White, Yosiane & Roberts, Gareth. 2022. Variability in speaker expectations of morphosyntactic mutation in Welsh. *Glossa: a journal of general linguistics* 7(1). pp. 1–44. DOI: https://doi.org/10.16995/glossa.8730

Variability in speaker expectations of morphosyntactic mutation in Welsh

Yosiane White, Radboud University, NL, yosiane.white@ru.nl
Gareth Roberts, University of Pennsylvania, US, gareth.roberts@ling.upenn.edu

Like all modern Celtic languages Welsh exhibits initial consonant mutation with both lexical and morphosyntactic triggers. Owing to the complexity of the system and the sociolinguistic situation of Welsh, change and variation in the system seems inevitable, and evidence for change has been observed in production. Less work, however, has focused on speakers' attitudes and expectations in perception. We used an online auditory acceptability judgement survey to investigate expectations for different morphosyntactic soft-mutation triggers. Respondents listened to sentences with canonical and non-canonical mutation patterns and used Likert scales to indicate for each sentence whether they would use the same pattern themselves and whether they would expect it from others. Almost all respondents expected some variation, even in their own production, but two main clusters of respondents could be identified: "Conservative" respondents whose expectations were close to canonical mutation patterns and "Variable" respondents whose expectations were considerably more flexible. First-language status was the only demographic variable to predict respondent attitudes, suggesting that L2 Welsh speakers accept noncanonical mutation to a greater extent than L1 Welsh speakers. We also compared different mutation triggers, with the tentative conclusion that apparently identical triggers may not in fact be identical for all speakers, and that trigger *transparency* may be an important factor in predicting variability.

Welsh is one of the six modern Celtic languages. All six exhibit a number of typologically unusual features, including VSO constituent order,[1] inflected prepositions, and *mutation*, a phenomenon whereby the initial phoneme of a word changes in phonologically systematic ways in response to a variety of morphological, syntactic, and lexical—but not phonological or phonetic—triggers. Owing to the complexity of the mutation system and the current sociolinguistic situation of Welsh, it is reasonable to expect the system to be undergoing change, and for it to vary both stylistically and between speakers. Production and acquisition studies have provided evidence for change; so far, however, there has been little work from a perceptual perspective. In this paper we report the results of an online survey study in which we asked speakers for expectations and acceptability judgements on canonical and noncanonical mutation patterns. This study allows us to ask the following question: With regard to mutation, what are the boundaries of what Welsh speakers expect Welsh to be?

In Section 1 we lay out what we will call the canonical mutation system; in Section 2 we explain how this canonical system is changing using evidence from the production and acquisition literature; in Section 3 we discuss the reasons why the mutation system is undergoing change. Finally, in Section 5 and 6, we present our survey and its results.

1 The canonical Welsh mutation system

Mutation involves a change in the initial phoneme of a morpheme, usually triggered by an adjacent morpheme. There are three basic forms of mutation in canonical registers, traditionally named *soft*, *nasal*, and *aspirate*, along with a *mixed* category that combines aspirate and soft mutation. The outcomes of the various mutation types are summarised in **Table 1**; in addition to the listed effects on consonants, certain triggers also cause h-prothesis before initial vowels (e.g., *ein* “our” + *afal* “apple” -> *ein hafal*).

As can be seen in **Table 1**, soft mutation—of the three basic forms of mutation—affects the most consonants; it also has around 30 different lexical and grammatical triggers (along with several prefixes)—roughly double the number of triggers for the other mutation types combined. Given this, and given that its triggers include all inflected personal verbs, all subject NPs, and all singular feminine nouns and articles, soft mutation accounts for the vast majority of mutations.

[1] VSO constituent order is the third most common constituent order typologically, but the two most common orders account for most of the world’s languages, and VSO order is very unusual among Indo-European languages (Dryer 2013).

unmutated		soft (SM)		nasal (NM)		aspirate (AM)		mixed (MM)	
orthog.	**IPA**	**orthog.**	**IPA**	**orthog.**	**IPA**	**orthog.**	**IPA**	**orthog.**	**IPA**
p	p	b	b	mh	m̥	ph	f	ph	F
t	t	d	d	nh	n̥	th	θ	th	θ
c	k	g	g	ngh	ŋ̊	ch	χ	ch	χ
b	b	f	v	m	m	—	—	f	v
d	d	dd	ð	n	n	—	—	dd	ð
g	g	∅	∅	ng	ŋ	—	—	∅	∅
m	m	f	v	—	—	—	—	f	v
ll	ɬ	l	l	—	—	—	—	l	l
rh	r̥	r	r	—	—	—	—	r	r

Table 1: Canonical Welsh mutation patterns. The symbol ∅ indicates that the consonant in question disappears, while a dash (—) indicates that the consonant remains unchanged.

Mutation occurs within complex words (e.g., *an-* “non-” + *dibynnol* “dependent” -> *annibynnol* “independent”) but also—and more notably—across word boundaries (e.g., *fy* “my” + *cath* “cat” + *gwirion* “silly” -> *fy nghath wirion* “my silly cat”). Historically mutation arose out of phonetically conditioned alternations. Soft mutation, for instance, is chiefly a vestige of intervocalic lenition (Ball & Müller 2002: pp. 53–72). In the modern language, however, the original phonetic conditioning environments have generally long since disappeared, mainly as a result of the loss of final unstressed syllables as Old Welsh emerged from its ancestor Brythonic (Ball & Müller 2002; Willis 2010). This means, for example, that soft mutation may be triggered by morphemes ending in voiceless consonants (as with *cath wirion* [ka:θ ˈwɪrjɔn] “silly cat”). Another consequence of this process is that homophonous morphemes may trigger different mutation patterns. A good example of this is provided by the third-person possessive proclitics *ei* “her”, *ei* “his”, and *eu* “their”, all of which are pronounced [i] in ordinary speech.[2] They are distinguished, however, by the mutation of the following word. This is shown in Examples 1–3, which also serve to illustrate that the scope of a mutation trigger consists only of the immediately adjacent word.

[2] These forms are sometimes given diphthongal pronunciations in more formal or careful speech, which in the north yields a distinction between *ei* [ei] and *eu* [əi]. The orthographically identical forms are, however, never distinguished. In most southern varieties, all three forms are pronounced identically even if diphthongized (usually as [əi]).

(1) i χa:θ ˈwɪrjɔn
POSS.3FS AM\cat SM\silly
ei chath wirion
“her silly cat”

(2) i ga:θ ˈwɪrjɔn
POSS.3MS SM\cat SM\silly
ei gath wirion
“his silly cat”

(3) i ka:θ ˈwɪrjɔn
POSS.3P cat SM\silly
eu cath wirion
‘their silly cat’

Mutation triggers can be either lexical or morphosyntactic. Most triggers are in the former category, which includes a variety of prepositions, numerals, determiners, conjunctions, clitics, particles, and other function words. Because the original phonetic contexts triggering mutation have disappeared, lexical mutation triggers cannot be identified by form; they must be learnt individually. Nor can they be identified by morphosyntactic category. Many prepositions trigger soft mutation, for instance, but some trigger nasal or aspirate mutation, while others trigger no mutation at all.

Our survey focused on morphosyntactic triggers. Morphosyntactic triggers constitute a considerably smaller set than lexical ones;[3] by their nature, however, they involve a relatively large number of lexical items (e.g., all singular feminine nouns or inflected personal verbs) so are not infrequent. Membership of a mutation-relevant morphosyntactic class is not necessarily predictable, however; aside from some suffixed words, for instance, feminine nouns in Welsh are not reliably recognisable by form.

Morphosyntactically triggered soft mutation occurs on:[4]

1. nouns following pre-nominal adjectives
2. adjectives following feminine singular nouns

[3] Given that mutation arose out of phonetically conditioned alternations, morphosyntactic triggers are necessarily the result of reanalysis, often statistically motivated (cf. Bybee & Hopper 2001). The fact that adjectives undergo soft mutation following feminine singular nouns, for instance, results from the high proportion of feminine nouns that historically ended in a vowel.

[4] There is also a set of mutations, the so-called “free mutations” (Oftedal, 1962), which might potentially be analysed as syntactic and can occur utterance initially. These include the soft mutation of nouns used vocatively, the mixed mutation of negated verbs, and the soft mutation of interrogative verbs. As Ball and Müller (2002: p. 32) argued, however, these are better treated as having “zero lexical triggers”, as all have a stylistic alternative in formal registers in which the mutation is triggered by an overt particle.

3. singular feminine nouns following the definite article
4. noun phrases used adverbially
5. items directly following an overt subject NP
6. the direct object of an immediately preceding finite personal verb
7. items following an intervening phrase (e.g., a sentence modifier intervening between a verb and its object)

Such triggers form an interesting test case for looking at variation in the mutation system for a number of reasons. One reason is that their application across many lexical items means that they are frequent across sentences but abstractly defined. This has theoretical implications. It is fairly typical to treat triggers 5 and 6 as the same trigger: subject NPs, which can be overt or (especially in more formal and literary registers) not overt. This suggests that Welsh speakers should treat both environments as equivalent with regard to mutation. Some researchers have gone further and accounted for 5–7 by arguing that any c-commanding maximal projection (overt or not) triggers soft mutation (for a review and discussion of this and related accounts, see Borsley, Tallerman, & Willis 2007; pp. 226–230). If so, we should expect that users of Welsh should be equally likely to implement and to expect soft mutation in response to any such phrase. On the other hand, we might expect that some triggers are "stronger" than others, that is, more likely to trigger mutation reliably. A potentially relevant feature here is the *transparency* of the trigger in question, by which we mean the degree to which the trigger is easy to learn and identify, especially by L2 speakers (see Section 3.1). The adverbial use of noun phrases is, for example, a particularly opaque trigger. That is, the mutation is triggered not (as for most instances of mutation) by an adjacent overt element, but by the abstract role the noun phrase is playing in the sentence. There are also transparency differences within triggers. Natural-gender feminine nouns (like *merch* "girl"), for instance, can be thought of as more transparently feminine than inanimate feminine nouns (like *pont* "bridge"). If, as a result of change, speakers mutate less reliably, then there may be evidence of variation between triggers with respect to their "strength" or reliability, which in turn may be related to their relative transparency. If the maximal-projection account is correct, however, then relevant triggers should not vary in reliability (because, underlyingly, they are the same trigger), regardless of transparency.

In the study presented here we investigated morphosyntactic triggers 2–6 listed above. For the gender-based triggers (2 and 3) we used natural-gender nouns and inanimate nouns to investigate transparency. We also compared mutation of direct objects of finite personal verbs with overt and covert subject noun phrases so as to investigate the maximal-projection account. Our key questions are thus as follows: To what extent do Welsh speakers' expectations of and preferences for mutation vary, to what extent does variation depend on demographic factors, and

to what extent is there variation between triggers? There is certainly long-standing evidence of variation and change in Welsh speakers' production of mutation, which we discuss in Section 2.

2 How is the Welsh mutation system changing?

The idea that the mutation system might be changing is not novel. It was stated in the November 1904 issue of the literary magazine *The Athanæum* (p. 621) that "If Welsh loses its mutations as South Wales is doing slightly, we shall be sorry." Nor is change surprising. Linguistic systems are typically dynamic and subject to change. We should thus expect some degree of change in the Welsh mutation system even in the absence of contact with English. All the same, Welsh is a minority language whose speakers are almost without exception also speakers of major world languages; such levels of contact are well established to cause change in the minority language, including the simplification of complex grammatical systems (Dorian 1978; 1980; Wolfram 2004; Matras 2020).

Changes to the Welsh mutation system have been systematically documented only since the 1970s. Before then, relevant observations were to be found mostly in theses documenting specific regional dialects of Welsh and were typically strictly descriptive, with little attempt to draw broader inferences from the data (M. C. Jones 1998, p. 3). One of the first systematic approaches came with the work of M. Jones and Thomas (1977), who (among other points) noted the variability of aspirate mutation in the spoken language. They also commented on the shift from mixed to soft mutation for negated verbs and the replacement in some varieties of nasal mutation with soft mutation. Similarly, Awbery (1986) suggested that the Welsh mutation system was simplifying from a four-way system to a two-way system, whereby aspirate and nasal mutation fall gradually out of use, with soft mutation replacing them in at least some environments. M. C. Jones (1998) introduced an apparent-time dimension, finding that—in two different communities—older adults retained and used the traditional mutation system better and more consistently than younger adults.

Besides the loss of aspirate and nasal mutation in favour of soft mutation, the best-explored change to the mutation system involves grammatical gender (Thomas, 2001). The loss of grammatical gender is well-documented in many languages, and it is therefore unsurprising that grammatical gender has been a topic of interest for language change in Welsh (Gathercole & Thomas, 2005; Gathercole, Thomas, & Laporte, 2001; Hammond, 2016; Thomas & Gathercole, 2005). That this coincides with studying change in the mutation system provides an exciting opportunity for understanding change in each of the two systems in the context of the other (in line with Weinreich, Labov, and Herzog's, 1968, embedding problem). However, it also creates a potential confound for empirical research: If speakers do not mutate a given feminine noun or adjective, is this because they do not mutate feminine nouns or adjectives or because they do not treat the noun in question as feminine?

Gathercole et al. (2001) studied the productive abilities of children aged 4–10 in Welsh-only homes, using a battery of tests to investigate children's interpretation and use of the Welsh gender system. They found that children both overextended the mutation of adjectives to masculine contexts and underextended the mutation of nouns in feminine contexts. The authors suggested that their participants might be acquiring mutation patterns item by item rather than as generalisable rules. Thomas and Gathercole (2005) used a story-telling task to elicit semi-naturalistic speech data from Welsh adults and children. Looking at article + noun + adjective constructions, they found that adults had a relatively good productive knowledge of the mutation and gender system, producing soft mutation for feminine nouns 90% of the time. They nonetheless observed a certain amount of variability for soft-mutation, noting that even the same speaker might mutate a feminine noun in one sentence but not mutate the same noun in the next. Other speakers were found to mutate the noun after the article, but not the following adjective (e.g., *y gath du* for *y gath ddu* "the black cat"), with a production experiment indicating that only around 60% of adults mutate adjectives after feminine nouns. They further noted that a few of the adults had what were canonically soft mutated forms as base forms (*dylluan* and *gwningen* instead of *tylluan* and *cwningen*), and therefore "double mutated" the nouns when soft mutation was appropriate (resulting in *y ddylluan* or *yr wningen* instead of *y dylluan* "the owl" and *y gwningen* "the rabbit"). Gathercole and Thomas (2005) studied the production of soft mutation on native Welsh nouns, and on borrowed nouns from English. They found that participants were very reluctant to mutate borrowed nouns (see also Bellin, 1988). They also compared animate and inanimate nouns (the gender of the former being more transparent) and found that this influenced responses for some tasks.

As noted above, how gender is treated matters, as it brings with it a potentially serious confound, given that mutation constitutes the only way in which gender is reliably morphologically marked in Welsh (Hammond, 2016). This means that change in mutation behaviour in the context of gender may reflect change in the gender system, change in the mutation system, or both. This is exemplified in a study by Hammond (2016) showing through statistical natural language processing techniques that soft mutation is the only remaining factor that allows learners to disambiguate grammatical gender in Welsh. Hammond argued that the weakening of the mutation system will mean that learners guess masculine more and more often, eventually leading all nouns to move into the masculine category, removing the functional utility of mutation as a marker for gender altogether. The decline in use of grammatical gender and decline in use of mutation are thus closely connected.

As well as addressing this question in their 2005 work by comparing responses to animate and inanimate nouns, Thomas and Gathercole (2007) suggested that to understand the acquisition of the grammatical gender system, we must look at the acquisition of the mutation system in non-gender-related contexts. By comparing results from studies on mutation in these

contexts (e.g., Ball, 1988; Ball & Müller, 2002) to studies on the grammatical gender system, it becomes possible to compare the two. While this is useful in allowing us to get a global idea of how the trajectories of these two systems compare, the comparison is not ideal as non-gender-related mutation triggers differ in frequency, salience, and complexity from gender-related triggers. It is very possible that some mutation subsystems are changing less rapidly than others.

Taken together, production studies on the acquisition and use of mutation in Welsh suggest that the system is undergoing change: aspirate and nasal mutation are declining in use, and the use of the soft mutation system is becoming more variable, including for many adults. However, some evidence, such as Thomas and Gathercole's (2005) and Gathercole et al.'s (2001) over-mutation finding, suggest that, instead of simply declining in use, the system may instead be undergoing systematic restructuring and simplification. This can be better understood if we understand the reasons for change in the system.

3 Why is the Welsh mutation system changing?

In bilingual contexts where the utility of the minority language is in decline due to a prestige language like English, simplifications of grammatical structures is common, especially if they are not shared with the prestige language (e.g., Dorian, 1978 on Scottish Gaelic). Welsh speakers are almost categorically bilingual, and since Welsh is the minority language it fits this profile well. Additionally, the Welsh mutation system is complex to learn and carries a low functional load. Simplification of the mutation system should therefore not be surprising. Section 3.1 discusses simplification of linguistic structures in bilingual contexts. Section 3.2 discusses the present sociolinguistic situation in Wales. Section 3.3, finally, explains how the learnability and functional load of the mutation system also contribute to why it is changing.

3.1 Minority language simplification in bilingual contexts

In cases of extensive contact between a prestige and minority language, with near categorical bilingualism among the minority language speakers, changes to the minority language are common. Changes include the simplification of complex structures in the minority language (Dorian, 1973, 1977, 1978, 1981; Dressler, 1972; Lavandera, 1978) and linguistic transfer (Gafter & Horesh, 2020; Nagy, 2015), both of which are well attested in Welsh (Boon, 2014; M. C. Jones, 1998; Morris, 2013, 2021; Thomas & Gathercole, 2005). This simplification can involve both the loss of complex grammatical structures altogether, and the generalisation of certain forms to a wider range of contexts (Maher, 1991; Silva-Corvalán, 1986). The question of which changes happen has been tied to efficiency demands, such that harder-to-learn structures that are complex to use are likely candidates for simplification (Boon, 2014; Hymes, 1971).

Although these types of changes are often interpreted as a step on the way to language death, Dorian (1994) argued—with reference to numerous case-studies—that such simplification and restructuring can in fact make minority language more versatile and viable for survival. In such situations, older speakers tend to be the most conservative, with younger speakers more variable.

This research suggests that, ultimately, we should expect the Welsh mutation system to transform into a fairly reliable but considerably simpler system, although its complete disappearance is certainly a possible alternative (Dorian, 1973; Hymes, 1971). On the level of specific triggers, the loss of grammatical gender argued to be occurring in Welsh (Hammond, 2016; Thomas & Gathercole, 2005) removes the basis of soft mutation of at least most feminine nouns and adjectives. We should expect one of three things to happen. First, mutation of nouns following articles and of adjectives following nouns could disappear entirely (as Hammond, 2016, suggested is happening). Second, gender-based mutation could remain, but be restricted to the most transparently feminine nouns (primarily natural-gender nouns). The third possibility is that mutation could be extended to all nouns following articles and all adjectives following nouns. Dresher (2000) argued that many factors can influence the direction of levelling, and that it is possible that the assignment is random, which could lead to widespread variation (Dorian, 2010), consistent with Gathercole et al.'s (2001) observations of productions by young Welsh learners.

In the study presented here we provide a complement to such production-based work by collecting acceptability judgements from Welsh speakers of canonical and non-canonical mutation patterns. This allows us to identify to what extent speakers (a) accept or expect extension of mutation to non-canonical contexts (*over-mutation*), and b) accept omitting mutation in canonical contexts (*under-mutation*). This work provides a meta-linguistic perspective on the simplification of a complex grammatical system in a bilingual minority-majority language contact context. In Section 3.2 we discuss that context in more detail.

3.2 The health and sociolinguistic situation of Welsh

Although certainly a minority language, Welsh is by far the healthiest of the Celtic languages. It is the only one not listed by the UNESCO Atlas of World Languages in Danger (UAWLD) as endangered, being listed merely as vulnerable (Moseley, 2010). The 2013–2015 Welsh-Language Use Survey found that 318,800 people in Wales (11% of the population; 47% of the total number of speakers, as estimated by the same survey) considered themselves fluent speakers of Welsh, while a further 21% of speakers considered themselves to know "a fair amount"; 360,000 (13% of the population of Wales) claimed to speak Welsh daily; only 5% of speakers claimed to never use the language. This contrasts clearly with the situation of all the other Celtic languages. The

next healthiest is Irish, for which the 2016 Irish census (Central Statistics Office, 2020) reported that only 10% of speakers used Irish at least weekly, while 23.8% said that they never speak it.

Nonetheless, there remains cause for concern for the future of Welsh. **Figure 1** shows the proportion of the Welsh population that speaks Welsh between 1891 and 2011. In spite of a recent plateau, the overall trajectory since 1891 has very clearly been downwards. The number of fluent speakers as recorded in the 2013–2015 Welsh-Language Use Survey was lower than in the 2004–2006 survey (although the number of non-fluent speakers had increased).

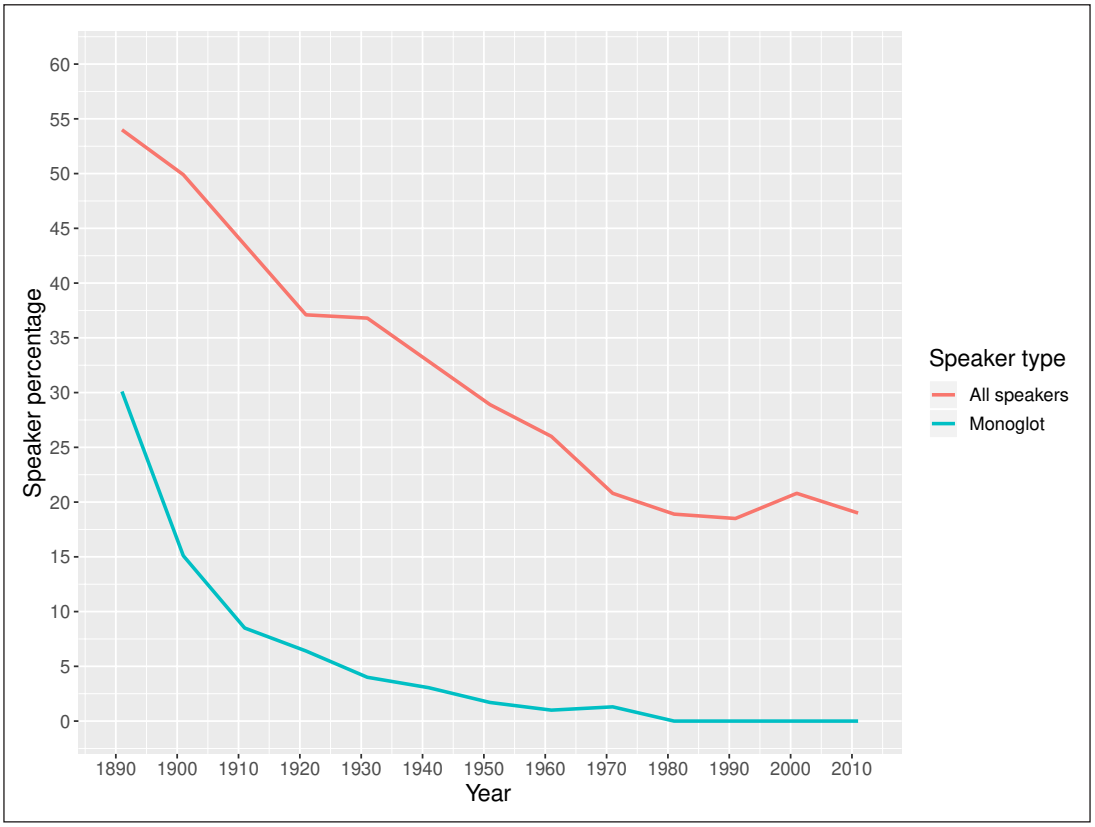

Figure 1: Welsh speakers as percentage of Welsh population (including all speakers and monoglot speakers) based on census data for every decade between 1891 and 2011 (no census was conducted in 1941). To save space the y axis runs from 0 to 60%.

Not surprisingly, decline in the use of Welsh is accompanied by an increase in the use of English. With the exception of some preschool children and dementia sufferers with second-language attrition, all Welsh speakers in Wales also report speaking English. Over 50% of Welsh speakers lack a parent who speaks Welsh fluently, and a little over 50% of speakers report being more comfortable speaking English than Welsh. This situation poses a danger to the long-term survival of Welsh, and a more immediate danger to the survival of complex grammatical systems such as mutation.

The main counter-pressure to these demographic forces is education. All school pupils attend Welsh classes, and in the 2019/2020 school year 108,638 school pupils (23%) were educated in other subjects through the medium of Welsh (StatsWales, 2020).[5] Moreover, the success of Welsh-medium education and the Welsh government's commitment to supporting the language mean that the trend is for more Welsh-medium schools to open (or for traditionally English-medium schools to introduce more Welsh instruction) rather than the reverse. While this is a positive trend in terms of not only maintaining but even increasing the number of Welsh speakers, it has the inevitable consequence of increasing the proportion of Welsh speakers who acquired the language in school rather than at home. As noted above, less than half of Welsh speakers have a parent who speaks the language fluently.

Differences between earlier and later learners of minority languages have been noted elsewhere. In a study of the revitalisation of Scottish Gaelic, Nance (2015) showed that young learners (or "new speakers") of Scottish Gaelic sound different from "traditional speakers" of the language. She concludes that new speakers use different forms of the language and use the language for different communicative reasons than traditional speakers.

This work is supported by game theoretic models of minority language shift that show that when bilinguals of a majority and minority language communicate they tend to use the majority language if they are unsure about whether their interlocutor is bilingual, even if they are motivated to use the minority language as often as possible (Iriberri & Uriarte, 2012; Uriarte & Sperlich, 2021). Within bilingual populations, speakers converge to a stable equilibrium in which many bilinguals shift to use the majority language even if the minority language would have been communicatively available for use. Uriarte and Sperlich (2021) showed that this communicative convention aligns with data on the daily use of Welsh as determined by The Census of Wales 2004–06 and 2013–15.

There are two main routes for young L2 Welsh learners to acquire mutation: through explicit instruction in school and through exposure to conservative varieties of Welsh. The first is straightforward: all school pupils in Wales have Welsh classes, in which the mutation system is taught explicitly at least to some extent. The second is more complex; in some regions exposure to conservative speakers may be rare. Formal literary Welsh would provide the most consistent exposure to canonical mutation patterns, but—even assuming reliable exposure to literary Welsh on the part of young learners—it differs substantially from colloquial registers and in that respect provides a poor model for them (cf. Ball, 1992). Less formal standardised varieties, as used in broadcasting by the BBC and the Welsh-language television channel S4C, provide better and

[5] This figure is for both primary and secondary schools and covers a range of school types, including some in which Welsh and English are both used for some subjects; 96,770 pupils (89%), however, attended schools in which most classes were taught only in Welsh; 77,635 of these attended schools in which Welsh was also the language of day-to-day business in the school.

more accessible models. However, broadcast Welsh varies considerably not only in register but also in the extent to which canonical mutation patterns are observed. (It is notable in this context that subtitles on BBC videos often "correct" non-canonical mutation patterns).

Regional variation may be relevant here. Welsh is divided into a number of regional dialects, with the clearest distinction being a north–south divide. Although, according to the 2011 census, roughly 58% of Welsh speakers are located in the more densely populated south, the areas with the highest proportions of Welsh speakers are located in the north, and the north of Wales is typically perceived as being more Welsh-speaking. Some work has related changes in the mutation system to specific regional varieties. Awbery (1986), for instance, treated simplification of the mutation system—realised as the loss of aspirate and nasal mutation accompanied by the spreading of soft mutation to new contexts—as a feature of (some) southern dialects. However, simplification also occurs in northern varieties, and it is not clear whether the apparent association with southern varieties might be better accounted for in terms of other sociological variables, such as parental fluency or age of acquisition, which also vary regionally.

In summary, the bilingual sociolinguistic situation in Wales seems likely to explain a great deal of the variability observed in studies like that of Gathercole et al. (2001). In the perception study we report in this paper we thus expected to find more variable acceptability judgements from younger "new" speakers, and more conservative judgements from older "traditional" speakers (Dorian, 1994; Nance, 2015; O'Rourke & Pujolar, 2015). It may also be (though it is less certain) that there are regional differences between north and south.

3.3 Functional load and learnability

In Section 3.1 we discussed how grammatical structures that are hard to learn or have a weak communicative value are particularly susceptible to simplification in bilingual language-contact contexts. The Welsh mutation system is a good example of this: Most mutation serves no disambiguating function. The sentence in example 4, for instance, does no better job of conveying the information that the speaker is moving to Bangor than *dwi'n symud i Bangor*. In general, in other words, mutation seems likely best treated as a grammatical convention serving little communicative function. (Though it might serve a social function as an indicator of linguistic prestige; Prys 2016; see also Roberts & Fedzechkina, 2018 for discussion of social niches and the maintenance of redundant linguistic forms.) However, there are a handful of contexts in which mutation might be argued to bear some functional load in distinguishing an intended proposition from similar, unintended ones. Direct-object mutation is a particularly obvious example of this. Given that Welsh has VSO constituent order and allows pronominal subjects to be dropped, the mutation of direct objects potentially allows subjects to be distinguished from objects in sentences like those in examples 5 and 6.

(4) dw i ‘n symud i Fangor
be.PRES.1S SBJ.1SG PROG move.INF to SM\Bangor
‘I’m moving to Bangor’

(5) llyncodd modryb Aled wrth glywed y newyddion
swallow.PST.3SG aunt Aled at hear.INF the news
‘Aled’s aunt swallowed when she heard the news’

(6) llyncodd fodryb Aled wrth glywed y newyddion
swallow.PST.3SG SM\aunt Aled at hear.INF the news
‘he/she/it swallowed Aled’s aunt on hearing the news’

The mutation of feminine nouns and adjectives might also be said to play some disambiguating role, as it allows homophones to be distinguished, provided they differ in gender. For instance, the Welsh words for the absolute direction “south” and the relative direction “right” are in origin the same word and are identical except that the former is masculine and the latter feminine, allowing *y de* to be distinguished from *y dde*. A similar point can be made about some lexical triggers, such as the possessive clitics in Examples 1–3.

In fact, in a recent study of adult Welsh heritage language speakers, Boon (2014) found that simplification of the mutation system among heritage speakers living in London depends on the functional load of individual forms. She suggested that mutations that facilitate successful communication and serve a disambiguating function are maintained and that functional load is a better predictor of retention than frequency of trigger. On the whole, however, the functional load of Welsh mutation is generally very low. While it is easy to imagine situations in which it is helpful that mutation allows “turn right” to be distinguished from “turn south”, mutation is no help in distinguishing many other homophones (e.g., *ysgol*, “ladder”, “school”), and in some cases it creates them (e.g., *moch* “pigs” and *boch* “cheek” both become *foch* as a result of soft mutation). Even when mutation does disambiguate, it tends to be hard to contrive examples of genuine ambiguity. Indeed, language users generally are extremely tolerant of lexical ambiguity, employing context and repair mechanisms to navigate potential miscommunication (Galantucci, Roberts, & Langstein, 2018; Tzanne, 2000; Vitello & Rodd, 2015). Furthermore, Welsh has alternative grammatical mechanisms for clarification in many cases. Possibly one of the strongest cases for mutation playing an important functional role concerns the three possessive clitics pronounced /i/, and it is certainly not hard to come up with examples where a lack of mutation leaves the gender and number of the possessor ambiguous. However, echoing pronouns can be (and, in speech, very often are) added post-nominally (**Table 2**). It is also worth noting that gender distinctions in this domain are not made at all in quite a number of languages, suggesting that their absence does not pose a serious communication problem.

English	standard Welsh	colloquial Welsh (without mutation)
her car	ei char (hi)	ei car hi
his car	ei gar (o)	ei car o
their car	eu car (nhw)	eu car nhw

Table 2: Examples of possessive constructions in standard registers (where mutation is expected) and colloquial registers (where its use is variable).

Similar points can be made about distinguishing subjects from objects following inflected personal verbs. If there is indeed a risk of an object being taken for a subject, the subject can be inserted. In any case, constructions with inflected personal verbs are relatively uncommon in many (chiefly northern) colloquial varieties, where periphrastic constructions are preferred.

In summary, Welsh mutation might play some functional role, but it bears an extremely light functional load when frequency and the availability of alternative disambiguation-avoidance strategies are taken into account. To the extent that we might expect functional pressures to support the retention of a grammatical convention (see, e.g., Koplenig, Meyer, Wolfer, & Mueller-Spitzer, 2017, on the role of functional pressures in typological patterns of constituent order and case marking), there is little reason to expect this to play a very important role in the retention of the mutation system in Welsh. To be retained, it needs rather to be acquired as an arbitrary grammatical convention, and the system is non-trivial to learn. While the changes themselves (essentially frication, voicing, and nasalisation) are rather systematic and straightforward, they are triggered by a diverse and unpredictable range of lexical items and morphosyntactic contexts, with few phonological or morphosyntactic clues as to what item will trigger what kind of mutation, if any. The continued existence of mutation over many centuries confirms that it is acquirable. However, those were centuries in which transmission was predominately to young children; as the average age of acquisition rises, we might expect a change in what parts of the grammar are easily learnable (cf. Dale & Lupyan, 2012; Trudgill, 1986; Wray & Grace, 2007). Given this, the number of late learners among the Welsh-speaking population matters quite a lot for the future of mutation.

4 Predictions

Previous production studies on the Welsh mutation system have found variability in the system, which is in line with long-standing literature on language change in bilingual minority-language contexts (Dorian, 1978, 2010; Dressler, 1972; Lavandera, 1978). Our study is an online acceptability judgement survey focusing on respondents' perceptions and expectations of mutation in Welsh.

We had several predictions. The first concerns demographic patterns. Although change occurs in the absence of contact, we should expect the influence of bilingualism to be key to understanding change and variability in the modern Welsh mutation system. This leads us to predict that respondents' L1 status (i.e., whether they consider Welsh or English, or both, to be their first language) should play an important role in explaining variability in our data. In line with work on Scottish Gaelic by Nance (2015), among others, we expect that L1 Welsh speakers will show more conservative acceptability judgements than L2 Welsh speakers and bilinguals. Other demographic factors that can be expected to play a role include age and regional background. We should expect age to matter if there is a change in progress, with older speakers preferring more conservative mutation patterns to younger speakers. Finally, we should expect regional background (i.e., whether a respondent is from the north or the south of Wales) to matter (a) if there is regional variation in the mutation system, as suggested by Awbery (1986), and (b) because Welsh speakers make up a larger proportion of the population in the north of the country.

Our second prediction concerns the nature of the mutation trigger. The maximal-projection account, as discussed in Section 1, holds that a number of potentially distinct triggers are in fact examples of the same trigger, specifically c-commanding maximal projections (Borsley et al., 2007; Tallerman, 2006). This implies that triggers that can be treated as falling under this category should be treated the same by speakers. Alternatively, it could be that the triggers involved are learnt and treated as distinct, at least by many speakers. To get at this we included sentences in our data in which direct objects directly follow inflected personal verbs with no overt subject and sentences in which they directly follow an overt subject. If responses differ for these two sentence types, that suggests that these may operate as distinct triggers.

Our third prediction concerns trigger *transparency*. Given the demographic prediction above and in the context of ongoing change in the mutation system, it seems likely that not all triggers are changing to the same extent. One possibility is that this is due to the functional load of the trigger (Boon, 2014). Another (though not mutually exclusive) possibility is that this is due to the transparency of the trigger. That is, how straightforward is it to learn and identify the trigger, especially for L2 speakers? We investigated this with respect to gender triggers by including natural-gender nouns (where natural gender matches grammatical gender and may aid recognition of the latter) and inanimate nouns. The natural-gender nouns, we anticipate, should—by virtue of greater transparency—be a more reliable trigger than the inanimate nouns. We also included sentences with adjectives following natural-gender nouns; since the adjectives themselves are gendered only on the basis of the noun being modified (and thus often occur unmutated in the same grammatical context, i.e., directly following a noun) we predicted that this would be an even less reliable trigger. For comparison, we also included adverbial noun

phrases, already known as a variable trigger (even in canonical contexts), and one that is very abstractly defined.

We took two approaches with our analysis. With respect to our first prediction, we looked at the entirety of the data and tested the role of different demographic factors using a linear mixed-effects model. For the second and third predictions, which required that we separate out data for different triggers, we did not have enough data for any given trigger for the linear-model analysis to be appropriate. In this case we therefore performed a k-means clustering analysis to identify patterns in the data for comparison. The goal was primarily to provide preliminary observations as a basis for future investigation focusing on specific triggers with larger samples.

5 Method

In this study, 49 respondents took an online acceptability judgement survey in which they listened to a series of Welsh sentences with canonical and non-canonical mutation patterns and, for each sentence, judged on a 5-point Likert-scale whether (1) they would use the mutation pattern themselves and (2) they would be surprised if another speaker used it.

Besides providing a perception complement to the production literature on this topic, there are several further benefits to using this design. First, we were able to recruit a relatively broad sample of Welsh speakers from across Wales, and of varying ages and first language backgrounds (L1 Welsh, L2 Welsh, and balanced bilinguals). Also, by contrast to traditional methods of separately collecting acceptability judgements from consultants, our design ensured that each respondent heard exactly the same test sentences, with no variation that might influence judgements.

5.1 Respondents

A total of 49 respondents were recruited online. 18 respondents were recruited through the crowdsourcing platform Prolific, and 31 respondents were recruited through social media platforms such as Twitter and Facebook, as well as via Welsh online student groups at universities in the UK.

All respondents needed to be fluent speakers of Welsh to participate, as all of the written and spoken materials (consent, instructions, audio files, response options, etc.) were in Welsh. We further collected demographic information on the respondents' Welsh language background (see Section 5.3). All respondents reported being fluent speakers of Welsh; 26 (53% of) respondents reported learning Welsh as their first language at home, 12 (24%) were raised bilingually from birth, and ten (20%) learnt English first at home. Further, 60% of respondents reported using Welsh either daily or weekly, with the remaining respondents using it either once a month or a few times a year. When asked how proficient the respondents

thought they were in Welsh on a scale from 1 (not at all proficient) to 5 (highly proficient), the mean score for speaking was 4.0 (SD = 1.01), and for writing 3.6 (SD = 1.16). Five respondents reported a self-proficiency in Welsh of 2 or lower on both writing and speaking, and one respondent did not report any demographic information; these six respondents were therefore excluded from the study. The age range of the remaining 43 respondents spanned 17 through 73 years of age, with a median age of 31 (mean 36.2) years. There were 24 female respondents (mean age: 37.9 years) and 18 male respondents (mean age: 34.1 years); one respondent declined to inform us of their gender identity. Over 65% of respondents reported having attended university. Respondents were evenly distributed regionally, with 21 (49%) being from the north of Wales, 17 from the south, two from mid-Wales, and two from outside Wales; one gave no location.

5.2 Stimuli

Stimuli for the survey consisted of 110 sentences spoken by a speaker of a northern variety of Welsh. The sentences were recorded in a sound-proof booth using a Logitech UB H390 headset and the recording software Audacity (Audacity Team, 2014). The sound files were amplitude normalized using the phonetics software Praat (Boersma & Weenink, 2019). Each sentence contained one of the following words potentially subject to mutation.

1. A singular noun following the definite article
   (a) Natural gender nouns (12 sentences)
   (b) Nouns gender-marked by feminine suffix *-en* or masculine suffix *-yn* (12 sentences)
2. An adjective following an indefinite noun (12 sentences)
3. A direct object of a personal inflected verb (12 sentences)
4. An adverbial noun phrase (8 sentences)

Our study differs from previous studies of mutation in that it is focused on the *acceptability* of different mutation patterns among adult speakers as opposed to acquisition and production. In order to avoid the confound mentioned above with regard to gender-based triggers, we took an approach similar to that of Thomas and Gathercole (2005). That is, rather than remove gender from the picture entirely and compare gender triggers to non-gender triggers (as suggested by Thomas & Gathercole, 2007), we used animate nouns for which natural gender provides a strong clue to grammatical gender. By natural-gender nouns we mean nouns for which the grammatical gender depends on the semantics of the word. For example, the words *queen, princess, witch, wife, aunt, niece, nun, hen* and *ewe* all denote female referents and their equivalents in Welsh are all feminine nouns. If people do not mutate these nouns where expected, it seems a reasonable conclusion that this reflects change in the mutation system rather than the gender system (or

uncertainty as to the gender of the noun). For comparison, we also investigated mutation of inanimate nouns ending in *-en* and *-yn*, feminine and masculine diminutive/singulative suffixes respectively. If mutation behaviour for these nouns differed from behaviour for natural-gender nouns, this would likely be due to change in the gender system (though not, given the historically rather reliable gendering of these suffixes, *long-standing* variation in noun gendering).

As described above in Section 1, mutation would be canonically expected in the first two sentence types for feminine, but not masculine, nouns. In order to gauge speakers' judgements about canonical mutation, over-mutation (i.e., mutation with masculine nouns) and under-mutation (i.e., non-mutation with feminine nouns), four types of sentences were used for these triggers. First, Masculine Unmutated (*MascUnmutated*) and Feminine Mutated (*FemMutated*) follow traditional mutation patterns. Second, the converse of each, Masculine Mutated (*MascMutated*) and Feminine Unmutated (*FemUnmutated*), targeted over-mutation and under-mutation, respectively. Examples are shown in **Table 3**. (A full set of stimuli is available, along with results, at https://osf.io/hqrz6/.)

Condition	**Canonical?**	**Type 1a example**	**Type 1b example**	**Type 2 example**
Masc-Unmutated	Yes	Mae'r brawd yn byw yno *The brother lives there*	Y blodyn sy'n tyfu *It's the flower that's growing*	Mae tad cyfeillgar rhywun yn helpu *Someone's friendly father is helping.*
Masc-Mutated	Over-mutation	Mae'r dywysog yn gweithio *The prince is working*	Mae'r fochyn yn cysgu *The pig is sleeping*	Mae brawd fach rhywun yn cwyno *Someone's little brother is complaining*
Fem-Mutated	Yes	Mae'r fam yn teithio *The mother is travelling*	Y golomen sy'n hedfan *It's the dove that's flying*	Mae angen mam brofiedig *We need an experienced mother*
Fem-Unmutated	Under-mutation	Mae'r modryb yn gweu *The aunt is knitting*	Roedd y cwningen yn sboncio *The rabbit was bouncing*	Roedd tywysoges pwysig yn dod *An important princess was coming*

Table 3: Example sentences for gender-related mutation triggers.

Sentence types 3 and 4 are not related to grammatical gender. Therefore, unlike for the gender-related triggers, there is no opportunity for over-mutation, so there were only two sentence types: Mutated (canonical) and Unmutated (non-canonical). Of the 12 Type-3 sentences (involving the direct object of a verb), six had no overt subject and six included an overt subject pronoun that directly followed the verb. Soft mutation would be expected canonically in both of these contexts.

5.3 Procedure

The survey was created using the PennController extension of IbexFarm, an online experimental platform (Zehr & Schwarz, 2018). All materials, both written and auditory, were in Welsh. Aside from the appropriateness of this decision given that the stimuli were in Welsh, it also served to discourage less fluent speakers (whose answers might be influenced by unfamiliarity with certain lexical items or grammatical constructions) from taking part.

The survey was prefaced with a written consent form and a test for working audio. An introductory text followed in which the goal of the survey was explained. This survey was presented in a completely transparent format: Respondents were informed from the outset that this was a survey about mutation. They were told that we were interested in how mutation is used and perceived in day-to-day Welsh (as opposed to how it is expected according to prescriptive norms). We elaborated that natural language use often differs from what is taught in school about the language, and that, as linguists, we were not interested in whether people know the “correct” mutation patterns, but in how they would personally use mutation, and expect mutation to be used. This high level of transparency was motivated by two things. First, because Welsh speakers learn the prescriptive mutation rules in school, we were concerned that—if we did not specify what we were looking for—respondents would be inclined to try to answer “correctly” according to these prescriptive rules. Second, the speaker who recorded the audio files was from the north of Wales. As there are substantial differences between regional dialects in Wales (and no such thing as a regionally unmarked accent), we did not want respondents to say that they would not say a given sentence as in the audio file simply because of dialectal differences, as opposed to the mutation pattern. After careful consideration we saw no reason not to be transparent about the goal and purpose of this survey.

Respondents listened to the sentences in a random order. Each sentence played automatically at the start of a trial, and could be replayed an unlimited number of times using a “Play” button. The respondents were required to make judgements on two 5-point Likert scales (Likert, 1932). Each point on the Likert scale was labelled for each question. The two questions, with their corresponding Likert-scale labels, were as follows:

Question 1. Focusing on the mutation pattern in the sentence: would you yourself say this sentence like this?

1. I would always mutate like that
2. I would often mutate like that
3. I don't know
4. I would occasionally mutate like that
5. I would never mutate like that

Question 2. Still thinking about the mutation pattern: How surprised would you be if you heard another Welsh speaker say this sentence like this?

1. I would be extremely surprised. I have never heard a mutation pattern like that.
2. I would be quite surprised. It sounds weird to me.
3. I am unsure whether it would surprise me or not.
4. I wouldn't really be surprised. I have heard people mutate in the same way.
5. I would not be at all surprised. It sounds normal to me.

The two Likert-scale questions were accompanied by an optional free-response box in which respondents could elaborate on their choices, or could comment on the stimuli in more detail.[6]

After responding to all the stimuli, respondents completed a demographic questionnaire that contained questions about age, gender, profession, where they were from in Wales, etc. It further asked when they had learnt Welsh, whether they learnt English or Welsh first, how proficient they considered themselves to be in Welsh, and how often they used Welsh. Finally, respondents were asked to elaborate on whether they thought the mutation system might be changing in modern Welsh, whether they thought that age or region plays a role in this change, whether they have children and notice that they use mutation differently, and whether they have any other thoughts about mutation. On completion of the survey, respondents were compensated £10 for their time.

[6] Interestingly, when words had been mutated non-canonically (e.g., a mutated masculine noun after the definite article), some respondents used this box to indicate that they thought a non-existent word had been used. This strongly suggests that, for those speakers at least, the context for the mutation in question had not expanded.

5.4 Analysis

The data were analysed using R (R Core Team, 2017). As each respondent responded on a 5-point Likert scale to each question, each item was given a score from 1 ("I would always mutate like that") to 5 ("I would never mutate like that"). Our main research questions for this study concerned the strength of Welsh speakers' preference for or tolerance towards canonical and non-canonical mutation patterns. For this reason, we wanted to compare respondents' preference for sentences that contained canonical mutation with their preference for sentences that did not. We therefore transformed the Likert-scale scores into two different scores. We call the first a *Non-normativity score*. This simply involved moving from a 1–5 scale to a 0–4 scale and taking into account the nature of the sentences that respondents were responding to. In the survey itself, a low score on the Likert scale would always indicate a high level of approval of the mutation pattern in the sentence, whether this was canonical or not. In constructing the Non-normativity score, we transformed scores so that a 0 would indicate high approval of canonical mutation and 4 would indicate high approval of non-canonical mutation. For example, if someone rated both "mae'r brawd yn byw yno" (canonical) and "mae'r fochyn yn cysgu" (non-canonical) as 5 on the likert scale, the first would get a non-normativity score of 0 and the second would get a score of 4.

Linear mixed effects models were fit to analyse the differences in Non-normativity scores between the different mutation triggers. The variables Sentence type, Trigger, Age, Gender, Region (North or South), and First Language were fit as fixed effects, and random intercepts for Item and respondent were used. The data were visualised using the ggplot2 library in R (Wickham, 2016); this visual presentation using structured graphs was used to examine general trends and mutation preferences, and differences between the triggers.

The second score, which we call an *Extended non-normativity score*, was used in graphing results. Again, this involved first transforming the original Likert-scale responses so that they ranged from 0 to 4 (with 0 indicating a low rating of the mutation pattern in question and 4 indicating a high rating). A respondent's mean score for sentences with canonical mutation was then subtracted from their mean score for sentences with non-canonical mutation to produce a score ranging from –4 to 4. For instance, a respondent who rated canonical sentences highly (e.g., at 3) and non-canonical sentences low (e.g., at 1) would get a low score (1–3 = –2). This score could be applied to different kinds of sentence so that (as in Fig 2) a respondent's attitude towards including or omitting mutation when canonically expected could be plotted against their attitude to including or omitting mutation when canonically not expected.

We then conducted a k-means clustering analysis of the Extended non-normativity score. K-means clustering is an unsupervised non-linear algorithm for identifying clusters of data in

a multidimensional (in this case two-dimensional) space. Given the coordinates for a set of observations and a pre-determined parameter *k*, the algorithm attempts to find the best way to break up the observations into *k* clusters (Steinley, 2006). We established a value for *k* via silhouette analysis, using the factoextra library in R (Kassambara & Mundt, 2017). This analysis performs clustering with different values of *k* and establishes which value of *k* gives optimal results, by which is meant maximising the distance between clusters while minimising the distance between observations within clusters (Shahapure & Nicholas, 2020). We used 25 initial configurations in the k-means clustering analysis to ensure reasonable optimisation.

6 Results

Data are available at https://osf.io/hqrz6/. In what follows we will start by presenting an analysis of general patterns in the data, collapsing all triggers together (Section 6.1). This will include both a clustering analysis (Section 6.1.1) and a linear-model-based analysis (Section 6.1.2). We will then present a clustering analysis of each trigger separately (Section 6.2).

6.1 General patterns

Our first analysis was an analysis of the overall non-trigger-specific patterns in which we collapsed respondents' responses to the different sentences, and simply asked how strong respondents' preference for canonical mutation patterns was, both when mutation is canonically required and when it is not.

6.1.1 Clustering analysis

Figure 2 displays mean extended non-normativity scores for all respondents across sentence types. We used shape and colour to indicate the language that respondents reported as their first language. The x-axis of this graph represents extended non-normativity scores for sentences in which mutation would be expected (e.g., feminine nouns following the definite article). The leftmost end of this axis denotes a strong preference for mutating canonically, while the rightmost end denotes a strong preference for "under-mutation", that is, omitting mutation that would canonically be required. The y-axis represents extended non-normativity scores for sentences in which non-mutation is canonically required (e.g., masculine nouns for the gender-based triggers). The bottom end of the axis denotes a strong preference for canonical non-mutation, while the topmost end denotes a preference for "over-mutation", that is, extending mutation to a non-canonical context.

This axis structure gives us a graph that can be divided into four quadrants. Respondents who prefer canonical mutation patterns should fall into the bottom-left quadrant of this graph

(the “Conservative” quadrant). Speakers who use very little mutation in general should fall into the bottom-right quadrant (the “Under-mutate” quadrant). Speakers who over-mutate, such as mutating masculine nouns after the article, should fall into the top-left quadrant (the “Over-mutate” quadrant). Speakers who vary in whether or not they mutate canonically should fall in the middle of the graph, around the centre-crux of the quadrants (the “Variable” area). We should not expect the top-right quadrant of this graph to be occupied by many speakers, since it denotes a preference for not mutating when mutation is canonically expected *and* applying mutation when it is not canonically expected; such a direct reversal of canonical patterns would be surprising.

We plotted responses for our two questions (Q1. Would you mutate like this yourself? Q2. Would you be surprised if someone else mutated like this?) separately. We expected responses to Q1 and Q2 to be rather similar but for respondents to have stronger opinions about their own production (Q1) than about the use of mutation by others (Q2). This was the pattern we found. (See supplementary materials for overall responses to Q2 at https://osf.io/hqrz6/.) Therefore the analysis in the remainder of this paper will focus primarily on Q1 data.

The clustering analysis identified two clusters in the overall Q1 data, which are displayed in **Figure 3**.[7] This figure shows the same data as **Figure 2**, except that colour is now used to distinguish clusters (to which we also fitted ellipses based on a 95% confidence level for a multivariate t-distribution). One cluster is located in the bottom of the graph, mostly in the bottom-left Conservative quadrant, though including a few respondents with some tendency for under-mutation (all located towards the bottom left of the under-mutation quadrant, indicating that this tendency is not strong). Broadly speaking, respondents in this cluster prefer canonical mutation rules, rarely over-apply mutation, and only occasionally under-apply it. The second cluster is a diagonally oriented cluster of respondents running through the Variable middle section of the graph. This cluster is smaller than the other cluster, but not very much smaller. It constitutes 42% of the data, an indicator of the extent of variability in the current mutation system. In summary, if all triggers are considered together, our respondents tend to be either rather conservative (i.e., canonical) in their use of mutation, or variable. Very few respondents consistently over-mutate or under-mutate. But are there demographic patterns behind this result?

[7] The results for the Q2 data are similar and are displayed in the appendix included in the supplementary materials for this paper, available at https://osf.io/hqrz6/.

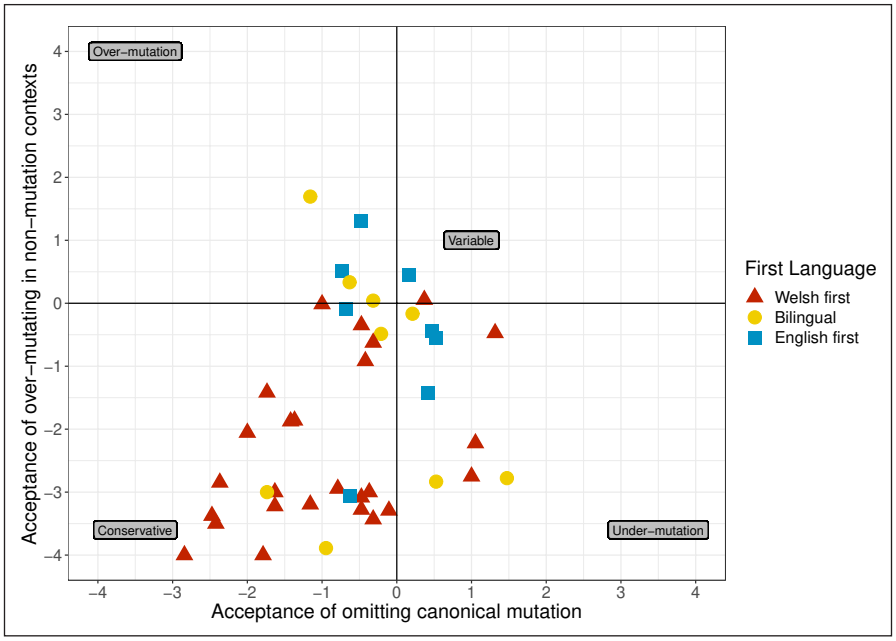

Figure 2: Respondents' overall mutation preferences; colours and shapes indicate first language. Each point on the graph represents a respondent's mean extended non-normativity score. Respondents in the bottom-left prefer canonical mutation patterns. Respondents in the bottom-right do not mutate in canonical contexts. Respondents in the upper-left do mutate in non-canonical contexts. Respondents around the centre of the graph show variable preferences.

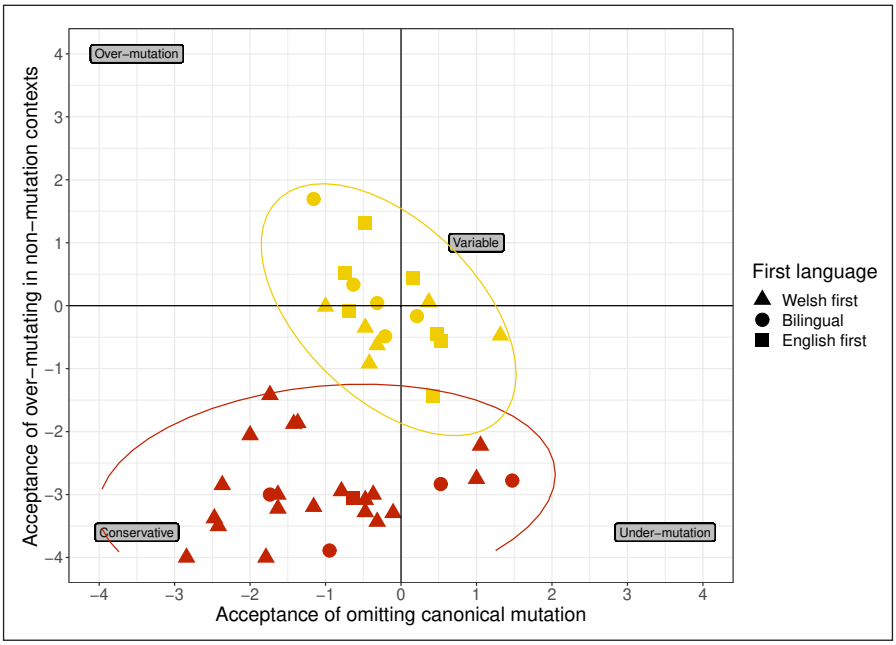

Figure 3: Respondents' overall mutation preferences with k-mean clusters indicated by colour and ellipses; shapes indicate first language.

6.1.2 Linear model analysis

We ran a linear mixed effects model on the data from Q1 with Sentence Type, Age, Gender, Region (North vs. South), and First Language as fixed effects. We found no effect of Age (β = –0.005, d = 0.002, p = 0.12), Gender (β = –0.04, d = 0.02, p = 0.85), or Region (β = –0.06, d = 0.03, p = 0.93) on people's mutation preferences. There was, however, a significant effect of First Language with respondents who learned Welsh first differing significantly from respondents who learned English first (β = 0.50, d = 0.23, p = 0.01). This suggests that speakers whose first language was Welsh were more conservative, while speakers who grew up with English as their first language were more likely to accept non-canonical mutation patterns. Bilingual speakers' responses did not differ from those of respondents who learned English first (β = 0.18, d = 0.08, p = 0.41) or from those from respondents who learned Welsh first (β = –0.32, d = 0.14, p = 0.07). The dependent variable was the Non-normativity score (Section 5.4) for which lower values indicate a greater preference for canonical mutation patterns and higher values indicate a greater preference for non-canonical mutation patterns. The mean scores for different mutation patterns can be seen in **Table 4**. The overall mean score was 1.52, close to the middle of the range. There are, however, important differences consistent with the cluster analysis discussed above. For instance, the Non-normativity score is 0.75 points lower for canonically unmutated sentences than for over-mutated sentences (β = 0.75, d = 0.33, $p < 0.001$), for instance, and the score for canonically mutated sentences is 0.98 points lower than for under-mutated sentences (β = 0.98, d = 0.44, $p < 0.001$). In general, under-mutation was preferred to over-mutation. This is suggested by the lower Non-normativity scores in **Table 4** for sentences that are canonically unmutated and the higher scores for sentences with canonical mutation. It can also be seen in **Figure 2**. By contrast, the higher scores for canonically mutated sentences and under-mutated sentences suggest that when mutation was expected canonically, speakers were rather accepting of sentences without it. This is consistent with an account in which mutation is, or is becoming, variable, perhaps (though we cannot infer this from our data) conditional on register.

Mutation pattern	Non-normativity score
Not mutated canonically/ unmutated in experiment	0.90
Not mutated canonically / mutated in experiment	1.64
Mutated canonically / mutated in experiment	1.28
Mutated canonically / unmutated in experiment	2.25

Table 4: Non-normativity scores for different mutation patterns (collapsed across triggers) for Q1. Non-normativity scores range from 0 to 4, where 0 indicates a strong preference for canonical mutation and 4 indicates a strong preference for non-canonical mutation patterns.

The above results are all based on data from Q1, which asked whether respondents themselves use the mutation pattern they were presented with. We also generated the same graphs and results for Q2 (see the appendix in the supplementary materials at https://osf.io/hqrz6/), which asked whether respondents would be surprised if a conversation partner used the heard mutation pattern. We expected more variation for this question, and this is indeed what was found. Respondents are located less towards the corners of the plot and more towards the Variable central area. This suggests that respondents are used to hearing variable mutation, and are not surprised to hear non-canonical mutation patterns. This demonstrates that even speakers who report conservative mutation patterns in their own language are quite familiar with non-canonical patterns from others, especially under-mutation.

We ran the same model on the Q2 data, and again found no evidence for an effect of Gender ($\beta = 0.03$, d = 0.02, $p = 0.74$), Region ($\beta = –0.05$, d = 0.04, $p = 0.63$), or Age ($\beta = –0.007$, d = 0.005, $p = .059$) on respondents' responses. There was an effect of First Language suggesting that L1 Welsh speakers were more conservative than L2 speakers ($\beta = 0.32$, d = 0.23, $p = 0.019$). Again respondents who grew up in a bilingual home did not differ from respondents who learned English first ($\beta = 0.16$, d = 0.11, $p = 0.28$) or those who learned Welsh first ($\beta = –0.16$, d = 0.11, $p = 0.20$) Again we found significant differences between responses to canonical non-mutation and to non-canonical over-mutation ($\beta = 1.34$, d = 0.95, $p < .001$) as well as between canonically mutated and non-canonically unmutated sentences ($\beta = 1.98$, d = 1.4, $p < 0.001$). Overall, results suggested very high familiarity with non-canonical mutation (see graph in the appendix in the supplementary materials, available at https://osf.io/hqrz6/).

6.2 Trigger-specific results

The above analysis collapsed different mutation triggers together, but examination of the data suggested interesting differences between responses to different triggers. we therefore split the data out by trigger and analysed each of them in isolation. Since there is not enough data per trigger to conduct an analysis on the data using the linear model, we provide only a clustering analysis. We also used only Q1 data for the trigger-specific analyses, since they apply more directly to respondents' own use of mutation. In all of the graphs for the different triggers (**Figures 4–9**) we have numbered the respondents so that they can be tracked between triggers; we also use shape to indicate first-language status (the only significant demographic predictor of responses in our linear model).

6.2.1 Natural-gender singular nouns following the definite article

Singular feminine nouns are canonically mutated after a definite article, but masculine nouns are not. Non-normativity score results are shown in **Figure 4** using the same plot structure as for **Figure 2** in Section 6.1. The silhouette analysis identified seven clusters in the data (indicated

using colour in **Figure 4**), though an alternative two-cluster analysis (indicated using ellipses) was almost as highly ranked. The two cluster analysis is similar to the two clusters identified for the overall data (**Figure 3**), with a cluster of Conservative speakers in the bottom left and a diagonally oriented cluster of Variable speakers. The seven-cluster analysis breaks this down into smaller subsets, including a cluster of respondents in the top right of the Conservative quadrant, bridging the gap between the more conservative clusters in the bottom left (of which there are three) and the central variable cluster. Overall, this trigger could be characterised as a relatively strong or reliable trigger that is still used canonically by many speakers. This is suggested by the fact that four of the seven clusters, accounting for over 50% of respondents, are primarily located in the Conservative quadrant. Aside from these, the largest cluster is the variable cluster in the middle of graph. The Over-mutation cluster has only four members, and there are similarly few participants in the Under-mutation quadrant.

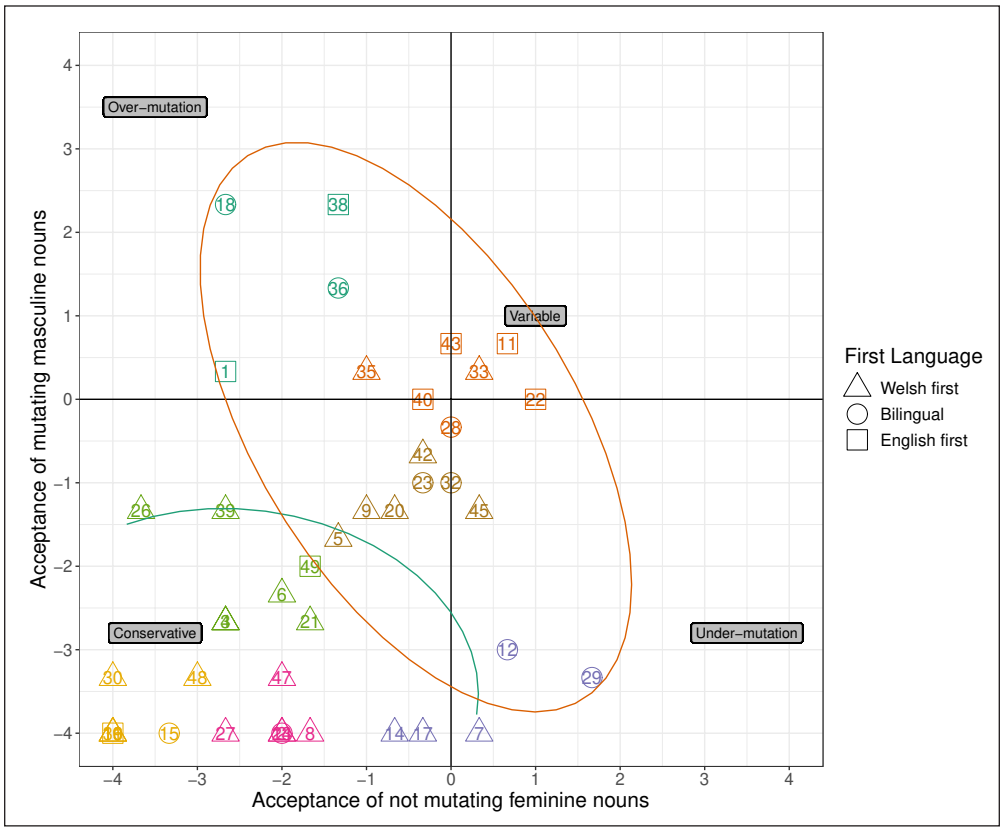

Figure 4: Preferences regarding mutation of definite natural-gender singular nouns. X-axis indicates respondents' acceptance of under-mutating feminine nouns. Y-axis indicates acceptance of non-canonically mutating masculine nouns. Each numbered point in the graph represents a single respondent, coloured according to which of the seven clusters they were identified as belonging to. Shapes indicate first language status. Ellipses indicate clusters from on an alternative two-cluster analysis (ranked similarly to the seven-cluster analysis.

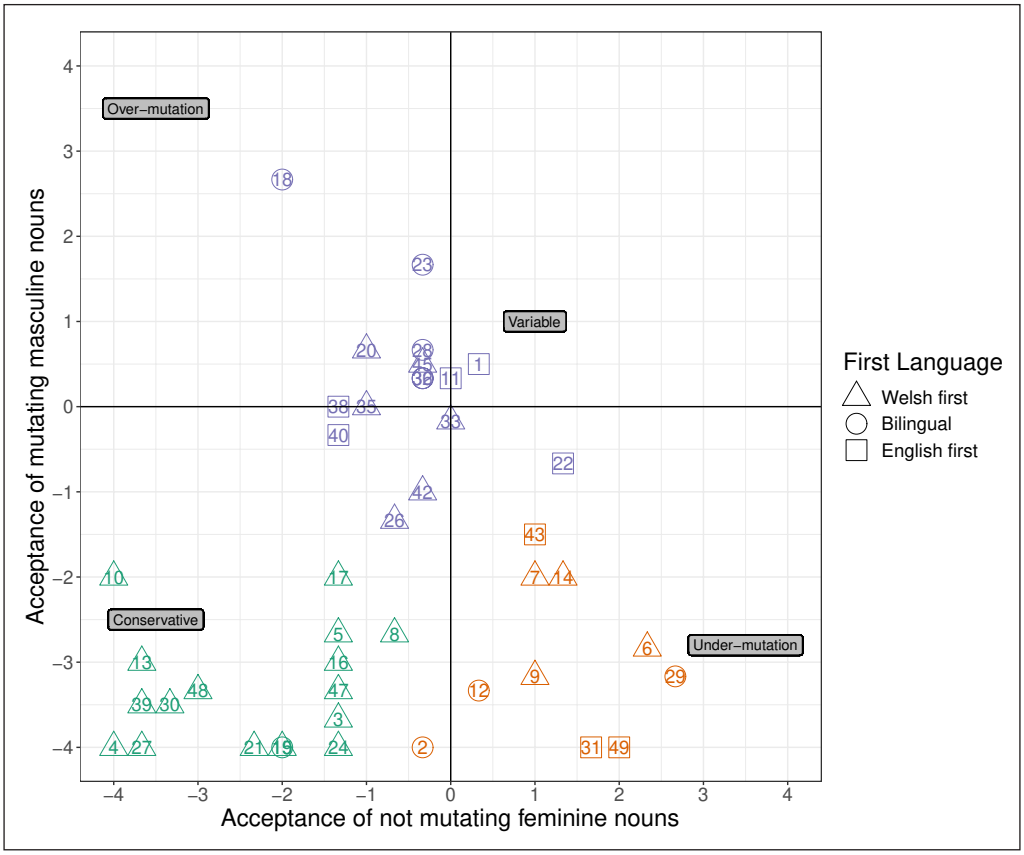

Figure 5: Preferences regarding mutation of suffixed definite singular nouns. X-axis indicates respondents' acceptance of under-mutating feminine nouns. Y-axis indicates acceptance of non-canonically mutating masculine nouns. Colours indicate clusters. Shapes indicate first language status.

6.2.2 Suffixed singular nouns following the definite article

This trigger is the same as the preceding trigger (natural-gender nouns), except that the grammatical gender is less transparent. In these sentences, the nouns are inanimate, though they are morphologically marked (with the masculine suffix *-yn* or the feminine suffix *-en*, both of which typically contribute diminutive or singulative meaning). The cluster analysis identified three main groups of respondents. The two largest are a Conservative group in the bottom left (comprising 39.5% of respondents) and a Variable group in the center (comprising 37% of respondents). The most notable difference from the natural-gender trigger is the presence of a rather more substantial cluster of under-mutating respondents in the bottom right quadrant. This is consistent with the expectation that attitudes towards mutating feminine nouns should vary according to how obvious the gender of the noun is, and that the gender of natural-gender nouns is likely to be more obvious than that of nouns marked by an arbitrarily gendered suffix.

Unlike for natural-gender nouns, there is also much more evidence of under-mutation than of over-mutation.

There are different possible interpretations of this result with regard to the dynamics of gender and mutation in Welsh. One is that the language might be shifting to a natural-gender system, in which nouns with inanimate referents are less likely to be perceived as feminine. It has previously been reported that such a change may indeed be underway for Welsh, at least for some speakers, with masculine pronouns being widely used to refer to all inanimate anaphors (B. M. Jones, 1993; M. C. Jones, 1998). This is also discussed by Thomas and Gathercole (2005), who did not find evidence for a shift towards a natural-gender system among speakers in northwestern Wales. An alternative interpretation of our data is that, to the extent that grammatical gender still operates in Welsh, there is uncertainty and variation as to which nouns are considered feminine, with animate nouns being the most stable. This second interpretation is not mutually exclusive with the first interpretation, and could be taken to constitute an early stage in the loss of grammatical gender. The difference in responses to natural-gender and inanimate nouns, however, suggests that mutation of feminine nouns (but not masculine nouns) following the definite article remains a feature of the Welsh mutation system for at least some speakers and that it varies to some extent independently of variation in the identification of gender.

6.2.3 Adjectives following indefinite natural-gender nouns

This trigger concerns mutation of adjectives after feminine nouns. This trigger does not rely on the presence of a definite article but affects adjectives directly following any feminine noun. In order to restrict the sentences to having a single mutation context, we only used unmutated indefinite nouns (all natural gender). The trend we saw for natural-gender and inanimate nouns, in line with work by Thomas and Gathercole (2005), was consistent with the expectation that mutation preferences might be influenced by gender transparency, with inanimate feminine nouns less likely to be mutated than animate ones. This trigger can be seen as less transparent still, as the gender of the adjective is dependent on that of another (in this case) unmutated, word. If this is right, we should expect respondents to be even less likely to prefer canonical patterns for gendered adjectives than for inanimate nouns.

This seems to be borne out in the data, as shown in **Figure 6**. The cluster analysis revealed five clusters of respondents. The most Conservative cluster is not only smaller than for inanimate nouns but also spread out horizontally in the bottom-left quadrant, with far fewer respondents in the bottom left corner than for either natural-gender or inanimate nouns. There is, however, a more substantial “Variable Conservative” cluster located in the top right of this quadrant (accounting for 28% of the data). There is a further Variable cluster located in the right side of the graph and a somewhat substantial cluster of under-mutating respondents in the bottom-right quadrant. There is no cluster of Over-mutating respondents. This is all consistent with this

being a trigger for which the trend is to mutate less. Taken together with the other triggers, the data are consistent with the existence of a positive relationship between trigger transparency and preference for mutation. It is important in this regard to bear in mind that, for the adjective trigger, we looked only at adjectives following *unmutated* indefinite nouns. It is possible that we would have found different results for adjectives following mutated definite nouns, for which the mutation of the noun might prime mutation of the adjective.

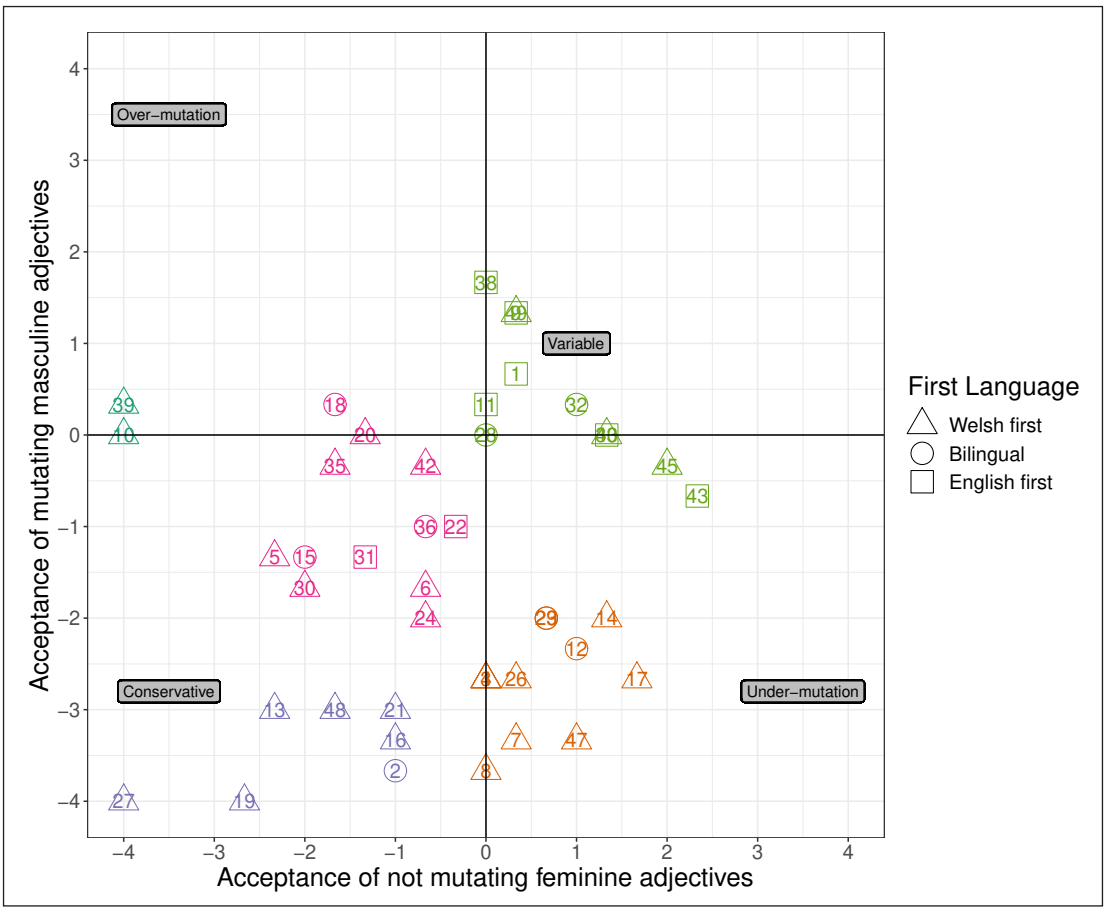

Figure 6: Preferences regarding mutation of adjectives following masculine unmutated and feminine mutated nouns. X-axis indicates respondents' acceptance of under-mutating feminine adjectives. Y-axis indicates acceptance of non-canonically mutating masculine adjectives. Colours indicate clusters. Shapes indicate first language status.

6.2.4 Direct objects of inflected verbs

Direct objects directly following inflected personal verbs are canonically mutated whether or not an overt subject intervenes. Many accounts treat overt and covert subject NPs as the same trigger, and some go further and treat any c-commanding maximal phrasal projection as a soft mutation

trigger (Borsley et al., 2007; Tallerman, 2006). If responses vary depending on the presence of an overt subject, this might shed interesting light on how the mutation rule is encoded.

Results for sentences without an overt subject can be seen in **Figure 7**. For this trigger, over-mutation is impossible as mutation is always canonically required, so the quadrants have different meanings from the quadrants in **Figures 4–6**. The Conservative quadrant is now in the bottom-right; the top-right quadrant indicates speakers who are Variable and who accept both mutation and non-mutation for this trigger. The top-left quadrant contains respondents who prefer under-mutation and the bottom-left quadrant contains respondents who gave low ratings both to sentences with mutation and to sentences without mutation. The axis labels also reflect the wording on the original Likert scales.

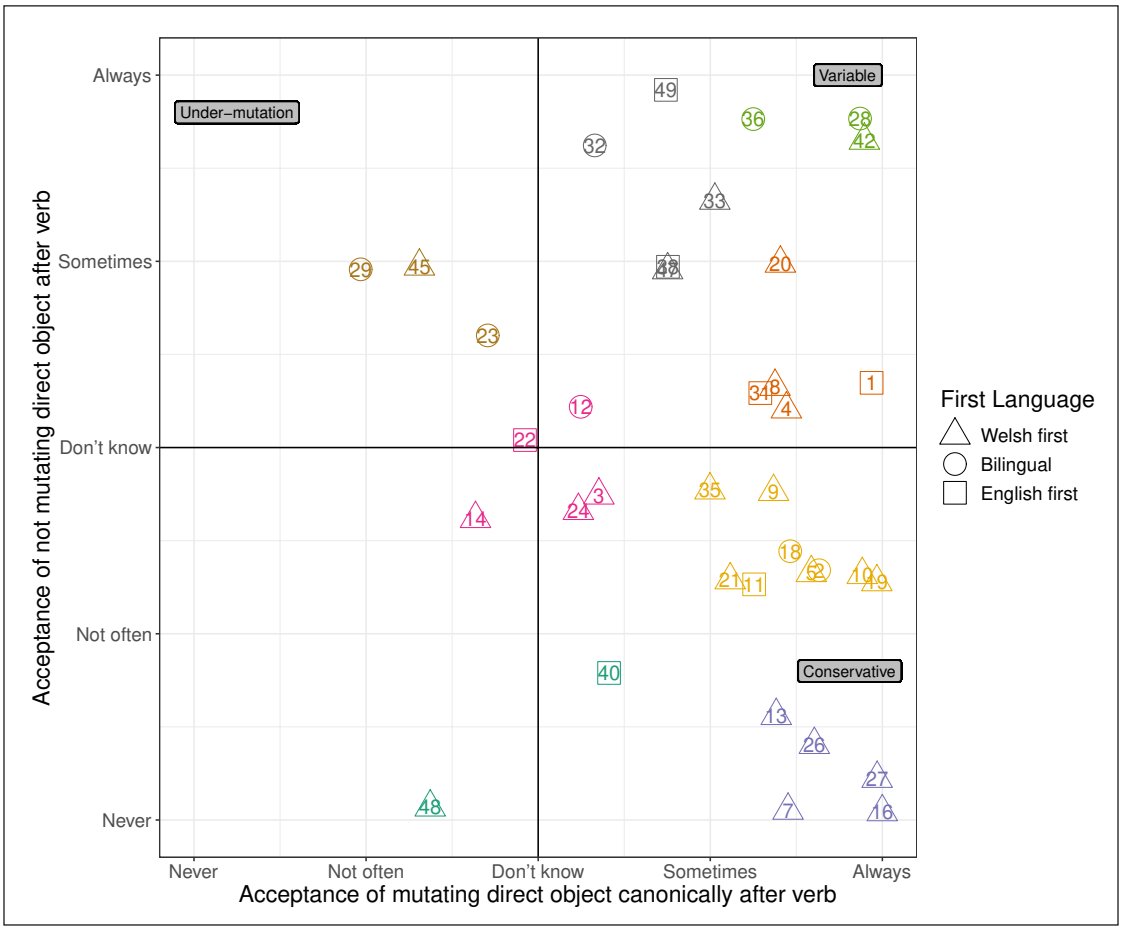

Figure 7: Preferences regarding mutation of direct objects following an inflected verb without an overt subject. Respondents in the bottom-right quadrant prefer conservative mutation. Respondents in the top-left quadrant prefer under-mutation. Respondents in the top-right quadrant are variable. Colours indicate clusters. Shapes indicate first language status.

The cluster analysis identified eight clusters in total, but the most striking feature of the graph is that the vast majority of respondents are located in the right half of the graph, with only six respondents (14%) clearly located in the left half. In fact, there is only one small cluster (of four respondents) wholly located in that half, specifically in the right half of the Under-mutation quadrant. In other words, the vast majority of respondents can be identified as Conservative or Variable, suggesting that this trigger remains rather strong. There is a cluster of five respondents located in the bottom-right of the Conservative quadrant, suggesting particularly canonical preferences. This is suggestive. However, it is important to note that sentences without overt subjects are rather more common in formal and literary registers (where canonical mutation is to be expected) than in colloquial registers, with the consequence that speakers are somewhat less likely to encounter non-mutation with this trigger in everyday life. Prescriptive norms are also likely to play a more important role here, so differences here should be treated with caution.

Figure 8 shows the results for sentences with an overt subject. The cluster analysis identified seven clusters in this case. Overall, however, the pattern is very similar to the pattern in **Figure 7**. There are two main differences. The first is that, for sentences with overt subjects, the cluster of under-mutating respondents is located further to the left of the graph, suggesting a greater preference for not mutating. The second difference is that the Conservative respondents tend to be located closer to the midlines, with only one respondent in the bottom-right corner (as compared with the five-respondent cluster located in this region for the overt-subject trigger). In other words, respondents seem overall more accepting of non-mutation after overt subjects than after inflected personal verbs with no overt subject, suggesting that—at least for some speakers—these might not behave as the same trigger, contrary to many accounts. It should also be noted, however, that personal inflected verbs with overt subjects are considerably more common in colloquial registers than verbs without them. This variation between registers may well account for at least some of the differences observed.

One other cluster of six respondents is worth commenting on. This is located around the lower half of the vertical midline. These respondents have a preference for non-mutation in this context (as indicated by their low position on the y-axis) but seem also unsure about applying mutation (as indicated by their position in the middle of the x-axis). If the verb form should be considered the primary mutation trigger in this context, sentences with intervening overt subjects can be seen as involving less direct triggers, analogous with adjectives.

6.2.5 Adverbial noun phrases

Adverbial noun phrases were our final trigger. According to canonical mutation rules, the first element in a noun phrases (e.g., *mis nesaf* “next month”) is mutated when it is used adverbially (*fis nesaf*). Historically this trigger was sufficiently consistent that certain noun phrases which occur only rarely outside an adverbial context (e.g., *ddoe* “yesterday”) are unfamiliar to many

speakers in their unmutated form. However, the trigger is also particularly abstract, being essentially triggered by the role of the noun phrase, not by any overt morpheme; furthermore, it is canonically considered to be optional if sentence initial (Williams, 1980), and in all positions is more observed in written registers than colloquial ones. We therefore expected respondents to be rather tolerant of nonmutation for this trigger, and perhaps to be less accepting of mutation.

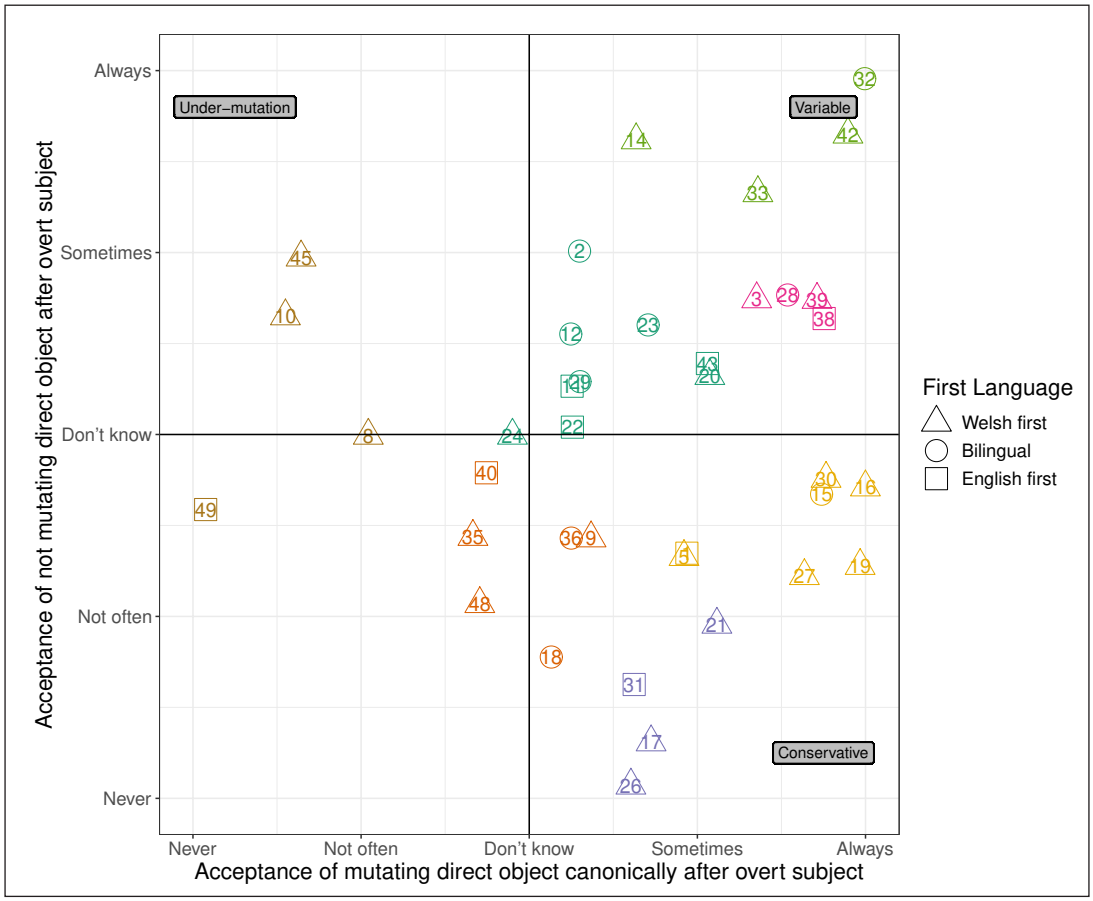

Figure 8: Preferences regarding mutation of direct objects following an inflected verb with an overt subject. Respondents in the bottom-right quadrant prefer conservative mutation. Respondents in the top-left quadrant prefer under-mutation. Respondents in the top-right quadrant are variable. Colours indicate clusters. Shapes indicate first language status.

Figure 9 shows the results for this trigger, and a strikingly different pattern is in evidence compared with the other triggers. The cluster analysis identified two main clusters, one mainly located in the Under-mutation quadrant and the other mainly located in the Variable quadrant. There are few participants located in the Conservative quadrant (and all are identified by the cluster analysis as belonging to the Variable cluster). None are located in the most conservative

corner of the quadrant. The pattern, in other words, is for variability or non-mutation, consistent with the view that this trigger is a relatively "weak" one, for which the vast majority of respondents are comfortable with non-mutation. To a small extent this may by influenced by the position of the noun phrase in question. Respondents gave their highest approval (with a mean Non-normativity score of 0.898) to sentences with an unmutated sentence-initial noun phrase, which is consistent with mutation in such sentences being canonically optional. However, sentences with mutation received similar ratings wherever they occurred (1.39 for sentence-initial and 1.61 for non-sentence-initial noun phrases; unfortunately data are too sparse for statistical comparison), and it is clear from the pattern of data in **Figure 9** that the overwhelming picture was simply of high acceptance of *not* mutating.

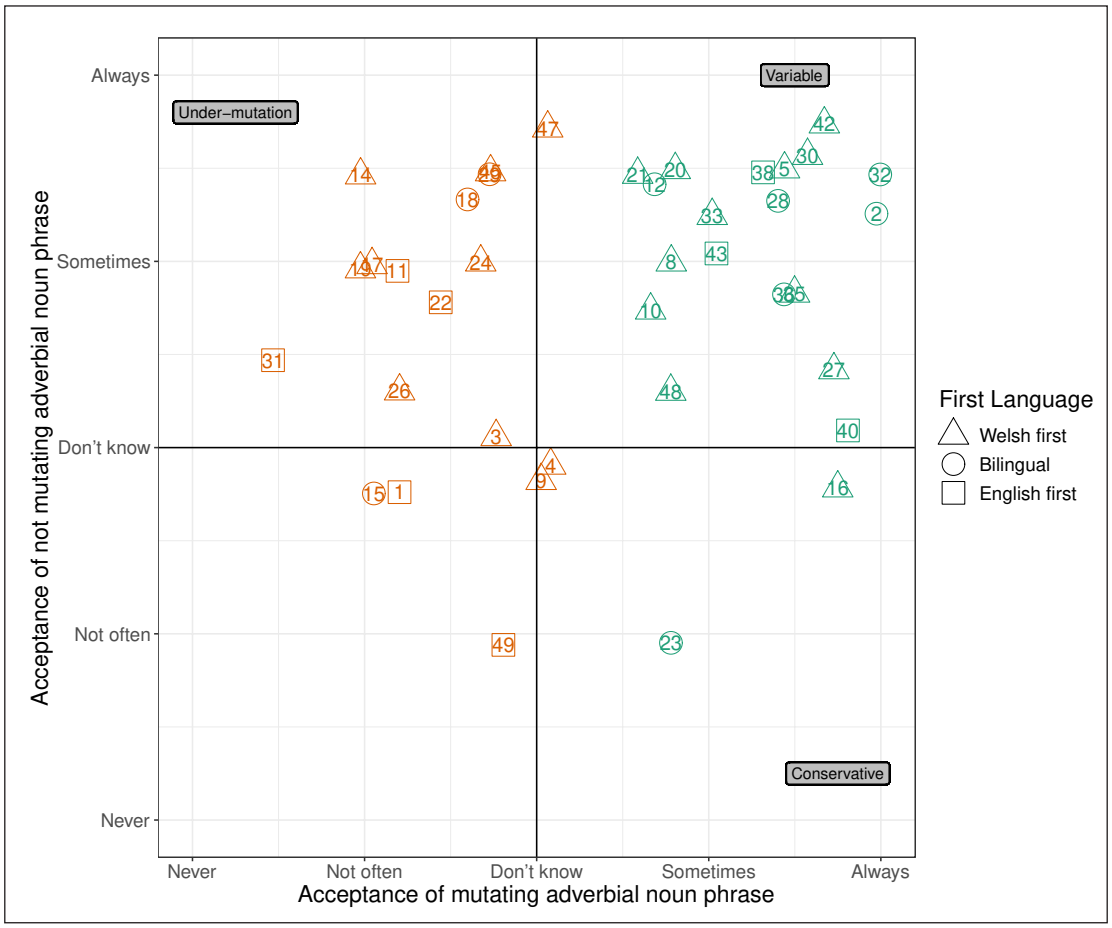

Figure 9: Preferences regarding mutation of adverbial noun phrases. Respondents in the bottom-right quadrant prefer conservative mutation. Respondents in the top-left quadrant prefer under-mutation. Respondents in the top-right quadrant are variable. Colors indicate clusters. Shapes indicate first language status.

7 Discussion

We used an online survey to investigate Welsh speakers' perceptual acceptability judgements of mutation and non-mutation across a variety of morphosyntactic triggers. We compared judgements of (a) sentences with canonical mutation patterns, (b) sentences with under-mutation (i.e., no-mutation where mutation is expected according to canonical rules), and (c) sentences with over-mutation (i.e., mutation where not canonically expected).

We had three main predictions. The first prediction was that demographic factors, primarily L1 status, but also age and region, would explain variation in the data. In fact, only L1 status had a significant effect. This was also consistent with our cluster analysis. The overall pattern, taking all triggers together, was of two similarly sized clusters of respondents: one Conservative cluster who tend to prefer canonical mutation patterns and a Variable cluster accepting of a range of different patterns. The Conservative cluster was dominated by respondents who reported having grown up speaking Welsh as a first language, and who made up 80% of this cluster (compared with 30% of the Variable cluster). Only one respondent (4%) in the Conservative cluster spoke English as a first language and only two gave both English and Welsh as first languages, compared with seven (39%) and five (28%) respondents respectively in the Variable cluster. This suggests a rather clear (and not surprising) picture: Change and variation in the Welsh mutation system seems to be primarily driven by the use of English in Wales. This general conclusion aligns with prior studies in bilingual minority-language contexts that bilingual and L2 speakers of the minority language use accept much more variability in the use of a complex language system like mutation than more traditional L1 speakers do (Dorian, 1994; Nance, 2015). This literature also led us to expect to find a similar effect for age, but we did not find one in our sample.

If we consider the mutation preferences of the Variable cluster of respondents, we find a preference for under-mutation as compared with over-mutation. The minority-language change literature suggests that complex grammatical systems can simplify either by losing features or by extending them to new contexts (i.a. Maher, 1991; Silva-Corvalán, 1986). The preference of Variable respondents for under-mutation compared with over-mutation suggests that in Welsh, omission is more common than extension. This is particularly relevant to the gender system, for which soft mutation is the only useful cue. Hammond (2016) suggested that inanimate feminine nouns are becoming masculine, and our results are consistent with this.

Our second prediction concerned direct objects directly following overt and covert subjects of inflected personal verbs. This is typically treated as the same trigger, and what we have called the maximal-projection account groups the two contexts together with other contexts to treat all c-commanding maximal-projections as triggers of soft mutation (Borsley et al., 2007; Tallerman, 2006). Contrary to this, we predicted that the presence of an overt subject NP might make a difference, at least for some speakers, implying that they treat these as distinct triggers.

Our cluster analysis did indeed suggest that these two contexts are not treated identically by our respondents. This result must be treated with a certain degree of caution, however. Covert subjects in Welsh are more a feature of written or formal than of colloquial registers, so it may be that register is playing an important role here. On the one hand, if these triggers are acquired as distinct, then that may be partly a consequence of register-based differences in their distribution. On the other hand, it may be (as an anonymous reviewer suggested) that some respondents' judgements partly reflect judgements of register-appropriateness rather than the appropriateness or naturalness of the mutation per se. While we sought to avoid this kind of effect with our instructions, it remains a real possibility, and a concern for future work.

Our third prediction concerned trigger transparency. We predicted that the remaining triggers would be ordered as follows with respect to their reliability, that is, their likelihood of eliciting more conservative responses:

natural-gender nouns > inanimate nouns > adjectives following nouns

The results of our clustering analysis were indeed consistent with this. The least conservative responses were also associated with the least transparent trigger, adverbial noun phrases, although this trigger exhibits variability even in formal registers (being optional sentence-initially). In future work researchers might like to focus on comparisons of frequency, transparency, and functional load (cf. Boon, 2014) in predicting trigger-dependent patterns of variation and change.

There are three limitations to this study that do present options for future work on this topic. First, our results on individual triggers must necessarily be considered somewhat preliminary and exploratory. While we consider our clustering analysis to have been enlightening, the amount of data for each trigger was relatively small, limiting our statistical options. We consider our results to provide a suggestive and intriguing basis for future work targeting particular triggers and variables with more power.

A second limitation concerns the stimuli. While the speaker, a Welsh-speaking linguist, was careful to aim for a natural style, as judged by the authors (and many sentences were recorded more than once so as to improve naturalness), recruitment difficulties prevented us from piloting the sentences with independent samples. This should be a consideration for future work, as should the use of a diversity of voices with different demographic characteristics.

The third limitation has more to do with the respondents themselves. As an anonymous reviewer pointed out, respondents may vary in their linguistic awareness and, consequently, in their ability to judge the appropriateness of mutation as distinct from other aspects of the sentence. Again, this is a place where future work—by asking questions in a wider variety of ways—may make progress. It is worth adding that we also provided our own respondents with prompts for open-field responses. In many cases their responses are somewhat enlightening, and we discuss them in what follows.

Several responses are relevant to the importance of L1 status. A number of respondents, for instance, indicated that they thought education and exposure to "correct" mutation patterns in the home likely play an important role. Many suggested that exposure to canonical mutation patterns at home is crucial as mutation is not formally taught until secondary school, by which time it is harder to learn. One respondent stated, for instance:

> I learnt the standard patterns in secondary school (rather than in primary school), and by then my mutation patterns in speech had been 'established'.[8]

Other respondents described how the mutation patterns they learnt at home from their parents often differed from the canonical patterns taught in school. One might expect this to result in identifiable regional variation too. As discussed, we did not find evidence for this, which might mean that region is in fact irrelevant to variation in mutation (except to the extent that it intersects with numbers of first-language speakers). However, it might also be that our regional variable (north vs. south) was simply too coarse-grained, or that our sample size was too small to identify existing patterns.

A related point concerns quality of input. Multiple respondents suggested that they learnt how to mutate through being corrected by their parents, and that their children are learning to mutate from them. One respondent commented:

> I think the environment is a factor and that she'll learn partly from how I mutate, but I certainly don't always mutate correctly!

The variability in production implied here is very much consistent with our data: respondents who patterned as "Variable" (who tended to give high ratings both to sentences with mutation and to sentences without mutation) were one of the largest categories, and the vast majority of our respondents exhibited some variation in their responses. Nonetheless, it is worth noting that, overall, there were as many respondents in the Conservative cluster (who tended to rate canonical mutation higher than non-canonical mutation) as in the Variable cluster.

As noted above, register may play an important role in explaining some of our results. We have discussed this in the context of certain triggers in particular, but it may also be that some of our respondents associate mutation in general with more formal registers, at least to an extent. We did not manipulate register, so our survey data cannot provide much beyond hints of this. However, there is evidence from our questionnaire, in which we asked respondents if they use mutation differently in different social contexts. Many described a distinction between "lazy"/ "messy"/ "untidy"/ "relaxed" Welsh and "formal"/ "correct"/ "standard" Welsh, and suggested that in the former contexts, not paying attention to mutation is common. Crucially,

[8] Responses have been translated from Welsh by the second author.

most respondents described *not paying attention* to mutation rather than not mutating at all. It was also clear that respondents perceived a distinction between canonical, prescriptive mutation rules and acceptability in colloquial registers. Many, for instance, described themselves as using mutation "incorrectly" in informal contexts, while acknowledging (as one respondent put it) that "People use words they know are officially incorrect but are more natural to use."

Other respondents offered support for the idea that non-canonical mutation is acceptable in informal contexts, for example:

> […] nor do they worry if they don't speak correctly.

> When I speak with the family, there's less pressure to get it right.

A methodological point should be raised in this context. The sentences used in our survey contained many natural gender nouns. Since the list of natural-gender nouns is limited, and not all nouns begin with a sound that is subject to mutation, it was necessary for the sake of variety to include a number of nouns that are well known, but might not be especially common in everyday speech. These include nouns referring to such individuals as princesses, kings, witches, monks, and maids. It is possible that such nouns, bringing with them associations of fairytales, influenced respondents' interpretation of register. In future work it would be important to control for such effects, and interesting to manipulate register, such as with a matched-guise task (cf. e.g., Campbell-Kibler, 2008).

Related to register is accommodation, whereby speakers adjust their language in response to their interlocutor's linguistic behaviour or even in response to *expected* linguistic behaviour (Wade, 2020). In the questionnaire, multiple respondents mentioned the influence of speakers' first- and second-language status on mutation patterns, and one respondent commented that how they mutate "depends on who is in the conversation. I adapt to go along with the person. So I would say something less correct/formal to help a learner to understand, for example." In this context it is important to note that we asked respondents to imagine that they heard each sentence in a conversation; responses might have differed had we framed the task differently.

One respondent commented explicitly on the topic of over-mutation, in response to a question on whether they think mutation is used differently by different groups of people:

> I have second-language friends in their 30–40s who over-mutate quite often, that is mutate where there's no need, or even remutate a word that's already been mutated.

This comment recalls a finding of Thomas and Gathercole (2005), who found adults double-mutating certain nouns, and confirms that there is metalinguistic awareness among some Welsh speakers of over-application of mutation as a phenomenon. Our overall pattern of results,

however, suggested that respondents were generally more tolerant of under-mutation that over-mutation.

In general, we consider the present work to constitute an exciting starting point for new research on the status of mutation in Welsh, while providing a perceptual and judgement-based perspective to complement earlier work focused on production. The study shows that the variation in the Welsh mutation system that has been detected in production studies of Welsh children, can also be seen in the expectations of adult Welsh-English bilinguals. Balanced bilingual and L2 Welsh speakers gave lenient acceptability judgements to non-canonical mutation, especially (though not exclusively) under-mutation. L1 Welsh speakers gave more conservative judgements. These judgements pattern with expected directions of change for minority languages in bilingual contexts (e.g. Dorian, 1994).

The method—an online survey in which respondents simply gave judgements to recorded sentences—allowed acceptability-judgement data to be collected relatively easily without requiring in-person interaction with a consultant or requiring respondents to record themselves. The completely transparent approach (in which respondents were made fully aware of the purpose of the research) was also successful—we see no evidence of serious demand characteristics. In spite of the study's simplicity, it produced suggestive results that we hope will form the basis of further investigation, ideally with larger samples. As discussed above, future work might focus in a more controlled way on the role of variables such as register and trigger transparency. This methodology could also be easily adapted to involve different voices (e.g. to examine dialect differences), to manipulate respondents' expectations of the speakers, and to investigate the role of different kinds of triggers, or triggers of different frequencies. We hope that our study and its results might be of interest not only to researchers interested in Welsh in particular, but also to researchers with an interest in grammatical change in bilingual situations of prolonged contact between a minority and prestige language more broadly.

Abbreviations

1 = 1st person, 3 = 3rd person, AM = aspirate mutation, F = feminine, INF = infinitive, M = masculine, PST = past, PL = plural, PRS = present, PROG = progressive, SG = singular, SM = soft mutation, SBJ = subject

Data availability and supplementary materials

The stimuli, raw survey results, demographic data, and a summary of results for Q2 are available via the Open Science Foundation, along with a key describing the variables in each of these datasets. DOI: 10.17605/OSF.IO/HQRZ6

Supplementary file 1: Appendix. Overall results for Q2

Supplementary file 2: Demographic data

Supplementary file 3: Results file

Supplementary file 4: Stimuli list

Supplementary file 5: Key for reading supplementary files 2–4

Acknowledgements

The authors wish to thank Dr. Ruaridh Purse and Ryan Budnick for their thoughts and input on how best to analyse the data for this project. We also wish to thank Dr. Rolf Noyer for his feedback on an earlier version of this paper, and the two anonymous reviewers for their insightful and detailed feedback on this work.

Competing interests

The authors have no competing interests to declare.

References

Audacity Team. 2014. *Audacity (R): Free audio editor and recorder.*

Awbery, Gwenllian M. 1986. Moves towards a simpler, binary mutation system in Welsh. In Andersen, Henning (ed.), *Sandhi phenomena in the languages of Europe*, 161–166. Berlin: Mouton de Gruyter.

Ball, Martin John. (ed.). 1988. *The use of Welsh: A contribution to sociolinguistics*. Clevedon: Multilingual Matters.

Ball, Martin John. 1992. Diglossia and its effect on the Welsh mutation system. *Études Celtiques* 29(1). 453–453.

Ball, Martin John & Müller, Nicole. 2002. *Mutation in Welsh*. London/New York: Routledge. DOI: https://doi.org/10.4324/9780203192764

Bellin, Wynford. 1988. The development of pronunciation. In Ball, Martin J. (ed.), *The use of Welsh: A contribution to sociolinguistics,* 213–228. Clevedon: Multilingual Matters.

Boersma, P., & Weenink, D. 2019. *Praat: Doing phonetics by computer*.

Boon, Erin Diane. 2014. *Heritage Welsh: a study of heritage language as the outcome of minority language acquisition and bilingualism.* (Unpublished doctoral dissertation). Harvard University, Boston.

Borsley, Robert D. & Tallerman, Maggie & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486227

Bybee, Joan L. & Hopper, Paul J. 2001. *Frequency and the emergence of linguistic structure* (Vol. 45). Amsterdam: John Benjamins Publishing. DOI: https://doi.org/10.1075/tsl.45

Campbell-Kibler, Kathryn. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, 637–659. DOI: https://doi.org/10.1017/S0047404508080974

Central Statistics Office. 2020. *Census of population 2016 – profile 10 education, skills and the Irish language.* Ireland. Retrieved from https://www.cso.ie/en/releasesandpublications/ep/p-cp10esil/p10esil/ilg/

Dale, Rick & Lupyan, Gary 2012. Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems* 15(03n04). 1150017. DOI: https://doi.org/10.1142/S0219525911500172

Dorian, Nancy C. 1973. Grammatical change in a dying dialect. *Language* 49. 413–438. DOI: https://doi.org/10.2307/412461

Dorian, Nancy C. 1977. The problem of the semi-speaker in language death. *Linguistics* 15(191). 23–32. DOI: https://doi.org/10.1515/ling.1977.15.191.23

Dorian, Nancy C. 1978. The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language* 54(3). 590–609. DOI: https://doi.org/10.1353/lan.1978.0024

Dorian, Nancy C. 1980. Language shift in community and individual: The phenomenon of the laggard semi-speaker. *International Journal of the Sociology of Language* 1980(25). 85–94. DOI: https://doi.org/10.1515/ijsl.1980.25.85

Dorian, Nancy C. 1981. *Language death: The life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press. DOI: https://doi.org/10.9783/9781512815580

Dorian, Nancy C. 1994. Purism vs. compromise in language revitalization and language revival. *Language in Society* 23(4). 479–494. DOI: https://doi.org/10.1017/S0047404500018169

Dorian, Nancy C. 2010. *Investigating variation: The effects of social organization and social setting*. Oxford: Oxford University Press.

Dresher, B. Elan. 2000. Analogical levelling of vowel length in West Germanic. In Lahiri, Aditi (ed.), *Analogy, levelling, markedness*, 47–70. De Gruyter Mouton.

Dressler, Wolfgang 1972. On the phonology of language death. *Chicago Linguistic Society* 8. 448–457.

Dryer, Matthew S. 2013. Order of subject, object and verb. In Dryer, Matthew S. & Haspelmath, Martin (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from https://wals.info/chapter/81

Gafter, Roey J. & Horesh, Uri. 2020. Two languages, one variable? pharyngeal realizations among arabic–hebrew bilinguals. *Journal of Sociolinguistics* 24. 369–387. DOI: https://doi.org/10.1111/josl.12373

Galantucci, Bruno & Roberts, Gareth & Langstein, Benjamin. 2018. Content deafness: When coherent talk just doesn't matter. *Language & Communication* 61. 29–34. DOI: https://doi.org/10.1016/j.langcom.2018.01.001

Gathercole, Virginia C. Mueller & Thomas, Enlli M. 2005. Minority language survival: Input factors influencing the acquisition of Welsh. In *Proceedings of the 4th international symposium on bilingualism*, 852–874.

Gathercole, Virginia C. Mueller & Thomas, Enlli M. & Laporte, Nadine I. 2001. The acquisition of grammatical gender in Welsh. *Journal of Celtic Language Learning* 6. 53–87.

Hammond, Michael. 2016. Predicting the gender of Welsh nouns. *Corpus Linguistics and Linguistic Theory* 12(2). 221–261. DOI: https://doi.org/10.1515/cllt-2015-0001

Hymes, Dell. 1971. *Pidginization and creolization of languages*. Cambridge: University Press.

Iriberri, Nagore & Uriarte, Jose-Ramon. 2012. Minority language and the stability of bilingual equilibria. *Rationality and Society* 24(4). 442–462. DOI: https://doi.org/10.1177/1043463112453556

Jones, Bob Morris. 1993. *Ar lafar ac ar bapur: Cyflwyniad i'r berthynas rhwng yr iaith lafar a'r iaith ysgrifennedig*. Aberystwyth: Y Ganolfan Astudiaethau Addysg.

Jones, Mari C. 1998. *Language obsolescence and revitalization: Linguistic change in two sociolinguistically contrasting Welsh communities*. Oxford University Press.

Jones, Morris & Thomas, Alan R. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.

Kassambara, Alboukadel & Mundt, Fabian. 2017. *Package 'factoextra'*.

Koplenig, Alexander & Meyer, Peter & Wolfer, Sascha & Mueller-Spitzer, Carolin. 2017. The statistical trade-off between word order and word structure–large-scale evidence for the principle of least effort. *PloS one* 12(3). DOI: https://doi.org/10.1371/journal.pone.0173614

Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7(2). 171–182. DOI: https://doi.org/10.1017/S0047404500005510

Likert, Rensis. 1932. A technique for the measurement of attitudes. *Archives of psychology* 22. 55.

Maher, J. 1991. A crosslinguistic study of language contact and language attrition. *First Language Attrition* 67. 84. DOI: https://doi.org/10.1017/CBO9780511620720.005

Matras, Yaron. 2020. *Language contact* (Second ed.). Cambridge: Cambridge University Press.

Morris, Jonathan. 2013. *Sociolinguistic variation and regional minority language bilingualism: An investigation of Welsh-English bilinguals in North Wales* (Unpublished doctoral dissertation). The University of Manchester, Manchester, UK.

Morris, Jonathan. 2021. Social influences on phonological transfer: /r/ variation in the repertoire of welsh-english bilinguals. *Languages* 6(97). DOI: https://doi.org/10.3390/languages6020097

Moseley, Christopher. (ed.). 2010. *Atlas of the world's languages in danger* (3rd ed.). Paris: UNESCO Publishing. Retrieved from http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/

Nagy, Naomi. 2015. A sociolinguistic view of null subjects and vot in Toronto heritage languages. *Lingua* 164. 309–327. DOI: https://doi.org/10.1016/j.lingua.2014.04.012

Nance, Claire. 2015. 'New' Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalisation. *Language in Society* 44(4). 553–579. DOI: https://doi.org/10.1017/S0047404515000408

Oftedal, Magne. 1962. A morphemic evaluation of the Celtic initial mutations. *Lochlann* 2. 93–102.

O'Rourke, Bernadette & Pujolar, Joan. 2015. New speakers and processes of new speakerness across time and space. *Applied Linguistics Review* 6(2). 145–150. DOI: https://doi.org/10.1515/applirev-2015-0007

Prys, Myfryr. 2016. *Style in the vernacular and on the radio: Code-switching and mutation as stylistic and social markers in Welsh* (Doctoral dissertation, Bangor University). Retrieved from https://research.bangor.ac.uk/portal/files/20578879/null

R Core Team. 2017. R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Roberts, Gareth & Fedzechkina, Maryia. 2018. Social biases modulate the loss of redundant forms in the cultural evolution of language. *Cognition* 171. 194–201. DOI: https://doi.org/10.1016/j.cognition.2017.11.005

Shahapure, Ketan Rajshekhar & Nicholas, Charles. 2020. Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, 747–748. DOI: https://doi.org/10.1109/DSAA49011.2020.00096

Silva-Corvalán, C. 1986. Bilingualism and language change: The extension of estar in los angeles spanish. *Language* 62(3). 587–608. DOI: https://doi.org/10.1353/lan.1986.0023

StatsWales. 2020. *Pupils by local authority, region and Welsh medium type.* Retrieved from https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Schools-Census/Pupil-Level-Annual-School-Census/Welsh-Language/pupils-by-localauthorityregion-welshmediumtype

Steinley, Douglas. 2006. K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology* 59(1). 1–34. DOI: https://doi.org/10.1348/000711005X48266

Tallerman, Maggie. 2006. The syntax of Welsh "direct object mutation" revisited. *Lingua* 116(11). 1750–1776. DOI: https://doi.org/10.1016/j.lingua.2004.10.004

Thomas, Enlli M. 2001. *Aspects of gender mutation in Welsh* (Unpublished doctoral dissertation). University of Wales, Bangor.

Thomas, Enlli M. & Gathercole, Virginia C. Mueller. 2005. Minority language survival: Obsolescence or survival for Welsh in the face of English dominance. In *Isb4: Proceedings of the 4th international symposium on bilingualism*, 2233–57.

Thomas, Enlli M. & Gathercole, Virginia C. Mueller. 2007. Children's productive command of grammatical gender and mutation in Welsh: An alternative to rule-based learning. *First Language* 27(3). 251–278. DOI: https://doi.org/10.1177/0142723707077056

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

Tzanne, Angeliki. 2000. *Talking at cross-purposes: The dynamics of miscommunication*, Vol. 62. Amsterdam: John Benjamins Publishing. DOI: https://doi.org/10.1075/pbns.62

Uriarte, Jose-Ramon & Sperlich, Stefan. 2021. A behavioural model of minority language shift: Theory and empirical evidence. *PLoS ONE* 16(6). DOI: https://doi.org/10.1371/journal.pone.0252453

Vitello, Sylvia & Rodd, Jennifer M. 2015. Resolving semantic ambiguities in sentences: Cognitive processes and brain mechanisms. *Language and Linguistics Compass* 9(10). 391–405. DOI: https://doi.org/10.1111/lnc3.12160

Wade, Lacey. 2020. *The linguistic and the social intertwined: Linguistic convergence toward southern speech* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

Weinreich, Uriel & Labov, William & Herzog, Marvin. 1968. Empirical foundations for a theory of language change. In Lehmann, W. P. & Malkiel, Y. (eds.), *Directions for historical linguistics*, 95–195. Austin: University of Texas Press.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. Retrieved from https://ggplot2.tidyverse.org. DOI: https://doi.org/10.1007/978-3-319-24277-4_9

Williams, Stephen Joseph. 1980. *A Welsh grammar*. University of Wales Press.

Willis, David. 2010. Old and Middle Welsh. In Ball, Martin John & Müller, Nicole (eds.), *The Celtic languages*, Second edition, 117–160. London/New York: Routledge.

Wolfram, Walt. 2004. Language death and dying. *The Handbook of Language Variation and Change*, 764–787. DOI: https://doi.org/10.1002/9780470756591.ch29

Wray, Alison & Grace, George W. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117. 543–578. DOI: https://doi.org/10.1016/j.lingua.2005.05.005

Zehr, Jeremy & Schwarz, F. 2018. *PennController for Internet Based Experiments (IBEX)*.