

Possessive pronouns in Welsh: Stylistic variation and the acquisition of sociolinguistic competence*

Katharine Young^{1,2,3}, Mercedes Durham¹ and Jonathan Morris²

¹School of English, Communication and Philosophy, Cardiff University, Cardiff, UK; ²School of Welsh, Cardiff University, Cardiff, UK and ³School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, UK

Corresponding author: Katharine Young. Email: kyoung2@cardiffmet.ac.uk

*The online version of this article has been updated since original publication. A notice detailing the change has also been published

Abstract

This paper examines possessive pronoun forms in Welsh, a feature thought to be undergoing change (Davies, 2016). First, we seek to add to the understanding about how and in which stylistic contexts these forms are used. Second, we examine whether students in Welsh-medium schools with different home language backgrounds show the same sociolinguistic competence. In contrast to what is prescribed in many grammar books, the colloquial form *mam fi* 'my mum' is used at much higher rates than the traditional literary *fy mam* and sandwich variants *fy mam i*. This is particularly the case in more casual styles. We also find differences between north and south Wales in overall rates of use, but within the two schools studied, the English home language students broadly show the same patterns and constraints as the Welsh home language students, underlining that language background does not affect the acquisition of sociolinguistic competence.

Keywords: Welsh; sociolinguistic competence; style; possessives; acquisition

Introduction

Welsh, like many other languages, shows marked grammatical differences between formal and informal language, meaning that some forms are primarily used in writing or in formal oral situations and others are used mainly or exclusively in informal oral contexts. The use of possessive forms is a case in point, with a colloquial form being left completely out of grammars and textbooks, but very much present in speech although it is considered nonstandard by some (Davies, 2016; Jones, 1998; Thomas, 1988). This paper will establish how the three possible variants pattern for a range of both known and hitherto unexamined internal and external factors, and in doing so, contribute to a better understanding of the feature.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Previous studies examining variation in the choice of possessive forms show that the colloquial variant is used more among younger speakers (e.g., Davies, 2016). However, they have not focused on sociolinguistic factors such as speech context and home language, so we do not know to what extent these might influence the use of the colloquial. Our study examines young people in Welsh-medium education in both north and south Wales and attempts to (1) examine the linguistic and social influences on the selection of different possessive variants and (2) ascertain the extent to which sociolinguistic competence is acquired, which is completely novel in the Welsh context. In a context such as Wales, where the Welsh Government (2017) hopes to count a million speakers by 2050¹ and where Welsh-medium schools teach students who have Welsh as their home language and ones who do not in the same classroom², insight into how different variants of a feature are used in different stylistic contexts is vital, as it will allow us to better understand to what extent the children without Welsh as a home language share their peers' home-acquired patterns.

Previous studies focusing on the acquisition of sociolinguistic competence in immersion classroom situations (Mougeon, Rehner, & Nadasdi, 2010; Regan, Howard, & Lemée, 2009) have tended to find that speakers of a second language (L2) do not fully acquire informal styles and consequently do not demonstrate the full linguistic range that native (L1) speakers have. This is in contrast to studies that have examined less conventional learning contexts (e.g., French in study abroad: Regan *et al.*, 2009; acquisition of English following migration: Diskin, 2017; Drummond, 2011, 2013; and English in a lingua franca context: Durham, 2014), which have found that learners are able to match L1 speakers' social and linguistic constraints in some cases. The teaching situation in Wales is different from the immersion contexts studied, however, and this may affect the acquisition of sociolinguistic competence. As noted, students who speak Welsh at home are in the same Welsh-medium classes as children who do not, giving us an opportunity to examine a situation in which the acquisition of stylistic variation might be influenced by both peer-to-peer interaction and classroom instruction. This paper will examine how students at Welsh-medium schools from different linguistic backgrounds compare with respect to stylistic variation, and in doing so, contribute to our wider understanding of the acquisition of sociolinguistic competence.

Possessives in Welsh

There are three ways to form possessive pronouns in Welsh: a literary variant, a sandwich variant, and a colloquial variant (see Davies, 2016 for a discussion of these terms). Table 1 shows how each variant is formed.

The **literary** variant is formed by attaching a pronoun to the front of the noun phrase. These can come in the form of prefixed pronouns (e.g., *fy mam* 'my mum') or

Table 1. Possessive pronoun variants

Variant name	Prenominal pronoun	Noun	Post-nominal pronoun	English gloss
Literary	<i>fy</i> [və]	<i>mam</i> [mam]		'my mum'
Sandwich	<i>fy</i> [və]	<i>mam</i>	<i>i</i> [i]	'my mum'
Colloquial		<i>mam</i>	<i>fi</i> [vi]	'my mum'

infix pronouns (e.g., *i'm mam* 'to my mum'), depending on whether the preceding word ends with a vowel (Borsley, Tallerman, & Willis, 2007:157–158).

The **sandwich** variant is formed with a prenominal pronoun, like the literary variant, along with a post-nominal pronoun. The suffixed pronouns are described by Awbery (1994:1) as echo pronouns because they provide information already given in the phrase. The practice of “echoing” the pronoun of the possessor after the noun may have emerged due to the phonological similarity in certain dialects between *ei* ‘his’/‘her’ and *eu* ‘their’ (King, 2016:94).³

The **colloquial** variant is formed solely with a post-nominal pronoun. This variant is the newest of the three and has been found to be used increasingly frequently among younger speakers of Welsh. For example, Jones (1998:74), examining Welsh speakers in Rhymney, found that while it was used about 25% of the time by speakers above 60, it represented 40% of instances for those 20–39 and 75% of instances for speakers 7–19. The link between stratified age variation and language change in Jones’ (1998) data is evidence of this variant becoming increasingly frequent.

Despite this variant showing evidence of being increasingly common in Welsh, this form has often been seen as ungrammatical and remains heavily stigmatized. Awbery (1976:16) argued that it was not permissible for the possessive pronoun to follow the head noun (e.g., *ci fe* ‘[the] dog [of] him’). King’s (2016:93–94) *Modern Welsh*, which is considered a comprehensive reference to colloquial and literary Welsh grammar, notes that “while it is common ... [it] is widely regarded as sub-standard.” Similarly, Borsley et al. (2007:159) noted that the construction was possible but deemed it nonstandard. They also hypothesized that development of the colloquial variant may be an extension of non-pronominal noun phrases with possession (e.g., *car Megan* ‘Megan’s car’) where the possessor follows the noun. This is relevant because, as will be discussed further below, the colloquial form has at some points been ascribed to language acquisition processes (in children, but particularly in non-L1 speakers of Welsh; Willis, 2016).

Nevertheless, all three variants are found in speech and writing, even if, as shown by Watkins (1977), the colloquial variant is indeed found more often in speech and the literary variant more in writing. This distinction between mediums is tied to formality, as the literary form is seen as more formal and the colloquial form as more informal. The sandwich construction sits between the two in terms of formality and has previously been described as a way of drawing literary language closer to spoken speech (Watkins, 1977:153).

While there is some research on possessive pronoun use in different styles (Davies, 2016; Thomas, 1988), across regions (Jones, 1998), class (Thomas, 1988), and age (Davies, 2016), these studies do not consider the relative influence of a number of linguistic and social factors, and none consider home language. This reveals a gap in terms of how speakers, particularly young speakers, vary according to sociolinguistic and linguistic factors such as home language and context of use.

Data and methods

The data presented here belongs to a wider project examining the acquisition of sociolinguistic competence in Welsh immersion classrooms.⁴ To assess the extent to which bilingual Welsh-English children in Welsh immersion classes from different backgrounds and different parts of Wales showed stylistic variation, two schools were

sampled: one in the north, in an area where the percentage of Welsh speakers is relatively high, and one in the south, where the percentage of Welsh speakers in the community is relatively low (this factor will be referred to as *region* in the analysis). Participating students (who were all 16 or 17 years old) in both schools were separated into two groups: those who mostly used Welsh as a home language (WHL) and those who mostly used English as a home language (EHL)⁵ (*home language*). Students were recorded in three contexts representing different levels of formality (*context*). Further details on these factors are presented in the following sections. All participants were considered Welsh-English bilinguals on the basis that they had attended Welsh-medium or bilingual secondary education (age 11–16) prior to attending their sixth form, where classes and examinations were held in Welsh as a first language.

Region

The project targeted two areas of Wales: Gwynedd in the north and Cardiff in the south (Figure 1). Gwynedd is generally perceived to be part of the Welsh-speaking “heartlands” (Coupland, 2012), with 64% of people in the area reporting that they are able to speak Welsh (Welsh Government, 2022). Cardiff is the capital city of Wales and is a densely populated metropolitan area. The proportion of Welsh speakers in Cardiff is lower than the national average due to different patterns of language shift across Wales. The Census results for Wales showed that 18% of the country’s population were



Figure 1. Map of counties in Wales with Gwynedd (at north) and Cardiff (at south) in gray.

able to speak Welsh, while in Cardiff, 12% of people reported that they could speak Welsh (Welsh Government, 2022).

In schools in Gwynedd, Welsh is the sole or main medium of instruction for most pupils (9,127 pupils in Welsh-medium schools and 6,178 pupils in bilingual schools where Welsh is the medium of instruction 80%+ of the time, out of a total of 17,038 [i.e., 90%] of pupils) (StatsWales, 2022). We examined a sixth form college in Gwynedd (students aged 16–18 preparing for their final school exams). A total of 10 pupils participated in the study from the school we had selected. They were from a class of 12 following a bilingual Psychology A Level course.

In Cardiff, there are 8,478 pupils in Welsh-medium education out of the total number of pupils, 56,837 (i.e., 15% of pupils) (StatsWales, 2022).⁶ This shows that pupils studying through the medium of Welsh are a minority in the capital. A total of eight pupils participated in the study from the Cardiff school selected. They were from across year 12, studying various A Level courses.

The data in both schools was collected in 2021 during the global COVID-19 pandemic. This delayed the collection process and meant that we recruited fewer students than we had originally planned. All the tasks were conducted online and will be detailed in the *Context* section.

Home language

In both locations, students were classified by whether they used mostly Welsh or English at home. We did this by asking students what language(s) they used with their parents and other family. Most were either fully Welsh- or English-speaking at home, but some used either Welsh or English with different family members. Students reporting more than 50% Welsh use were classified as WHL and those with less than 50% as EHL. As Table 2 shows, we had roughly similar numbers in the two communities and across EHL and WHL groups⁷.

Table 2. Speakers by region and home language

	English home language	Welsh home language	Total
Gwynedd	4	6	10
Cardiff	4	4	8
Total	8	10	18

Context

In order to offer something to the schools and students in exchange for their help and to deal with the fact that COVID-19 meant that the classroom and peer interactions we were hoping to obtain could not be collected, the first author put together a set of careers training recordings and exercises for students to watch and discuss in peer groups. Following this, she gave each student a mock job interview using the skills they had acquired from the recordings and exercises. A week or two after that, she had a more informal chat with them individually to obtain more casual data. The data collection consisted of two sessions (the conversation and interview contexts) for each

participant with the researcher and six workshop sessions with their peers, which participants completed in their own time over the course of 2–3 weeks (the peer-group context). These contexts represent different levels of formality:

1. The conversation context: sociolinguistic interviews with the researcher in which the participants talk about their lives.
2. The peer-group context: participants interact with their peers during six semi-spontaneous workshop sessions.
3. The interview context: participants undertake a mock job interview with the researcher.

The mock job interview was designed to be the most formal context. The students were asked to prepare for job interview questions such as “Tell me about a time when you showed good time management” in the run up to the interview.⁸ This was the first time they had met the interviewer. The other two contexts were relatively informal; the peer-to-peer context and sociolinguistic interviews were loosely structured conversations around set topics (careers training module for peer-to-peer interactions, and home life and hobbies for the sociolinguistic interview). The sociolinguistic interview and job interview with the researcher were conducted and recorded on Zoom. In the peer-to-peer context, the researcher was not present, and the students recorded their own conversations using their mobile phones. We found that the format of the workshops meant that some students were more formal in their language use than in the conversation context, despite speaking solely to their peers, while for others, the sociolinguistic interview was seen as less casual than the peer-to-peer workshops.⁹

We chose these three contexts to establish to what extent the students were able to style-shift in less formal contexts. As noted in the Introduction, previous research examining the acquisition of sociolinguistic competence found that while L2 speakers do acquire some variation patterns (Durham, 2014), they are less likely to acquire informal variants as classroom settings often privilege formal variants (Mougeon *et al.*, 2010; Regan *et al.*, 2009). We expect that the situation in Welsh classrooms is somewhat different (not least because Welsh home language children are present too), and having two separate casual contexts will allow us to fully establish what is going on.

The data were collected in June and July 2021; this was judged to be a suitable time for data collection by gatekeepers in the schools as it was after exams and before the summer holidays. Approximately 34 hours of speech data across the different contexts and in both schools were collected from participating students. This comes out to just over a quarter of a million words when transcribed.

Feature extraction

All possible possessive pronouns were extracted from the data. Fixed phrases were excluded, such as *yn fy marn i* ‘in my opinion’, as were any unclear tokens ($n = 2$). We also noted earlier that the literary variant can contain prefixed or infix pronouns; both these types were transcribed as prefixed pronouns, because of their rarity. For example, *’n chwaer* ‘my sister’ was transcribed as *fy chwaer* (the same is applied to infix pronouns in the sandwich variant, e.g., *’n chwaer i* was coded as *fy chwaer i*).

A total of 1,968 possessive pronoun tokens were extracted, 871 were from Gwynedd, and 1,097 from Cardiff. When broken down by home language, 1,226 tokens came from students from Welsh-speaking homes (WHL), while 742 possessive pronoun tokens came from students from English-speaking homes (EHL).

The sections below discuss the social and linguistic factors considered in our analysis and summarize previous findings (where they exist).

Social factors

Due to our focus on 16- to 18-year-olds and the fact that our data are not evenly distributed in terms of gender across the various groups (there are fewer males overall), we do not consider age or gender in our analyses. The colloquial variant has consistently been found to be used at higher rates in younger speakers (Davies, 2016; Hatton, 1988; Jones, 1998; Thomas, 1988; Watkins, 1977). We thus considered the social factors outlined below.

Context. The main research on style differences in the use of possessive pronoun variants focuses on written and oral differences. Borsley et al. (2007:158) found that literary Welsh primarily uses the literary variant, although Watkins noted (1977:153) that some authors make a “conscious effort” to use the sandwich form in modern literary language. The colloquial construction, as noted above, is considered ungrammatical by many and is found less often in writing. It is instead reported to occur more often in speech although it has been shown to depict L2 speech in literary works (Willis, 2016).

In addition, there is variation within the spoken register, as demonstrated in Davies (2016:41). Davies used a corpus of “spontaneous, informal speech” (2016:33) to analyze possessive pronoun variant use of 151 participants who were predominantly from north-west Wales and found that the three variants were used, even though the context remained consistently informal. Young (2019) examined Welsh teachers and focused on their reported use of different features across different formality contexts. Not only did the teachers report significant variation in their own use of the colloquial form in in-classroom and out-of-classroom contexts, but they also reported that they were more likely to correct their students’ use of the colloquial possessive variant as the context of use became more formal. This suggests that the colloquial variant will be used at higher rates in the more informal speech contexts. Using the three different contexts we collected, we aim to establish if this is the case and if all sets of speakers follow this pattern.

Home language differences. There is no previous research focusing on differences in use in possessive pronouns according to home language (or between early and late learners of Welsh), although Robert (2009:104) identified the colloquial variant as a potential indicator of new speaker speech. It is worth noting that learning materials for adult learners (National Centre for Learning Welsh, 2023) tend to focus on the literary and sandwich variants and that, as discussed above, many Welsh grammars do not list the colloquial form as an option. We aim to establish whether Welsh home language speakers are more or less likely to use the colloquial variant than non-Welsh home language speakers.

Regional differences. King (2016:93) reported that variation in variant use exists from region to region, although he also noted that the literary and sandwich variants tend to be seen as the standard forms even in spoken language. However, much previous work has been conducted at a single location. Awbery (1994) considered Pembrokeshire in south-west Wales, and Roberts (1988) examined Pwllheli in north-west Wales, making it more difficult to compare regions. Jones (1998) conducted the only comparative work thus far, considering Rhymney (south-east Wales) and Rhosllanerchrugog (north-east Wales). She found that the use of the literary form was more common in the south-east Wales cohort, while in north-east Wales, the colloquial variant was more frequent. We aim to establish whether there are differences between the north and the south in our data as well.

Linguistic factors

We considered two linguistic factors that have previously been found to affect the rates of the variants (grammatical person and lexical frequency). Given the lack of comprehensive variationist research on Welsh possessive pronouns, we also considered several other factors that have been found to be relevant for similar features in other languages (language of noun and noun category).¹⁰

Grammatical person. Davies (2016:55) compared rates of the three variants in third-person singular and first-person plural forms and found that the colloquial variant was more common in both for younger speakers, whereas the colloquial variant was less frequent in the third-person singular for older speakers. King (2016:94) suggested that the sandwich variant may be preferred to the literary variant in the third-person singular in order to distinguish between the prenominal masculine pronoun *ei* and the feminine pronoun *ei* (e.g., *ei gar o* ‘his car’ and *ei char hi* ‘her car’), which without the mutation and post-nominal pronoun would be identical.

Unlike previous work, we consider all grammatical persons, although overall token numbers mean that we will combine some forms together in the statistical analysis. Examples of each grammatical person with the three possessive variants are presented in Table 3 to demonstrate how this factor was coded in the data. The type of mutation that occurs with the sandwich and literary forms is also given.

Table 3. Coding grammatical person¹¹

	Colloquial	Sandwich	Literary	English gloss
First-person singular	<i>stryd fi</i>	<i>fy stryd i</i>	<i>fy stryd</i>	‘my street’
Second-person singular	<i>stryd ti/chdi</i> ¹²	<i>dy stryd di</i>	<i>dy stryd</i>	‘your street’
Third-person singular masculine	<i>stryd fo/o; fe/e</i> ¹³	<i>ei stryd o/e</i>	<i>ei stryd</i>	‘his street’
Third-person singular feminine	<i>stryd hi</i>	<i>ei stryd hi</i>	<i>ei stryd</i>	‘her street’
First-person plural	<i>stryd ni</i>	<i>ein stryd ni</i>	<i>ein stryd</i>	‘our street’
Second-person plural	<i>stryd chi</i>	<i>eich stryd chi</i>	<i>eich stryd</i>	‘your street’
Third-person plural	<i>stryd nhw</i>	<i>eu stryd nhw</i>	<i>eu stryd</i>	‘their street’

Lexical frequency. Davies (2016:44) found that certain frequently used possessed nouns (such as *tŷ* ‘house’) were only used with the colloquial variant and hypothesized that high frequency nouns that show limited variation could point to conventionalized “set phrases” (e.g., *tŷ ni* ‘[our] home’). Variationist sociolinguistic research has increasingly examined frequency effects on morphosyntactic variation (e.g., Erker & Guy, 2012; Linford & Shin, 2013), particularly in the field of L1 and L2 acquisition. As there is little research on high frequency nouns in Welsh and conventionalized “set phrases,” we felt it was important to consider this factor in our analysis.

Following previous research, we decided to determine lexical frequency within our own corpus (cf. Erker & Guy, 2012). To do this, we created a list of the most frequent nouns in the corpus (i.e., those occurring at a rate higher than two instances per 10,000 words of the corpus, which was about .02%) and then compared these to the nouns that occurred in our set of possessive pronoun tokens. For ease, instances of plural, singular, masculine, and feminine forms were counted as a single noun. For example, *athro* ‘male teacher’, *athrawes* ‘female teacher’, and *athrawon* ‘teachers’ were all coded as *athro*.¹⁴ Nouns that were found both in our possessive pronoun tokens and in the frequency corpus were counted as frequent, and all other nouns were counted as infrequent.

Out of the 362 different noun forms found in the possessive tokens, 51 were found to be frequent and to represent a total of 1,150 tokens, which is 58% of the overall possessive pronoun tokens. In the analysis below, we will establish whether, like in Davies (2016), frequent nouns are more likely to be used with the colloquial form.

Language of possessed noun. Welsh contains loanwords and calques from English, and speakers are also known to code-switch and insert English words when speaking Welsh (Deuchar, Donnelly, & Piercy, 2016). In order to see if this may have an effect on variant choice, we coded for recent loanwords using the following criteria: where an English loanword appeared in *Geiriadur Prifysgol Cymru* (the standard historical Welsh language dictionary) as a Welsh word, it was coded as Welsh. Other English words were coded as English. This is a commonly used criterion in codeswitching research in Welsh (Prys, 2016).

Possessed noun category. Nouns can be categorized according to their alienability or inalienability. The alienability of possession is determined by “whether the object can exist apart from its possessor” (Nichols, 1988:575), and thus nouns such as *calon* ‘heart’ and *brawd* ‘brother’ are examples of inalienable possession. On the other hand, if the possessed object can exist apart from its possessor, it is considered alienable, as in the case of nouns such as *ffôn* ‘phone’ or *swydd* ‘job’. Oceanic languages, spoken in Papua New Guinea, Melanesia, Polynesia, and Micronesia, distinguish grammatically between alienable and inalienable possessive constructions (Lichtenberk, Vaid, & Chen, 2011). English does not have such a grammaticalized distinction, and it is unknown whether Welsh does. We thus examined this to determine its potential effect. We initially coded using Lichtenberk et al.’s (2011) subcategories of alienability, but in our results, we present only the broad division between inalienable and alienable.

Statistical analysis

Mixed-effects logistic regression analyses were conducted in order to determine the effect of external and internal factors on the variation of the possessive pronoun variant. Statistical modeling was carried out using the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) packages for *R* (R Core Team, 2021). Fixed effects (such as home language and context) are factors that are replicable in further studies. By including random effects (speaker and word), which are sampled randomly, modeling can account for inter-speaker variation (Johnson, 2009). The mixed-effects analysis will present a two-way breakdown of the variants: the colloquial variants on the one hand and the other two (non-colloquial) variants on the other.

Summary of factors

Table 4 presents each of the random and fixed factors influencing the possessive construction. The levels presented here were built into the statistical model (in italics is the baseline factor).

Table 4. Factors included in the statistical modeling

Factor	Levels
Context	<i>Mock job interview</i> Peer-to-peer Sociolinguistic interview
Speaker's home language	<i>English</i> Welsh
Region	<i>Cardiff</i> Gwynedd
Possessed noun category	<i>Alienable</i> Inalienable
Possessed noun language	<i>Welsh</i> English
Frequency	<i>Frequent</i> Infrequent

Results

Factor-by-factor distribution

The overall distribution for the different groups and factors will be discussed individually before moving onto the statistical analysis. A total of 1968 tokens were extracted and coded in the data. As Table 5 shows, the colloquial variant is the most frequently selected form, followed by the literary form. Given the age of the speakers and the fact that previous studies have shown that there are higher rates of the colloquial form in younger speakers, this is not unexpected.

Region

We now turn to region, considering the data from the location in the north (Gwynedd, $n = 871$) and in the south (Cardiff, $n = 1097$) separately (Figure 2). Recall that

Gwynedd has a higher proportion of Welsh speakers in the community and that previous research had found that the north-west had higher rates of the colloquial form.

Figure 2 shows that while the rates for the colloquial variant are indeed higher in Gwynedd (71% versus 48% in Cardiff), the patterning remains the same. The colloquial form is the most frequent form in both, followed by the literary variant, while the sandwich variant is the least frequent. By examining home language and region together, we will be able to establish whether the difference is due to home language or whether it reflects a genuine difference in region. We will present this alongside context in Table 6.

Table 5. Overall distribution of possessive pronoun variants

Pronoun type	%	<i>N</i>
Colloquial	58%	1142
Literary	27%	524
Sandwich	15%	302
Total		1968

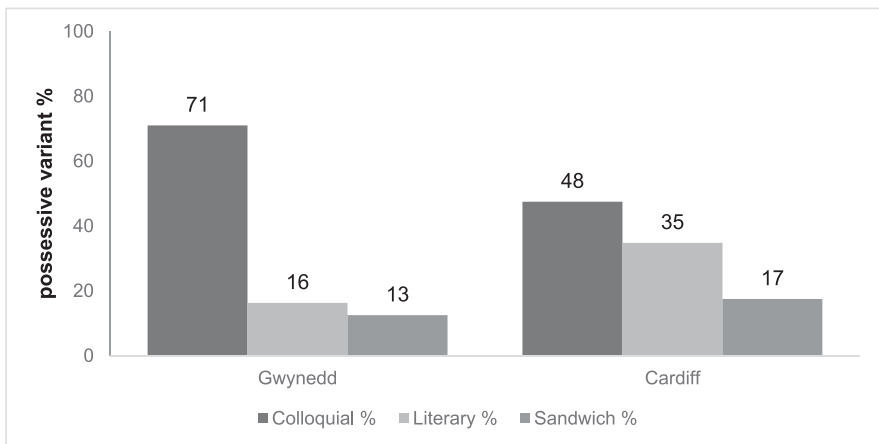


Figure 2. Rates of possessive pronoun variants in Gwynedd and Cardiff.

Context

The three contexts—a mock job interview, a peer-to-peer conversation, and a sociolinguistic interview—are presented from most formal to least formal. Table 6 presents the percentages of the three variants for the four region–home language groups in the three contexts, while Figure 3 presents a two-way split for the variants (colloquial versus literary + sandwich).

In terms of differences based on region and home language, Table 6 makes it clear that while the EHL students in Gwynedd are “extreme” colloquial variant users, the

Table 6. Possessive pronouns according to context, by speaker group

Speaker group	Context	Colloquial %	Literary %	Sandwich %	Total <i>N</i>
Gwynedd EHL	Job interview	84	11	4	44
	Peer-to-peer	80	13	7	61
	Sociolinguistic interview	98	1	1	148
	Total	91	6	3	253
Gwynedd WHL	Job interview	45	34	21	123
	Peer-to-peer	66	20	14	226
	Sociolinguistic interview	69	15	16	269
	Total	63	21	17	618
Cardiff EHL	Job interview	17	49	34	86
	Peer-to-peer	50	35	15	197
	Sociolinguistic interview	75	11	15	206
	Total	55	27	18	489
Cardiff WHL	Job interview	6	66	28	155
	Peer-to-peer	53	33	14	238
	Sociolinguistic interview	56	32	13	215
	Total	42	41	17	608

difference between regions is not due solely to them, as the WHL students are also higher than both groups in Cardiff.

More generally, the EHL students show higher rates of the colloquial than the WHL students, but it is nonetheless the most frequent form for all groups, followed by the literary form. The rates of use for the colloquial and literary forms for the Cardiff WHL students are very similar, which puts them at odds with the other three groups. Gwynedd's EHL students have far lower rates of the literary and sandwich forms due to their near-categorical use of the colloquial form.

Given our findings for region and home language, we present the results of each subsequent factor broken down across these four groups. In terms of context, [Table 6](#) shows that all four groups have the highest rate of colloquial variants in the sociolinguistic interview, but the difference between peer-to-peer and sociolinguistic interview is less marked for the two WHL groups than for the EHL groups. The job interview context is very different from the other two in Cardiff, underlining that the more formal context is triggering a shift away from the colloquial variant in favor of the other two variants (it is important to note that it is not purely a shift toward the literary form as the sandwich form is used roughly a third of the time in those contexts). This is the case for both the WHL children and the EHL children, which underlines that the EHL children have acquired at least some stylistic aspects.

For Gwynedd, the situation is less clear because of the higher rates of the colloquial form overall, but by and large, it does seem the WHL students have the same pattern

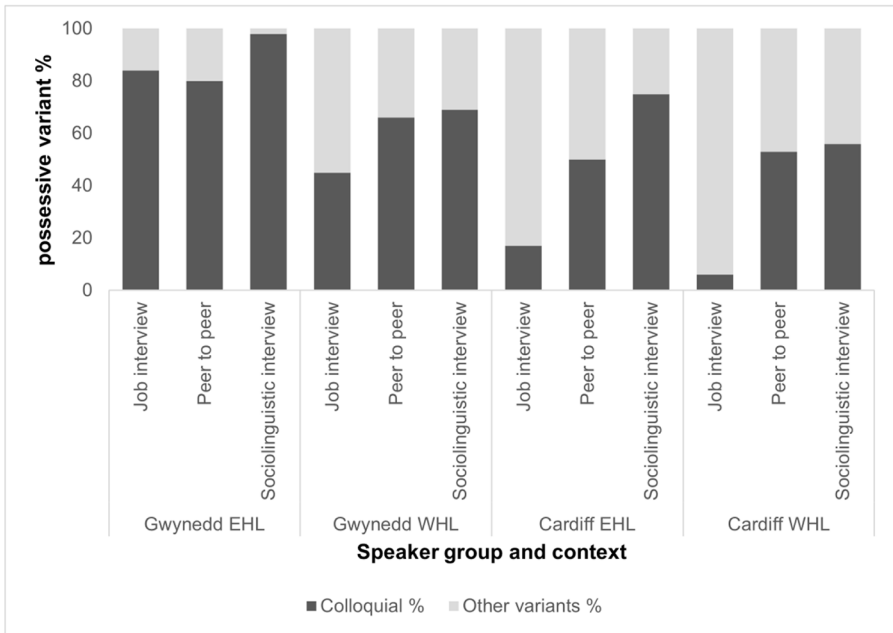


Figure 3. Distribution of possessive variants in each context by speaker group.

as the Cardiff students (job interview showing a sharper stratification than peer-to-peer and sociolinguistic interview). The EHL students show a shift, but for them, the division is between sociolinguistic interview versus job interview and peer-to-peer.

We will turn to the internal factors next, but at this point, the main findings are that (1) the colloquial variant is the most frequently used variant overall, (2) there is a substantial difference in use between Gwynedd and Cardiff, and (3) that the EHL students show style shifting depending on the level of formality of the context like their WHL peers, suggesting they have acquired L1-like sociolinguistic competence.

Lexical frequency

Table 7 presents the rates of the three variants for frequent and infrequent nouns for the four groups.

For all four groups (though only marginally for Cardiff Welsh home language students), infrequent nouns are more likely to occur with the literary and sandwich forms. The two Welsh home groups show that the decrease in the colloquial form is mainly linked to an increase in the sandwich variant as the rates of the literary variant barely shift. More frequent nouns appear, then, to show higher rates of colloquial use, and this is the case for both the WHL and the EHL speakers.

Grammatical person

We now move onto the second factor that has previously been studied, grammatical person.¹⁵ Tables 8 and 9 show the distribution for each group. Recall that Davies (2016)

had found that rates of the colloquial third-person singular and the first-person plural forms increased in younger speakers.

Low numbers in some categories (especially for the Gwynedd EHL group) mean that some rates may not be fully accurate, but for the two WHL groups, the two categories that Davies (2016) studied (third-person singular and first-person plural) have higher rates of the colloquial form than overall usage. With respect to the third-person singular, this is also the case for the EHL groups. The first-person singular generally has higher rates of the colloquial variant than the remaining pronouns for all four groups.

Table 7. Possessive pronouns according to noun frequency, by speaker group

Speaker group	Frequency	Colloquial %	Literary %	Sandwich %	Total <i>N</i>
Gwynedd EHL	Frequent	94	5	1	165
	Infrequent	86	8	6	88
Gwynedd WHL	Frequent	66	20	14	318
	Infrequent	60	21	19	300
Cardiff EHL	Frequent	60	24	16	292
	Infrequent	47	32	21	197
Cardiff WHL	Frequent	42	42	15	375
	Infrequent	41	39	21	233

Table 8. Possessive variants used by Gwynedd EHL and WHL by grammatical person

Grammatical person	Gwynedd EHL				Gwynedd WHL			
	Coll. %	Lit. %	Sand. %	<i>N</i>	Coll. %	Lit. %	Sand. %	<i>N</i>
First-person singular	96	3	1	162	67	19	14	349
Second-person singular	86	0	14	7	55	25	20	49
Third-person singular	98	0	2	41	69	15	17	89
First-person plural	71	19	9	21	79	13	8	24
Second-person plural	20	40	40	5	30	40	30	40
Third-person plural	77	24	0	17	57	24	19	67

Table 9. Possessive variants used by Cardiff EHL and WHL by grammatical person

Grammatical person	Cardiff EHL				Cardiff WHL			
	Coll. %	Lit. %	Sand. %	<i>N</i>	Coll. %	Lit. %	Sand. %	<i>N</i>
First-person singular	55	30	15	287	42	47	11	403
Second-person singular	41	41	18	17	33	54	13	24
Third-person singular	66	14	21	73	45	24	28	71
First-person plural	45	14	41	22	45	16	39	31
Second-person plural	54	35	10	48	35	45	20	20
Third-person plural	45	21	33	42	39	29	34	59

When considering the other two variants, the patterns seem to be tied to community, more than home language. For the Gwynedd speakers, second-person plural has the highest rates of both the literary and the sandwich forms. Instead, in Cardiff, the rates of the literary form are highest in second-person singular, and for the sandwich form, they are highest in first-person plural. Based on this, the EHL students share some, but not all, of the general patterning with their WHL counterparts, but do share the more local patterns of use.

Although we included this factor in initial models, we found that it was never significant and that removing it improved the fit. We have nevertheless provided the group breakdowns here, because it is one of the few factors that has been previously studied.

Possessed noun category

A total of 590 (30%) alienable constructions and 1,382 (70%) inalienable constructions were identified in the corpus. Table 10 shows the rates of the three possessive pronoun constructions in alienable and inalienable possessum words, once again by region and home language.

Except for Cardiff WHL, inalienable nouns seem to favor the colloquial form slightly more than the alienable contexts.

Table 10. Possessive pronoun variants by alienability and speaker group

Speaker group	Alienability	Colloquial %	Literary %	Sandwich %	Total <i>N</i>
Gwynedd EHL	Alienable	84	14	2	50
	Inalienable	93	4	3	203
Gwynedd WHL	Alienable	61	24	15	208
	Inalienable	65	19	17	410
Cardiff EHL	Alienable	47	32	21	148
	Inalienable	58	26	17	341
Cardiff WHL	Alienable	45	41	14	184
	Inalienable	40	40	19	424

Possessed noun language

Although all the most frequent possessed nouns in the corpus were Welsh, there were 71 English nouns with possessive forms in the corpus. Most English possessed nouns only appeared once in the corpus, but some were repeated (e.g., *cousin* [$n = 10$], *job* [$n = 7$], *boss* [$n = 7$]). Table 11 presents the possessive variants used by possessed noun language and speaker group.

Before turning to the differences in rates of the variants, it is worth pointing out that the rates of English versus Welsh nouns differ in the two communities. In Gwynedd, English nouns represent 8%–9% of the overall tokens, whereas in Cardiff, they represent 3%–4%. This is important for the overall distribution of the variants in the two communities, as English nouns are more likely to occur with the colloquial form than

Table 11. Possessive variants used by speaker group and possessed noun language

Speaker group	Possessed noun language	Colloquial %	Literary %	Sandwich %	<i>N</i>
Gwynedd EHL	English	91	4	4	23
	Welsh	91	6	3	230
Gwynedd WHL	English	85	10	6	52
	Welsh	61	22	18	566
Cardiff EHL	English	89	5	5	19
	Welsh	53	28	19	470
Cardiff WHL	English	79	16	5	19
	Welsh	41	42	18	589

Welsh nouns. The overall numbers of English nouns are relatively low, however, and all four groups show a lower rate of the colloquial form with Welsh nouns, which demonstrates that the pattern is shared. For the Cardiff WHL group, the literary form is in fact used more frequently (42%) than the colloquial form (41%) with Welsh nouns.

Initial trends

Across the factor groups analyzed, there are differences in rates and patterns across the two communities. Within each community, there are differences in rates for the EHL and WHL students, but for many of the factors considered, the patterning is shared. This suggests that the EHL speakers have acquired the constraints of their peers, but that it is vital to consider this with respect to their local peers and not with how this feature might pattern elsewhere. We now turn to the statistical analysis.

Statistical analysis

In mixed-effects modeling, the colloquial variant was used as the dependent variable, which means that the model shows the likelihood of a speaker producing the colloquial variant (possessed noun + post-nominal pronoun) compared to the literary (prenominal pronoun + possessed noun) and sandwich (prenominal pronoun + possessed noun + post-nominal pronoun) variants. A general-to-specific approach was taken to the statistical modeling (Baayen, 2008:205; Nance, 2015:565). The first model included all predictors as shown in the following R code: VARIANT ~ CONTEXT * HOME LANGUAGE * REGION + NOUN LANG + CATEGORY + FREQUENCY + (1|PARTICIPANT) + (1|WORD). Nonsignificant predictors and interactions were then removed from the model one at a time and compared to the first model using a series of ANOVAs. If the removal of a nonsignificant predictor or interaction improved the fit of the model, then this model was retained. Following the analysis of the entire dataset, four separate models were then conducted, one for each group (Gwynedd EHL, Gwynedd WHL, Cardiff EHL, and Cardiff WHL).

Regression tables for the best-fitting model contain an intercept corresponding to a baseline combination of levels. Results Tables 12–16 show the fixed factors

that were significant predictors of colloquial variant use. Regression coefficients (β) (labeled *Estimate* in these tables) for each term indicate deviations for the intercept and are included alongside z -values and p -values for the levels associated with each factor. A positive significant coefficient suggests that the named factor level was more likely to influence the production of the colloquial variant than the baseline factor level. Conversely, a negative significant coefficient indicates that the named factor level was less likely to result in the production of the colloquial variant than the baseline factor level. Grammatical person was not retained in the final model.

In the overall model, three factors show a significant effect: region, context, and noun language. For region, Gwynedd has higher rates than Cardiff, underlying that there is a clear regional dimension. For context, the job interview is least likely to demonstrate the use of the colloquial form, with the peer-to-peer and particularly the sociolinguistic conversation showing higher rates. With respect to noun language, colloquial possessive constructions are more likely to contain English nouns than Welsh nouns.

The model in Table 12 also shows a three-way interaction between context, home language, and region. The interaction suggests that there is a difference in the effect of the context within the home language and region groups. The effect is negative, meaning that the effect of the sociolinguistic interview on the production of the colloquial variant is less in the Gwynedd WHL group compared to the Cardiff EHL group. In other words, the effect of context on the production of the colloquial variant differs between the region and home language groups, and in some cases, these differences are significant. In order to examine this interaction further, and in order to investigate potential differences in the acquisition of sociolinguistic competence, we examined the data for the home language and region groups separately. Tables 13–16 present the model for each group. Note that in these models, noun category was not significant, and its omission significantly improved the model fit.

While nothing comes up as significant for the Gwynedd EHL group (this may be due to lower overall token numbers), context and noun language are significant for the other three groups and noun frequency is for the Cardiff WHL group. For noun language, the direction of the effect, that is, higher rates of the colloquial form with English nouns than Welsh nouns, is the same across all four groups. This is also the case for noun frequency: all groups have higher rates of the colloquial form for frequent nouns than for infrequent nouns. For context, however, there is a home language effect. Both Gwynedd and Cardiff WHL speakers increase their rates of the colloquial form from job interview to sociolinguistic interview to peer-to-peer, but for Cardiff EHL, the pattern is job interview to peer-to-peer to sociolinguistic interview. For Gwynedd EHL, the pattern is different again, peer-to-peer then job interview then sociolinguistic interview.

It seems then that the EHL speakers share the key internal factors of variation with their WHL peers, and they do vary in terms of context (unlike what was found in previous studies); however, they nonetheless do not completely share the hierarchy of the WHL speakers. Why might this be? We turn to this in the Discussion section.

Table 12. Overall model (total possessives $n = 1968$, 58% colloquial)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
(Intercept)	.29	1.10	.27	.79			
Region							
(Baseline: Cardiff)						1097	48
Gwynedd	3.76	1.45	2.59	.01	**	871	71
Context							
(Baseline: Job interview)						408	28
Sociolinguistic interview	2.15	.40	5.33	<.001	***	838	72
Peer-to-peer	1.56	.39	4.02	<.001	***	722	58
Noun language							
(Baseline: English)						113	86
Welsh	-1.48	.44	-3.37	<.001	***	1855	56
Home language							
(Baseline: English)						742	67
Welsh	-2.32	1.41	-1.64	.10		1226	52
Frequency							
(Baseline: Frequent nouns)						1150	61
Infrequent nouns	-.35	.27	-1.27	.20		818	54
Noun category							
(Baseline: Alienable)						590	54
Inalienable	-.03	.25	-.14	.89		1378	60
Context and home language							
(Baseline: Job interview and English)							
Peer-to-peer:Welsh	1.75	.59	2.99	.002	**		
Sociolinguistic interview:Welsh	1.09	.60	1.83	.07	.		
Context and region							
(Baseline: Job interview and Cardiff)							
Peer-to-peer:Gwynedd	-2.21	.77	-2.90	.003	**		
Sociolinguistic interview:Gwynedd	-.49	.90	-.56	.58			
Home language and region							
(Baseline: English and Cardiff)							
Welsh:Gwynedd	.54	1.95	.28	.78			

(Continued)

Table 12. (Continued.)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
Context and home language and region							
(Baseline: Job interview and English and Cardiff)							
Sociolinguistic interview:Welsh:Gwynedd	2.44	1.05	-2.31	.02	*		
Peer-to-peer: Welsh:Gwynedd	-.88	.95	-.93	.35			

Table 13. Mixed-effects logistic regression on the Gwynedd Welsh home language (total possessives $n = 618$, 63% colloquial)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
(Intercept)	1.61	.62	2.61	.01	**		
Noun language							
(Baseline: English)							
Welsh	-1.69	.52	-3.23	.001	**	556	62
Context							
(Baseline: Job interview)							
Peer-to-peer	1.08	.30	3.65	<.001	***	226	65
Sociolinguistic interview	.74	.30	2.34	.01	*	269	69
Frequency							
(Baseline: Frequent)							
Infrequent nouns	-.31	.31	-1.01	.31		300	60

Table 14. Gwynedd English home language (total possessives $n = 253$, 91% colloquial)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
(Intercept)	7.63	2.98	2.56	.01	*		
Context							
(Baseline: Job interview)							
Sociolinguistic interview	2.50	1.46	1.71	.09		148	98
Peer-to-peer	-.38	1.04	-.36	.71		61	80
Frequency							
(Baseline: Frequent)							
Infrequent nouns	-1.26	1.62	-.78	.44		88	86
Noun language							
(Baseline: English)							
Welsh	-.44	2.15	-.21	.84		230	91

Table 15. Cardiff Welsh home language (total possessives $n = 608$, 42% colloquial)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
(Intercept)	-1.28	.90	-1.42	.16			
Context							
(Baseline: Job interview)						155	6
Peer-to-peer	3.06	.44	6.99	<.001	***	238	53
Sociolinguistic interview	2.83	.45	6.32	<.001	***	215	56
Noun language							
(Baseline: English)						19	79
Welsh	-1.59	.78	-2.05	.04	*	589	41
Frequency							
(Baseline: Frequent)						375	42
Infrequent nouns	-.04	.36	-.10	.92		233	41

Table 16. Cardiff English home language (total possessives $n = 489$, 55% colloquial)

	Estimate	Std. error	z value	Pr(> z)	Sig.	N	% Coll.
(Intercept)	1.22	.99	1.22	.22			
Noun language							
(Baseline: English)						19	89
Welsh	-2.51	.89	-2.81	.004	**	470	53
Context							
(Baseline: Job interview)						86	17
Sociolinguistic interview	2.47	.40	6.17	<.001	***	206	75
Peer-to-peer	1.64	.37	4.39	<.001	***	197	50
Frequency							
(Baseline: Frequent)						292	60
Infrequent nouns	-.71	.32	-2.21	.03	*	197	47

Discussion

Although we initially expected peer-to-peer to be the most informal context (as no researcher was present), we found that this was not always the case for the students. First, the fact that the students were doing school-related tasks might have added to the formality of the situation. Second, in some of the sociolinguistic interviews, students commented that they did not use Welsh with some of their peers in informal situations. It stands to reason that if the students were using Welsh with peers, but that they speak to each other more naturally in English, then they might end up being slightly more formal than in situations where they would more frequently use casual Welsh (chatting to someone in Welsh). Although the students who reported not using Welsh with peers were from both WHL and EHL backgrounds, the majority of them spoke English at home. We suspect that this may partly account for the hierarchy differences found between the WHL and EHL groups.

On the whole, however, within individual communities, there is a large degree of shared patterns and constraints. The EHL students use the colloquial form at higher rates but share the patterning to at least some extent. The patterns of the WHL students have been acquired. The fact that there are clear regional differences and that they are found in the EHL groups too underlines this.

Other than noun language, none of the internal factors (including those discussed previously) came up as statistically significant. It may be that the generally high rates of the colloquial form left little room to uncover other effects or that none of them were relevant. Analysis of older speakers (who would be expected to have lower rates of the incoming colloquial form) may find that some did once contribute to the variation patterns.

Conclusions

Our aims in this analysis were twofold: to gain greater understanding of ongoing change in the use of the possessive pronoun variants in Welsh and to establish whether students in Welsh-medium schools from English-speaking home environments shared the patterns and constraints of their peers who came from Welsh-speaking homes (and who consequently had more opportunities to use Welsh and acquire sociolinguistic competence).

With respect to ongoing change, we can confirm what previous studies have suggested but also add new aspects to consider. For example, there is a significant difference in the rates of use of the colloquial variant between north and south Wales. Although rates of the colloquial variant were high throughout (58%), Gwynedd speakers used that variant 71% of the time and Cardiff speakers 48%. The overall high rates of the colloquial also serve to confirm that the use of this variant is increasing, especially in the younger age groups (which all our speakers belonged to). The effect of style, which had been previously discussed but not quantitatively analyzed, was confirmed in our analyses, with the most formal situation showing the lowest use of the colloquial form.

The high use of the colloquial variant underlines that despite its exclusion from some dictionaries and grammars (aimed at L2 speakers) and its supposed non-standardness, it is a part of the contemporary Welsh possessive pronoun system and, if the trend continues into future generations, may become the only variant. Currently, the literary and the sandwich variants are the only forms that tend to be discussed with Welsh learners. Our results suggest that it may be necessary to introduce the colloquial variant too, in order to improve the functional competence of students (see Auger, 2002 for a discussion on the use of colloquialisms in formal teaching in Canadian immersion). Even if it turns out that it is not used frequently in writing (our analysis did not examine this), it is a very frequent form in many different formalities of speech. Future work could examine the speech of adult and later learners of Welsh, who may not be exposed to the more colloquial forms.

In terms of the sociolinguistic competence perspective, the Welsh speakers with EHL backgrounds show awareness of the fact that some variants are more appropriate in more formal contexts. They also demonstrate the ability to style-shift broadly similarly to their WHL peers. The other factors constraining the use of the colloquial variant

are again shared across the groups. This is very different from earlier studies examining classroom acquisition of stylistic differences. It seems then that Welsh-medium classes where children with different backgrounds use Welsh together do enable children to acquire sociolinguistic competence.

There remain some differences between the groups, however; by and large, the EHL students view peer-to-peer settings, at least those of the type in our study, as more formal than WHL students and consequently use less of the colloquial form there than in the sociolinguistic interview. It may be, however, that with more opportunities to use Welsh with peers in non-classroom settings, the EHL would quickly shift to a more WHL pattern.

Competing interests. The authors declare none.

Notes

1. The 2021 Census suggests there are around 530,000 Welsh speakers now.
2. For further information, see Welsh Government's guidance on the categorization of schools. Available at: <https://www.gov.wales/written-statement-guidance-school-categories-according-welsh-medium-provision>.
3. This construction can also serve an emphatic purpose, with the conjunctive echo pronoun emphasizing the possessor (e.g., *fy mam innau* 'my mother'), but no emphatic cases were found in the data.
4. This project was funded by the Economic and Social Research Council and the Welsh Government under project number ES/P00069X/1.
5. English was the home language for all the non-Welsh home language students.
6. Pupils by local authority, region, and Welsh-medium type (<https://www.gov.wales/>).
7. It was decided at the data collection stage during the COVID-19 pandemic that a minimum of four speakers per cell would be sufficient, given the timescale of the PhD project.
8. The questions they were actually asked in the interview differed from the preparatory question set to avoid rehearsed, learned, or read answers.
9. The first author (who was the interviewer) has a local Gwynedd dialect, similar to the Gwynedd participants' own. Many of the Cardiff participants noted that their teachers also spoke with a Gwynedd dialect.
10. We also examined whether initial consonant mutation affected variant choice but found that speakers did not appear to have this as a factor constraining their use, despite the fact that the colloquial variant never required mutation, and that mutation was required for the literary and sandwich variants with some pronouns and some words. Thus, we do not present it in our analysis.
11. IPA: *stryd* [stri:d], *fi* [vi], *ti* [ti], *chdi* [χi:], *fo* [vo:], *o* [o:], *fe* [ve:], *e* [e:], *hi* [hi:], *ni* [ni:], *chi* [χi:], *nhw* [ɲu:], *fy* [və], *dy* [də], *ei* [ɛi], *ein* [ɛin], *eich* [ɛiχ], *eu* [ei].
12. *Chdi* is a Northern Welsh variant.
13. The forms *o* and *fo* are used in north Wales, and *e* and *fe* in south Wales.
14. Including possessed nouns where the plural form is lexically distinctive from the singular, such as *dwyllo* 'hands' and *llaw* 'hand', which were coded as *llaw*.
15. Although Welsh distinguishes between formal and informal second-person singular forms, only one example of the latter was observed in the data (where a participant asked a direct question of the researcher). For this reason, it has not been included in the analysis.

References

- Auger, Julie. (2002). French immersion in Montreal: Pedagogical norms and functional competence. In S. Gass, K. Bardovi-Harlig, S. Magan, & J. Walz (eds.), *Pedagogical norms for second and foreign language learning and teaching*. Amsterdam: John Benjamins. 81–101.
- Awbery, Gwen. (1976). *The syntax of Welsh: A transformational study of the passive*. Cambridge: Cambridge University Press.

- Awbery, Gwen. (1994). Echo pronouns in a Welsh dialect: A system in crisis? *Bangor Research Papers in Linguistics* 5:1–29.
- Baayen, Harald. (2008). *Analyzing Linguistic Data Using R*. Cambridge: Cambridge University Press.
- Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steven. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.
- Borsley, Robert D., Tallerman, Maggie, & Willis, David. (2007). *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Coupland, Nik. (2012). Bilingualism on display: The framing of Welsh and English in Welsh public spaces. *Language in Society* 41:1–27.
- Davies, Peredur. (2016). Age variation and language change in Welsh: Auxiliary deletion and possessive constructions. In M. Durham & J. Morris (eds.), *Sociolinguistics in Wales*. London: Palgrave Macmillan. 31–60.
- Deuchar, Margaret, Donnelly, Kevin, & Piercy, Caroline. (2016). ‘Mae pobl monolingual yn minority’: Factors favouring the production of code-switching by Welsh/English speakers. In M. Durham, and J. Morris (eds.), *Sociolinguistics in Wales*. London: Palgrave Macmillan. 209–239.
- Diskin, Chloe. (2017). The use of the discourse-pragmatic marker ‘like’ by native and non-native speakers of English in Ireland. *Journal of Pragmatics* 120:144–157.
- Drummond, Rob. (2011). Glottal variation in /t/ in non-native English speech: Patterns of acquisition. *English Worldwide* 32(3):280–308.
- Drummond, Rob. (2013). The Manchester Polish STRUT: Dialect acquisition in a second language. *Journal of English Linguistics* 41(1):65–93.
- Durham, Mercedes. (2014). *The acquisition of sociolinguistic competence in a lingua franca context*. Bristol: Multilingual Matters.
- Erker, Daniel, & Guy, Gregory. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88(3):526–557.
- Hatton, Lynfa. (1988). The development of the nasal mutation in the speech of school-children. In M.J. Ball (ed.), *The use of Welsh: A contribution to sociolinguistics*. Clevedon: Multilingual Matters. 239–257.
- Johnson, Daniel E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3:359–383.
- Jones, Mari C. (1998). *Language obsolescence and revitalization*. Oxford: Clarendon Press.
- King, Gareth. (2016). *Modern Welsh: A comprehensive grammar*. London: Routledge.
- Kuznetsova, Alexandra, Brockhoff, Per B., & Christensen, Rune. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13):1–26.
- Lichtenberk, Frantisek, Vaid, Jyotsna, & Chen, Hshin-Chin. (2011). On the interpretation of alienable vs. inalienable possession: A psycholinguistic investigation. *Cognitive Linguistics* 22(4):659–689.
- Linford, Bret, & Shin, Naomi L. (2013). Lexical frequency effects on L2 Spanish subject pronoun expression. In J. Cabreilla Amaro, G. Lord, A. de Prada Perez, & J.F. Aaron (eds.), *Selected proceedings of the 16th Hispanic Linguistics Symposium*. Somerville: Cascadilla Proceedings Project. 175–189.
- Mougeon, Raymond, Nadasdi, Terry, & Rehner, Katherine. (2010). *The sociolinguistic competence of immersion students*. Bristol: Multilingual Matters.
- Nance, Claire. (2015). “New” Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalisation. *Language in Society* 44(4):553–579.
- National Centre for Learning Welsh. (2023). Curriculum and course books. Available at <https://learnwelsh.cymru/learning/curriculum-and-course-books/>. Accessed March 14, 2023.
- Nichols, Johanna. (1988). On alienable and inalienable possession. In *Proceedings from the Haas Festival Conference on Native American Linguistics*. Berlin/New York: De Gruyter Mouton.
- Prys, Myfyr. (2016). *Style in the vernacular and on the radio: Code-switching and mutation as stylistic and social markers in Welsh*. Doctoral Dissertation, Bangor University
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. Available at <https://www.R-project.org/>
- Regan, Vera, Howard, Martin, & Lemée, Isabelle. (2009). *The acquisition of sociolinguistic competence in a study abroad context*. Bristol: Multilingual Matters.
- Robert, Elen. (2009). Accommodating “new” speakers? An attitudinal investigation of L2 speakers of Welsh in south-east Wales. *International Journal of the Sociology of Language* 195:93–115.

- Roberts, Anna E. (1988). Age-related variation in the Welsh dialect of Pwllheli. In M.J. Ball (ed.), *The use of Welsh: A contribution to sociolinguistics*. Clevedon: Multilingual Matters. 104–124.
- StatsWales. (2022). Annual population survey—Ability to speak Welsh by local authority and year. Available at: <https://statswales.gov.wales/Catalogue/Welsh-Language/Annual-Population-Survey-Welsh-Language>. Access March 14, 2023.
- Thomas, Alan R. (1988). *A spoken standard for Welsh: Description and pedagogy*. Cardiff: Welsh Joint Education Committee.
- Watkins, T. Arwyn. (1977). The Welsh personal pronoun. *Word* 28(1–2):146–165.
- Welsh Government. (2017). *Cymraeg 2050. A million Welsh speakers*. Cardiff: Welsh Government.
- Welsh Government. (2022). Welsh language in Wales (Census 2021). Available at: <https://www.gov.wales/welsh-language-wales-census-2021>. Access June 27, 2023
- Willis, David. (2016). Cyfieithu iaith y caethweision yn Uncle Tom’s Cabin a darluniadau o siaradwyr ail iaith mewn llenyddiaeth Gymraeg [Translating the language of the slaves in Uncle Tom’s Cabin and representations of second-language speakers in Welsh literature]. *Llên Cymru* 39:56–72.
- Young, Katharine. (2019). *Perceptions of socio-stylistic variation among teachers of Welsh*. Masters thesis, Cardiff University.

Cite this article: Young K, Durham M, Morris J (2024). Possessive pronouns in Welsh: Stylistic variation and the acquisition of sociolinguistic competence*. *Language Variation and Change*, 1–24. <https://doi.org/10.1017/S0954394523000273>