a ti-only speaker for that context. Speakers who did not produce an instance of a given context were removed from the analysis for that context.

The pilot project of the Syntactic Atlas of Welsh Dialects (SAWD) interviewed 98 speakers from north Wales. They were asked to repeat sentences, changing them so to the way that they would say them naturally in everyday speech. Of the 60 questionnaire sentences, 44 contained a second-person singular pronoun in some way or other. Each of the 10 contexts in (28) was tested using at least two sentences, with both the innovative option with chdi and the conservative form with ti being offered in one sentence. Speakers were localized using the same procedure as with the Siarad corpus. This time, 3 speakers could not be localized and were removed. Each speaker was categorized for each context into one of the three same categories as before. A speaker was counted as having chdi if they added it to at least one sentence where it was not offered by the fieldworker. If they removed it from at least one sentence where it was offered and did not add it to any sentence where it was not offered, they were counted as lacking chdi for the relevant context. Speakers meeting neither of these conditions were counted as not having provided relevant data. This could arise either if a speaker paraphrased sentences in a way that eliminated the need for a pronoun or because they repeated the fieldworker faithfully in all cases where the context arose. While this last consideration did not occur in more than a handful of instances, it was included to ensure that the data reliably reflect active competence rather than mere passive comprehension or a desire to please the fieldworker.

The result of these procedures is effectively to create a competence grid describing the grammars of the speakers for each context. Each speaker is represented once only in the data for each context, and the data points for each context are therefore independent of one another. The grid therefore expresses the diffusion of innovations through the grammars (competence) of individuals in the population, not increases in frequency of use (performance) of the innovative form. The y-axis of our S-curves will therefore express the proportion of the population whose grammars allow the innovation. Other conceptions are of course possible, as Denison (2003:59–62), for instance, notes. Our consideration in section 1 above focused on frequency of use throughout the population, an approach often adopted in studies based on historical materials or in phonological studies based on interview recordings. S-curves have also played a prominent role in lexical diffusion (Chen, 1972:475; Cheng & Wang, 1977), where the y-axis represents the proportion of the lexicon affected by an innovation.

These sources provide apparent-time data for ongoing change. We assume that data provided by speakers today reflects the system that they acquired as children. If we see differences between older and younger speakers today, this is to be interpreted as evidence of change in progress rather than a stable, recurrent pattern of speakers’ grammars changing as they age. The variant used by younger speakers is diffusing and will eventually come to supplant the older variant in that location. In this way, change in progress can be identified via a synchronic snapshot of variation within a community in the form of an apparent-time study. This view also entails that, if a variant diffuses to a new location, it does so by being
adopted by children, thereby initiating age variation that will eventually develop into language change. Such an assumption is widely made in studies of ongoing change and has stood up to scrutiny when tested (G. Bailey et al., 1991).

Note that ‘diffusion’ here is being used to mean ‘spread of an innovation’, a fairly neutral sense commonly adopted in the literature, for instance, by Wolfram & Schilling-Estes (2003) or Britain (2013). It is not intended to be understood in the sense of Labov (2007, 2010), who distinguishes transmission of an innovation via child language acquisition with concomitant maintenance of structural constraints from diffusion of an innovation by adult language (or dialect) acquisition accompanied by failure to preserve structural constraints. For Labov, all of the changes considered in the current article would be instances of transmission.

6 A non-geographic model

An initial global (not geographically weighted) logistic regression was carried out for each of the contexts using RStudio, with age as the independent variable. The results are given in Table 6.

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Age coef.</th>
<th>standard error</th>
<th>p-value of age coeff.</th>
<th>signif.</th>
<th>AIC</th>
<th>n</th>
<th>% deviance acc. for</th>
</tr>
</thead>
<tbody>
<tr>
<td>efo ‘with’</td>
<td>0.923</td>
<td>0.024</td>
<td>0.013</td>
<td>0.057</td>
<td>.</td>
<td>127.18</td>
<td>118</td>
<td>3.0%</td>
</tr>
<tr>
<td>independent</td>
<td>0.911</td>
<td>0.020</td>
<td>0.011</td>
<td>0.074</td>
<td>.</td>
<td>154.78</td>
<td>138</td>
<td>2.2%</td>
</tr>
<tr>
<td>i ‘to’</td>
<td>0.442</td>
<td>0.031</td>
<td>0.011</td>
<td>0.004</td>
<td>**</td>
<td>167.03</td>
<td>139</td>
<td>5.2%</td>
</tr>
<tr>
<td>infl. prep.</td>
<td>0.294</td>
<td>0.033</td>
<td>0.012</td>
<td>0.004</td>
<td>**</td>
<td>166.05</td>
<td>134</td>
<td>5.3%</td>
</tr>
<tr>
<td>dyla ‘should’</td>
<td>-0.107</td>
<td>0.027</td>
<td>0.012</td>
<td>0.039</td>
<td>*</td>
<td>101.19</td>
<td>74</td>
<td>4.5%</td>
</tr>
<tr>
<td>oedd ‘were’</td>
<td>-0.595</td>
<td>0.053</td>
<td>0.011</td>
<td>0.000</td>
<td>***</td>
<td>178.51</td>
<td>146</td>
<td>13.7%</td>
</tr>
<tr>
<td>rhaid ‘must’</td>
<td>-0.766</td>
<td>0.053</td>
<td>0.013</td>
<td>0.000</td>
<td>***</td>
<td>136.66</td>
<td>111</td>
<td>13.7%</td>
</tr>
<tr>
<td>(bua)sa ‘would’</td>
<td>-1.207</td>
<td>0.053</td>
<td>0.011</td>
<td>0.000</td>
<td>***</td>
<td>177.67</td>
<td>149</td>
<td>13.5%</td>
</tr>
<tr>
<td>bydd ‘will’</td>
<td>-1.954</td>
<td>0.046</td>
<td>0.020</td>
<td>0.024</td>
<td>*</td>
<td>67.78</td>
<td>64</td>
<td>8.5%</td>
</tr>
<tr>
<td>gan ‘with’</td>
<td>-2.236</td>
<td>0.032</td>
<td>0.013</td>
<td>0.011</td>
<td>*</td>
<td>136.63</td>
<td>163</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

Significance codes: *** p < 0.001, ** p < 0.01, * p < 0.05, . p < 0.1

Table 6. Results of global logistic regression of the distribution of chdi.

Speakers’ years of birth were recalibrated relative to 1950 (e.g. 1945 becomes -5, and 1970 becomes +20), so that the intercept shows the extent to which the change had progressed by that time, and can therefore be used as a guide to the relative date of onset of change for the different contexts. The age coefficient is the slope of the logistic curve, measured in logits per year. This procedure follows that of Kroch (1989:225), who fixes the zero point in time for his analysis of the diffusion of do-support in English to 1350. This produces more meaningful intercept values, since the intercepts represent a measure of the degree of progress of the change at the zero point, and 1950 is a more meaningful point of reference for the current changes than 0 AD. For instance, the estimated intercept value of 0.923 for efo ‘with’ in Table 6 equates to a global probability of 0.72 (p = exp(k)/1+exp(k)) that a speaker from the sample born in 1950 will use efo chdi.

The AIC (Akaike information criterion) shown in Table 6 is a measure of model performance that takes into account model complexity. Since a geographically weighted logistic regression is a more complex model than a global logistic regression, it is bound to account for more of the data. The AIC penalizes more complex models, hence it provides a way to evaluate whether the added complexity of the GWR model was justified by the improvement in fit. Models with smaller AICs for a given data set are deemed to provide
better data fit, and hence it will be possible to compare the AIC values in Table 6 with those for geographically weighted models introduced in section 8 below.

To test if social variables other than age were relevant, a further logistic regression was carried out adding gender as an independent variable. Gender proved to be statistically significant only in the case of oedd ‘were’, with female gender favouring use of chdi (p = 0.039). There was no obvious evidence of social-class variation in the data, and, given that previous studies have been skeptical about the existence of such variation in Welsh (Ball, 1988:74–75; Jones, 1998:49; Williams, 1987), neither class nor gender was pursued further. Absence of social conditioning suggests that prestige is unlikely to be the fitness bias driving the S-curve in this case, contra Croft (2000:8), who argues that fitness biases are always social (‘the mechanisms for propagation … are essentially social’, cf. also the discussion by Blythe & Croft, 2012:272–273). If the account sketched above (section 4.3) is correct, the bias is ease of acquisition, with the new analysis in both the shifts proposed there being more easily acquired against the background of the prevailing grammatical system than its predecessor which fails to be acquired.

Table 6 tells a story that is in keeping with what we know of the historical background to the changes, at least in terms of their relative order. The two contexts in which chdi is found in the mid-nineteenth century have high intercept values, indicating a more advanced change and likely earlier actuation. They also have low age coefficients that struggle to reach statistical significance. Next come two prepositional contexts, inflected prepositions and the semi-inflecting preposition i ‘to’, with somewhat lower intercepts and somewhat higher age coefficients, indicating more recent innovation and faster ongoing change. The object of preposition i was identified as a relatively early context for change in section 4.3 above, and the high intercept value here is consistent with that. These are followed by a group of essentially verbal contexts, involving tense auxiliaries and modals. These have intercepts spread over a considerable range, indicating successive spread of chdi to the different verbs at different times; however, the age coefficients are remarkably similar, perhaps indicating a single, common rate of change. Finally, gan ‘with’ seems to be a special case, showing a very low intercept, indicating a very recent innovation, and a lower age coefficient, more consistent with the other prepositional contexts than with the verbal ones.

The change with efo ‘with’ and independent chdi appears to have run to near completion, with perhaps only marginal ongoing diffusion. For all the other contexts, we can be confident that ongoing change is present at the present day.

While a global logistic regression of the type in Table 6 successfully establishes the order of innovations, it gives no indication of the geospatial processes involved in diffusion of change. It also fails to provide a good model of the attested data, accounting for rather a small percentage of deviance, vastly overpredicting the use of chdi in areas outside of the dialect area where the change is found, and underpredicting its use within this area.

### 7 An area-based model

The global model’s inadequacy is largely due to its failure to incorporate any geographic element. To overcome this, we need to introduce a geographical factor into the regression. In this section, we do this in fairly traditional style, by dividing the region under investigation into districts, adding the district where a speaker is from as a variable in the regression. The choice of district boundaries is, of course, essentially arbitrary, and this in itself represents a weakness of this type of approach. Welsh local government administrative districts are too large to be part of an effective model of variation and change, and are particularly arbitrary, since there is no reason a priori to expect patterns of variation and change to coincide with these boundaries. A more promising type of approach involves the travel-to-work areas (TTWAs) of the England and Wales census (Coombes & Bond, 2008). TTWAs, used
extensively in labour-market analysis, are defined as areas where:

(i) ideally, at least 75% of the resident economically active population work in the area;
(ii) ideally, at least 75% of everyone working in the area live in the area; and
(iii) a minimum working population of 3,500 is present.

Map 1. Travel-to-work areas for 1981 in north Wales.

The 1981 England and Wales census defined 11 such areas in north Wales, shown in Map 1. These have the advantage that they previously been used in dialectological work (Buchstaller & Alvanides, 2013), and, being based on commuting patterns, might be expected to correlate, in part at least, with stable patterns of variation. The 1981 definitions were the earliest currently available, and were chosen over later ones because they are likely to be closer to the patterns experienced when the average informant was growing up.

To test the usefulness of this approach in modelling geographical diffusion, consider its application for three of the syntactic contexts listed above, namely after efo ‘with’, after an inflected preposition, and as the subject of (bua)sa ‘would (be)’. In each case, a logistic regression was performed with year of birth (recalibrated relative to 1950 as before) and travel-to-work area as predictors. The results are shown in Table 7 and mapped in Maps 2–4: the piecharts show the raw data, while the colour of each travel-to-work area identifies the first birth cohort to have majority chdi usage in that area (dark blue if chdi never reaches a majority).
### Table 7. Results of an area-based logistic regression for the distribution of *chdi*.

<table>
<thead>
<tr>
<th>context factor</th>
<th><em>efo</em> ‘with’ coefficient</th>
<th><em>inflected prepositions</em> coefficient</th>
<th><em>(bua)sa</em> ‘would (be)’ coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>p =</em></td>
<td></td>
<td><em>p =</em></td>
</tr>
<tr>
<td>intercept value</td>
<td>2.068</td>
<td>1.726</td>
<td>-0.485</td>
</tr>
<tr>
<td>year of birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(logits per year)</td>
<td>0.018 0.00</td>
<td>0.051 0.00</td>
<td>0.058 0.00</td>
</tr>
<tr>
<td>Holyhead</td>
<td>17.084 1.00</td>
<td>-0.252 0.83</td>
<td>-0.162 0.80</td>
</tr>
<tr>
<td>Pwllheli</td>
<td>17.243 1.00</td>
<td>-1.595 0.04</td>
<td>-1.929 0.02</td>
</tr>
<tr>
<td>Porthmadog</td>
<td>17.133 1.00</td>
<td>-1.662 0.09</td>
<td>-0.808 0.30</td>
</tr>
<tr>
<td>Wrexham</td>
<td>-1.787 0.09</td>
<td>-1.959 0.15</td>
<td>-1.744 0.16</td>
</tr>
<tr>
<td>Llandudno</td>
<td>-1.273 0.19</td>
<td>-2.434 0.00</td>
<td>-1.786 0.05</td>
</tr>
<tr>
<td>Rhyl</td>
<td>-3.421 0.01</td>
<td>-3.406 0.00</td>
<td>-1.907 0.12</td>
</tr>
<tr>
<td>Denbigh</td>
<td>-3.499 0.00</td>
<td>-4.671 0.00</td>
<td>-18.893 0.99</td>
</tr>
<tr>
<td>Machynlleth</td>
<td>-1.960 0.20</td>
<td>-19.495 1.00</td>
<td>-17.814 1.00</td>
</tr>
<tr>
<td>Dolgellau</td>
<td>-3.654 0.00</td>
<td>-19.847 0.99</td>
<td>-18.605 0.99</td>
</tr>
<tr>
<td>Welshpool</td>
<td>n/a</td>
<td>n/a</td>
<td>-19.549 1.00</td>
</tr>
</tbody>
</table>

The speed of change (the slope of the S-curve) is assumed to be uniform across the whole region for a given context. The intercept values are calculated for the Bangor and Caernarfon travel-to-work area, and the coefficients for other areas represent divergence from this value. Since the Bangor and Caernarfon TTWA tends to be the most linguistically innovative, area coefficients are mostly negative. The intercept value of 2.068 for *efo* is equivalent to saying that a given speaker born in the Bangor and Caernarfon TTWA in 1950 has a 0.89 probability of using *chdi* in this context. The coefficient for the Llandudno TTWA is -1.273 relative to the Bangor and Caernarfon TTWA, hence the intercept value for the Llandudno TTWA for 1950 is 0.795 (2.068 - 1.273), which corresponds to a probability of 0.69 that a given speaker will use *chdi*. 
The age coefficient, in conjunction with the intercept value, can be used to establish when a majority of acquirers have *chdi* in their grammars. For instance, the intercept value of -0.485 for *chdi* after (*buas*) ‘would (be)’ in the Bangor and Caernarfon TTWA for speakers born in 1950 corresponds to a probability of using *chdi* of only 0.38. However, at a rate of increase of 0.058 logits per year, the change will reach a frequency of 0.50 after 8 years. The area is therefore classified in Map 4 as reaching majority *chdi* competence in this context by 1960. The remaining maps are plotted on this same basis (for fuller explanation, see section 8.1 below).
Map 3. Plot of intercept values for a travel-to-work-area model of the diffusion of *chdi* ‘you’ after inflected prepositions.

Compared to the global regression, this approach has a number of advantages. The area-based model accounts for much more of the data (for *efo*, null deviance = 126.90, residual deviance = 73.58, 42.0% of deviance account for; for inflected prepositions, null deviance = 171.00, residual deviance = 98.99, 42.1% of deviance account for; and for *bua(sa)*, null deviance = 200.90, residual deviance = 138.00, 31.3% of deviance accounted for), cf. Table 6 above. It is of course somewhat more complex than the global regression, but, even allowing for this, the model performance is far superior, with AIC values of 95.58 for *efo*, 123.00 for inflected prepositions, and 162.00 for *bua(sa)* (cf. the substantially higher values for the global model in Table 6).

Both approaches are consistent with the conclusion that variation with *efo* is stable, while change is underway in the other two contexts. For *efo*, the age coefficient of 0.018 logits per year is slightly lower than that produced by the global model (0.024 logits per year), but this is not statistically significantly different from zero in either case, giving us little reason to reject the null hypothesis that there is no change in progress. Both the other contexts produce statistically significant age coefficients, confirming change in progress, with the area-based model producing higher rates of change than the global model.
In addition to offering superior data fit, an area-based regression offers some insight into the geospatial processes involved in diffusion. With efo, in Map 2, it shows a sharp divide between the five northwestern TTWAs, where diffusion of chdi has already run to completion, and the three TTWAs to the south and east, where it has not even begun, suggesting a sharp, diachronically stable isogloss. However, we cannot be certain that the isogloss really runs along the TTWA boundaries, as the relevance of these boundaries is presupposed in the model. Discussion of the status of the intermediate Wrexham TTWA will be postponed until the geographically weighted model is introduced in section 8 below.

After inflected prepositions, in Map 3, this model reveals a more subtle pattern of diffusion, with change initiating in the Bangor TTWA, then spreading north, south and east in that order. Once allowance is made for this diffusion, the same contrast between the five northwestern TTWAs emerges as before. Again though, while the fact of diffusion emerges fairly clearly, it is unclear whether the order of diffusion reflects a staged process of diffusion or whether it is merely an artefact of the size and shape of the TTWAs.

Much the same issues arise with the much more recent diffusion of chdi after (bua)sa ‘would be’ in Map 4. The same pattern of diffusion from the Bangor TTWA northwards, then southwards then eastwards emerges, albeit at a later date and bypassing the Pwllheli TTWA to the southwest almost entirely. But once again, the TTWA boundaries are presupposed, and prevent us from seeing inside them to look at the diffusion process itself.

8 A GWR model
We have seen that, while it offers an improvement over a global model, an area-based model, with speakers spread continuously across geographic space, dividing speakers into geographic groups based on administrative or other geographical divisions is an essentially arbitrary process. In the current instance, it would make sense to remove speakers from areas that are not and will not undergo change from the analysis, but we do not know in advance what those
areas are. In this section, we consider a geographically weighted regression that offers a way to circumvent these problems.

8.1 Stable variation
First we will consider the application of geographically weighted regression to dialect variation in changes that have already been substantially completed, leading to something close to stable dialect variation. The global regression suggested that this was the case for two contexts: after invariant prepositions such as *efo* ‘with’ (confirmed in the area-based regression) and in independent (non-agreeing) syntactic environments. The method will be set out in full for the first of these cases, before we investigate what it can tell us about the underlying processes of linguistic innovation and diffusion.

8.1.1 Invariant prepositions: *Efo chdi* ‘with you’
We begin with replacement of *efo ti* ‘with you’ by *efo chdi*, which, as we saw in section 4.3 above, is demonstrably the historically primary context for innovation. The global logistic regression performed above assumes that the intercept does not vary geographically, the innovation having begun everywhere simultaneously. A more realistic model is clearly one in which the intercept varies geographically, modelling the fact that diffusion of an innovation simply means that it reaches some geographical locations earlier than others. For the moment, we set the age coefficient globally, assuming constant rate of change (we will briefly explore the effect of allowing this to vary in section 8.3 below). This is implemented via a geographically weighted logistic regression, performed separately using the GWR4 software package (Nakaya, 2014 [2009]) for each of the ten syntactic contexts identified above. This produces a ‘mixed’ GWR model, one where some coefficients are set globally, but others locally (Brunsdon, Fotheringham & Charlton, 1999).

In GWR, a regression is performed for a given geographical location (the regression point) considering only data inside a defined geographical window, known as the spatial kernel. Within the spatial kernel, data points are weighted using a weighting function that takes into consideration how far away the data point is from the regression point: closer data is weighted more heavily than distant data. The radius of the kernel is referred to as the bandwidth. The bandwidth may be a fixed distance (a fixed bandwidth) or may be defined so as to vary in size in order to include a specific number of data points (an adaptive bandwidth). A large bandwidth smooths out the data, limiting the impact of atypical values. Too large a bandwidth will eliminate too much genuinely local variation, while too narrow a bandwidth will wrongly model random variation as geographically motivated (Brunsdon et al., 1999:217–220; Fotheringham et al., 2002:44–47, 211–213). Various methods have been proposed to identify an optimal bandwidth, obviating these two dangers, for a given regression. For the current study, a fixed bandwidth was used, with optimal bandwidth selection performed using GWR4’s Golden Search procedure. This identifies the bandwidth that will yield the model with the best fit to the data (lowest AIC). In the case of *efo chdi*, the regression was performed with a bandwidth of 11.9 km. The bandwidth established by this method varied somewhat for the other syntactic contexts, but ranged only between 10.2 and 17.7 km. Following common practice for fixed bandwidths, a Gaussian weighting function was used to weight the contribution of individual data points to the regression (Gollini et al., 2013).

A mixed GWR model produces geographically variable values for local parameters, in this case the intercept. The slope of the S-curve is set globally via the age coefficient, this time estimated as 0.0198, slightly lower than the estimate of 0.024 produced by the global model in Table 6 above. As before, the intercept values are plotted for each data point, see Map 5, so as to produce meaningful statements about the progress of diffusion. Given that the
The origin of the x-axis (time) is set to 1950, an intercept value of zero means that by 1950, the probability of a newly born speaker acquiring a grammar with *efo chdi* had reached 0.5 in that location, and that, thereafter, such speakers were in a majority. Given a global age coefficient of 0.0198 logits per year, an intercept value of 0.198 would indicate a location that had reached this stage 10 years earlier in 1940. A full interpretation of the intercept values for this example is given in Table 8.

<table>
<thead>
<tr>
<th>Majority diffusion year</th>
<th>Minimum intercept value</th>
<th>Majority diffusion year</th>
<th>Minimum intercept value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>-0.793</td>
<td>1940</td>
<td>0.198</td>
</tr>
<tr>
<td>1980</td>
<td>-0.595</td>
<td>1930</td>
<td>0.397</td>
</tr>
<tr>
<td>1970</td>
<td>-0.397</td>
<td>1920</td>
<td>0.595</td>
</tr>
<tr>
<td>1960</td>
<td>-0.198</td>
<td>1910</td>
<td>0.793</td>
</tr>
<tr>
<td>1950</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8. Interpretation of intercept values for *efo* (age coefficient = 0.0198 logits per year).

Plotting the intercept values as in Map 5, then, represents the model’s estimate of the date by which the innovation had diffused or will diffuse among children to a given geographical location.

Map 5. Plot of intercept values for GWR model of the diffusion of *chdi* ‘you’ after *efo*.

The mixed GWR model provides a large improvement of fit as compared to the global model and a small improvement compared to the area-based one, accounting now for 42.9% of the deviance (null deviance 126.93, residual deviance 72.42), with an improved AIC of 91.44. The (global) age coefficient remains low, at 0.020, and is not significant (standard error 0.016, z value 1.24, p = 0.217), still consistent with stable variation.

This model presents a plausible narrative of the progress of diffusion. The
historical background to the change indicates that it was well underway in the nineteenth century, suggesting that it may well have now run to completion. The intercept values can be mapped to create an isogloss dividing a northwestern area with a positive intercept value (areas with a probability greater than or equal to 0.50 that a speaker born in 1950 uses *efo chdi*) from areas to the south and east with a negative intercept value (a probability less than 0.50). The area to the west of this isogloss was in fact subject to the change already by 1910, suggesting that it had more or less reached its current extent before this date, a fact which is in agreement with the historical evidence.

The geographical source, however, is not plausibly identified. Since almost all speakers of all ages in the northwest accepted *efo chdi*, the model simply attributes the highest intercept values to locations furthest from the isogloss running roughly from Llandudno southwards passing to the east of Betws-y-Coed and to the north of Dolgellau, that is, locations where the spatial kernel contained no speakers from east of the isogloss. The point of highest intercept value is marked with a black ring on Map 5. It seems unlikely that the far western tip of the Llyn Peninsula in the west was the origin of the innovation. Not entirely unexpectedly, the model is therefore unable to identify the origin of a completed change whose history is no longer present in the form of age variation in the data from current speakers.

The ongoing situation is plausibly modelled in the main. There is no reason to suspect that there is (much) change in progress here. The global age coefficient is not statistically different from zero. The only area where the model suggests change is in the far east of the region, where a small number of speakers who accepted *efo chdi* lead to increased intercept values. It is not clear if this is a real phenomenon: the number of speakers sampled in the east is small; the area is one where the proportion of native speakers among the population is low and where active revitalization efforts may be leading to dialect change. While it is possible that *efo chdi* is spreading to these areas, more data would be needed to justify this conclusion.

**8.1.2 Independent chdi**

We now turn to the independent use of *chdi*, also known to have been innovated in the nineteenth century. The global regression showed a very similar story to that with *efo chdi*, fitting well with the known history that independent *chdi* was innovated shortly after *efo chdi* in the nineteenth century. The intercept values from the GWR analysis are plotted on Map 6 using the same method as used above to create Map 5. Allowing the intercept to vary geographically (bandwidth 12.2 km) produces a 0.50 probability isogloss more or less identical to that for *efo chdi*, running slightly to the west of it in the south. The model is more successful than either of its predecessors in accounting for the data, accounting for 46.6% of the deviance (null deviance 154.10, residual deviance 82.24, AIC 102.80). The age coefficient remains non-significant (0.003, standard error 0.015, z value 0.187, p = 0.852), consistent with the hypothesis that there is no ongoing change here. Because the age coefficient is so small, there being next to no age variation in the current sample, the model concludes that the innovation had more or less reached its full extent in speakers born before 1910, and has not been diffusing further since then. Once again, the maximum intercept value is found at the far western tip of the Llyn Peninsula, which is unlikely to be the source of the innovation. As before, this has to be interpreted as an artefact of the model, confirming the unsurprising conclusion that non-age graded data collected after a change has run to completion cannot identify the geographic origin of that change.
8.2 Change in progress in GWR

We now turn to the more interesting case of change in progress. Mixed GWR models were created for each of the remaining eight morphosyntactic environments, all of which had shown statistically significant apparent-time change in the initial global logistic regression model. A summary of the overall performance of these models is given in Table 9. In all cases a GWR model performs much better than the global model and slightly better than the area- based model, accounting for a far higher percentage of the deviance, and, except in the last two cases, achieving a far superior AIC. These final two cases involve contexts where change is only just underway, and where chdi is possible in the grammars of relatively few of the informants. It is probably this that explains the relatively poor performance of the GWR model in those cases. The age co-efficients give estimates of the slope of the S-curve that are similar, but generally slightly lower than those produced by the global model.

This overall improved performance confirms that much of the variation in the data is geospatial in nature. GWR successfully models much of this variation without imposing arbitrary a priori boundaries on the data, thereby avoiding the problem of defining ‘locations’ faced by the area-based approach.

The most important advantage of GWR, however, is that it provides us with a method for investigating the diffusion of changes in progress across geographical space. Since it creates S-curves for each location, we can plot the parameters of those S-curves to visualize the progression of change. Geographic plots of the intercept values produced by the model are given in Maps 7–14. That these maps are equivalent to a series of S-curves can be seen from Figure 4, which shows (part of) the same information as Map 7, with S-curves plotted for a selection of points on the map, in this case, the major towns across the region.
Figure 4. S-curves corresponding to the GWR model for diffusion of *i chdi* ‘to you’ (Map 7).

In each case, the model identifies a point of innovation, the point with the highest intercept value, and this is identified on the maps by a black ring. A geospatial pathway of diffusion can be identified over the subsequent decades. The geographical pattern of diffusion that this model produces is on the face of it a highly plausible one. As before, dates refer to the generation acquiring the language at a particular time. In almost all cases, the diffusion follows a near perfect wave pattern through contiguous space (‘contagious diffusion’). This in itself is telling, since, while such a process has often been proposed (C.-J. N. Bailey, 1973:64–109; Schmidt, 1872:27), few instances of it have actually been demonstrated (Britain, 2013:623). G. Bailey et al. (1993:371–377) demonstrate that the quasi-modal *fixin to* is diffusing in this way through Oklahoma, as part of a wider trend in the southern United States, but other instances of this type of diffusion for morphosyntactic change are difficult to come by.

Map 7 (for *i chdi* ‘to you’), like the two plots discussed in section 8.1 above, identifies a secondary area of diffusion, implying a ‘leap’ from the northwestern core zone of innovation to Wrexham in the east. This may represent an instance of hierarchical diffusion (from primary urban area to secondary urban areas skipping intervening rural areas) in the sense of Trudgill’s (1974) adaptation of the gravity model (Hägerstrand 1967) to diffusion of linguistic innovation. However, as noted, above, further fieldwork would be necessary to substantiate this possibility.
Maps 7–14 are arranged in effective chronological order of the initial innovation, the overall pattern being entirely consistent with the documented historical evidence discussed in section 4.3 above, itself an encouraging result, since the two analyses are based on entirely independent data. While the date of innovation varies from context to context, the wave pattern of diffusion is constant across all of them. For the three earliest changes (after *i* ‘to’, inflected prepositions and *dyla* ‘should’), the model fails to identify a plausible point of actuation of change, producing its maximal intercept value at an edge of the map, as we saw above in cases of stable variation. This is presumably because, with changes that were already underway in the nineteenth century, there are many locations where the change has run to completion and where there is therefore no age variation through which to establish when the change began at that location. With the other changes, however, the model does produce plausible points of origin, those points being estimated from still-extant patterns of age variation in different districts.

Dialect syntax potentially holds out the prospect of providing a means to help decide between competing hypotheses about synchronic syntactic structure and competing accounts of change. If the changes in each context are a reflection of a single underlying change to the grammar, then we expect, by the Constant Rate Effect, the point of actuation to be the same for all contexts. In the current instance, all the points of actuation lie within the standard error of the others, hence we cannot demonstrate any of the changes to have distinct starting points. The usefulness of this finding is of course limited by the fact that, for the earliest changes, the method does not plausibly identify a point of actuation at all.

Secondly, dialect syntax more generally has been held up as a way to decide between competing analyses of synchronic phenomena (Barbiers, 2005; Craenenbroeck, 2014) or diachronic processes. According to the Constant Rate Effect, we expect the rate of change (slope of the S-curve) in two contexts to be the same if they are manifestations of a single underlying change. Fruehwald, Gress-Wright & Wallenberg (2013), for instance, apply this logic in arguing that, since word-final devoicing of stops in German proceeded at the same rate with /b/, /d/ and /g/, a single unitary historical devoicing change should be posited, rather than three distinct changes each with its own trajectory. In the current instance, the age coefficients for *efo chdi* and independent *chdi* are distinct from those for three of the verbal contexts, namely *oedd* ‘were’, *rhaid* ‘must’ and *(bua)sa* ‘would’, in that they lie outside of each other’s standard errors. This supports the view put forward in section 4.3 above that we

<table>
<thead>
<tr>
<th></th>
<th>Age coeff.</th>
<th>p-value of age coeff.</th>
<th>AIC</th>
<th>n</th>
<th>bandwidth (km)</th>
<th>null deviance</th>
<th>residual deviance</th>
<th>% deviance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>efo</em> ‘with’</td>
<td>0.020</td>
<td>0.217</td>
<td>91.44</td>
<td>118</td>
<td>11.9</td>
<td>126.93</td>
<td>72.42</td>
<td>42.9%</td>
</tr>
<tr>
<td>independent</td>
<td>0.003</td>
<td>0.852</td>
<td>102.80</td>
<td>138</td>
<td>12.0</td>
<td>154.10</td>
<td>82.24</td>
<td>46.6%</td>
</tr>
<tr>
<td><em>i</em> ‘to’</td>
<td>0.035</td>
<td>0.008 **</td>
<td>135.32</td>
<td>139</td>
<td>10.2</td>
<td>171.97</td>
<td>110.09</td>
<td>36.0%</td>
</tr>
<tr>
<td>infl. prep.</td>
<td>0.040</td>
<td>0.008 **</td>
<td>120.90</td>
<td>134</td>
<td>10.9</td>
<td>171.04</td>
<td>98.29</td>
<td>42.5%</td>
</tr>
<tr>
<td><em>dyla</em> ‘should’</td>
<td>0.017</td>
<td>0.260</td>
<td>82.67</td>
<td>74</td>
<td>14.9</td>
<td>101.72</td>
<td>69.53</td>
<td>31.6%</td>
</tr>
<tr>
<td><em>oedd</em> ‘were’</td>
<td>0.048</td>
<td>0.000 ***</td>
<td>150.86</td>
<td>146</td>
<td>12.4</td>
<td>202.15</td>
<td>132.03</td>
<td>34.7%</td>
</tr>
<tr>
<td><em>rhaid</em> ‘must’</td>
<td>0.041</td>
<td>0.005 **</td>
<td>117.86</td>
<td>111</td>
<td>11.5</td>
<td>153.65</td>
<td>97.15</td>
<td>36.8%</td>
</tr>
<tr>
<td><em>(bua)sa</em> ‘would’</td>
<td>0.049</td>
<td>0.000 ***</td>
<td>160.44</td>
<td>149</td>
<td>11.7</td>
<td>200.88</td>
<td>139.84</td>
<td>30.4%</td>
</tr>
<tr>
<td><em>bydd</em> ‘will’</td>
<td>0.030</td>
<td>0.132</td>
<td>67.76</td>
<td>64</td>
<td>15.8</td>
<td>69.70</td>
<td>59.02</td>
<td>15.3%</td>
</tr>
<tr>
<td><em>gan</em> ‘with’</td>
<td>0.025</td>
<td>0.052</td>
<td>133.49</td>
<td>163</td>
<td>17.7</td>
<td>139.70</td>
<td>122.09</td>
<td>12.6%</td>
</tr>
</tbody>
</table>

Table 9. Summary of GWR model for diffusion of *chdi* ‘you’ across various environments.
are dealing with (at least) two independent innovations here, namely the initial innovation of chdi as a pronoun in strong contexts (i.e. efo chdi and independent chdi) and the reanalysis of formerly weak contexts as strong (the three verbal contexts).

By producing an estimate of when a majority of speakers at different locations came to use the innovating form, the GWR model also provides a useful exploratory tool to guide further investigation. For instance, it tells us that, if we carried out a small-scale study of variation in the use of the second-person singular pronoun in the Caernarfon area, we would expect to find no age variation with i chdi ‘to you’, chdi after an inflected preposition and dyla chdi ‘you should’; only residual variation among the oldest speakers with oedda chdi ‘you were’ and rhai(d) chdi ‘you must’; but active, ongoing age variation with (bua)sa chdi ‘you would’, bydd(a) chdi ‘you will’ and gynno chdi ‘with you’. This could guide a decision as to whether this was an optimal location for such an investigation, and could guide the study design to focus on those contexts undergoing change there rather than contexts where the change had already run to completion.

Map 7. Plot of intercept values for GWR model of the diffusion of i chdi ‘to you’.
Map 8. Intercept values for GWR model of diffusion of *chdi* after inflected prepositions.

Map 9. Plot of intercept values for GWR model of the diffusion of *dyla chdi* ‘you should’.
Map 10. Plot of intercept values for GWR model of the diffusion of *oedda chdi* ‘you were’.

Map 11. Plot of intercept values for GWR model of diffusion of *rhai(d) chdi* ‘you must’.
Map 12. Plot of intercept values for GWR model of diffusion of (bua)sa chdi ‘you would (be)’.

Map 13. Plot of intercept values for GWR model of diffusion of bydd(a) chdi ‘you will’.
Map 14. Plot of intercept values for GWR model of the diffusion of *gynno chdi* ‘with you’.

9  The geospatial Constant Rate Hypothesis: Does rate of change vary geospatially?
The approach adopted in section 8 involves a ‘mixed’ model, one that assumes a constant rate of change at all locations, while allowing the onset of change (measured via the intercept value) to vary from place to place. GWR in principle allows both to be varied. Implicit in the Constant Rate Hypothesis is the claim that rate of change will be constant across all areas affected by a change. Other approaches to the diffusion of innovations also generally assume that locations do not overtake each other (C.-J. N. Bailey, 1973:83). Can a better data fit be achieved by relaxing this assumption and allowing rate of change to vary geospatially? To test this, we model the data using a geographically weighted regression where both intercept and rate of change vary geographically. In the current case, this produced no improvement in model fit (in terms of AIC) for any context: the extra complexity of the model outweighs marginal improvements. In short, we cannot be confident that allowing the rate of change to vary geospatially is not merely modelling random noise in the data. While it is not possible to demonstrate the negative conclusion that rate of change does not vary geospatially (Paolillo, 2011), this result is at least consistent with an extension of the Constant Rate Hypothesis to the domain of dialect variation and change (the ‘Geospatial Constant Rate Hypothesis’).

10  Conclusion
This article has examined how best to analyse patterns of ongoing change in geospatially dispersed linguistic data. Where a change diffuses across geographic space over time, we do not want to predefine the boundaries of areas for analysis and thereby assume the relevance of those boundaries for patterns of variation and change. We have seen that geographically weighted regression offers a way to circumvent this difficulty. In applying this approach to the diffusion of the innovative second-person singular northern Welsh pronoun *chdi* in various morphosyntactic contexts, we saw that the GWR model produced a narrative of change that was consistent with independent evidence drawn from nineteenth-century written sources, and which was rooted in standard assumptions about the propagation of linguistic
change (the S-curve model, the Constant Rate Hypothesis etc.). It also allowed us, more easily than an area-based model, to identify a pattern of successive wave-like diffusion of morphosyntactic innovation from a single core zone of innovation and therefore to reach a better understanding of the processes underlying that diffusion.

Acknowledgements
This work grew out of research done for my British Academy Mid-career Fellowship MC110579 ‘Syntactic variation in contemporary Welsh’. The fieldwork was funded in part by a Small Research Grant from the Newton Trust. This support is hereby gratefully acknowledged.

I would like to thank Bob Borsley, Margaret Deuchar, Jon Stammers and Maggie Tallerman, as well as the reviewers for *Journal of Linguistic Geography* for comments and contributions that have helped to shape this paper. Any remaining errors are my own.

The following copyright declaration applies to Map 1: ‘This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown.’ All maps contain Ordnance Survey data ©Crown copyright and database right [2014] (OS VectorMap™ District).

Primary sources
Evans, Beriah Gwynfe. 1898. *Dafydd Daquis*. Wrexham: Hughes a’i Fab.

All other textual citations are derived from Early English Books Online (EEBO, [http://eebo.chadwyck.com](http://eebo.chadwyck.com)), Eighteenth Century Collections Online (ECCO, [http://quod.lib.umich.edu/e/ecco](http://quod.lib.umich.edu/e/ecco)) or Welsh Newspapers Online (Beta) ([http://papuraunewyddcymru.llgc.org.uk](http://papuraunewyddcymru.llgc.org.uk)).

References


