

This is a repository copy of *Viewing dialect change through acceptability judgments: A case study in Shetland dialect*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/205083/>

Version: Published Version

---

**Article:**

Jamieson, E [orcid.org/0000-0001-8679-3360](https://orcid.org/0000-0001-8679-3360) (2020) *Viewing dialect change through acceptability judgments: A case study in Shetland dialect*. *Glossa: a journal of general linguistics*. ISSN 2397-1835

<https://doi.org/10.5334/gjgl.979>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

## RESEARCH

# Viewing dialect change through acceptability judgments: A case study in Shetland dialect

E Jamieson

University of Edinburgh, Dugald Stewart Building, Edinburgh, GB  
[e.jamieson@ed.ac.uk](mailto:e.jamieson@ed.ac.uk)

Acceptability judgments are the standard methodology for investigating syntactic variation. While acceptability judgments have been shown to be reliable in cases of assumed stable variation, there has been little discussion of how syntactic *change* plays out in judgment tasks. This is despite evidence from sociolinguistics that at the end of a change, speakers' behaviour in production is "unpredictable". How does this "unpredictability" play out in judgment tasks, where speakers are asked to perform metalinguistic assessments on their use of a changing variable? In this paper, I present the results of acceptability judgment tasks focusing on a particle, *-n*, available in some biased questions in the Shetland dialect of Scots. This variety has been claimed to be rapidly obsolescing (e.g. Smith & Durham 2011). Combining quantitative analysis of acceptability judgments with qualitative comments made by the speakers, I argue young speakers exhibit *perceptual hyperdialectalism* in their judgments: extending acceptability of the variable to contexts where older speakers don't accept it; giving higher ratings than expected given their qualitative comments when the variable is being lost, and generally rating examples which could be perceived as "dialectal" higher. I argue these patterns arise from Preston's (2013) definition of linguistic insecurity: younger speakers are aware their grammar diverges from a more "traditional" grammar, but the traditional grammar is an important identity marker. These speakers therefore attempt to demonstrate knowledge of an older grammar – but their knowledge is not accurate. These patterns highlight the importance of combining quantitative and qualitative analysis in dialect syntax.

**Keywords:** dialect syntax; acceptability judgments; linguistic insecurity; hyperdialectalism; Scots

## 1 Introduction

Acceptability judgments are the standard methodology for investigating morphosyntax (e.g. Schütze 1996). In recent years, acceptability judgment methodologies have been combined with best practice from sociolinguistics in order to investigate the morphosyntax of non-standard varieties (e.g. Cornips & Poletto 2005; Barbiers & Bennis 2007; Zanuttini et al. 2018; Smith et al. 2019) and have been shown to be reliable ways to capture variation within communities (Thoms 2014; Zanuttini et al. 2018).

While these methodologies have been successful at investigating cases of stable variation, work on syntactic microvariation using judgments does not tend to incorporate discussion of ongoing *change* in these varieties. From sociolinguistic research, we know that the final stages of a change can be rapid and less orderly than the earlier stages of a change, with unexpected production patterns (e.g. Dorian 1977; Campbell & Muntzel 1989; Hill 1989; Schilling-Estes & Wolfram 1999; Labov 2001). Furthermore, Dorian (1977) presents evidence that "semi-speakers" of East Sutherland Gaelic, those at the tail end of the loss of the variety, are "unreliable" when asked to perform elicitation tasks. However, Adger (2017) argues that the "unreliability" of elicitations in East Sutherland

Gaelic is due to the breakdown and subsequent reorganisation of the syntax in the ongoing change process, with semi-speakers using sparse data to reconstruct a system that fits the underlying “syntactic ecology” of the language.

We can therefore ask: how do speakers in the late stages of a syntactic change behave when completing judgment tasks? Is there unpredictability? Do we see an organised restructuring process? Understanding how speakers behave in these contexts can shed light on the nature of the grammatical system at the end of a change, and how this is affected by the potential social pressures of a metalinguistic task.

In this paper, I focus on one community where a number of forms in the dialect are said to be undergoing obsolescence: the Shetland dialect of Scots. Linguistic change and the loss of features in this community is argued to be happening rapidly – even from one generation to the next (Smith & Durham 2011, 2012; Durham 2013). Shetland dialect is thus an ideal testbed for seeing how a community of speakers who are undergoing potential language loss, and exhibiting the unpredictable patterns seen at the end of a change, behave when presented with acceptability judgment tasks which ask them to make explicit judgments on dialect variants – in this case, *-n*, a particle used in certain types of biased questions. I will also consider whether unpredictability extends to judgments on stable acceptable and unacceptable features.

## 2 Dialect syntax and acceptability judgments

Since at least Chomsky (1966), the standard data collection procedure for morphosyntactic data has been acceptability judgments. While these began as introspective binary judgments, considerable criticism of this methodology (Quirk & Greenbaum 1970; Labov 1972; Newmeyer 1983; Rizzi 1990; Nagata 1992; Bard, Robertson & Sorace 1996) has led to the development of alternative, gradient methodologies for syntactic acceptability, such as Likert scales or magnitude estimation tasks (e.g. Bard, Robertson & Sorace 1996; Schütze 1996; Sprouse 2007; Bader & Häussler 2010). It is these gradient methodologies that have been incorporated with best practice from sociolinguistics (e.g. Labov 1966; 1972) in more recent work that has aimed to tap into vernacular variation in morphosyntactic phenomena (e.g. Cornips & Poletto 2005; Barbiers & Bennis 2007; Zanuttini et al. 2018; Smith et al. 2019). In English varieties alone, acceptability judgment data has led to insights into a wide variety of syntactic phenomena, such as imperatives (e.g. Henry 1995; Weir 2013), preposition dropping (e.g. Myler 2013; Biggs 2015), datives (e.g. Horn 2013; Wood & Zanuttini 2018), possessives (Tsoulas & Woods 2019) and more.

Gradient judgment tasks have been shown to be reliable indicators of acceptability in cases where there is no expected variation. Here, inter-speaker judgment reliability is high (Sprouse, Schütze & Almeida 2013; Sprouse & Almeida 2017; Langsford et al. 2018, though see Linzen & Oseki 2018 on Hebrew and Japanese judgments). Gradient judgment tasks have also been shown to be reliable in cases of stable variation in a community, where speakers using both a dialect variant and a standard variant (Thoms 2014; Zanuttini et al. 2018). However, when there are multiple grammars of a non-standard dialect within an individual community, Henry (1995) observes that speakers with one grammar may form incorrect hypotheses about speakers of the other. She notes (1995: 56) that it is “particularly important to get grammaticality judgments from speakers who themselves use the dialect concerned” in order to avoid this.

How does this work in situations of linguistic *change*? There has been little discussion of the reliability of acceptability judgment methodologies in situations of linguistic change.<sup>1</sup>

<sup>1</sup> Heycock, Sorace, Hansen, Wilson & Vikner’s (2012) work on the loss of V-to-T in Faroese is an exception; Maling & Sigurjónsdóttir (2002) also use judgments to look the development of the impersonal passive in Icelandic.

On one hand, languages are always changing (e.g. Labov 1994, 2001). After an initial innovation, changes progress constantly and systematically (Weinrich, Labov & Herzog 1968). We would therefore expect that as long as we are able to target speakers of a particular grammar within the change, judgments should be reliable for that grammar. However, at the end stages of linguistic change, speakers' behaviour with regard to the feature being lost is argued to be somewhat unpredictable. This can be seen with individual linguistic features in an otherwise "healthy" linguistic variety (e.g. Ingason 2010), but is amplified when linguistic change is happening to multiple features at once, leading to potential language loss (e.g. Dorian 1977; Christian, Wolfram & Dube 1988; Campbell & Muntzel 1989; Hill 1989; Labov 1994, 2001; Schilling-Estes & Wolfram 1999).

This unpredictability is reflected in the behaviour of speakers at the end of a change in other types of linguistic data gathering methodologies. In recorded production data, we see "dramatic increases in variability due to incongruent and idiosyncratic change" (Cook 1989: 235) across phonological (e.g. Wolfram & Schilling-Estes 1995; Schilling-Estes & Wolfram 1999) and morphosyntactic (e.g. Dorian 1973) variables.

Furthermore, we see unpredictability in the results of more structured data gathering tasks with speakers undergoing change in Dorian's (1977) use of spoken elicitation tasks with "semi-speakers" of East Sutherland Gaelic. "Semi-speakers" were younger members of the East Sutherland community whose Gaelic was considered "imperfect", though they could still make themselves understood (Dorian 1977; 1994). Dorian presented 16 members of the East Sutherland community with 115 English sentences and asked them to translate the sentences into Gaelic. The semi-speakers produced analogical levelling across a range of paradigms, including verbal agreement and irregular noun formation, in an inconsistent manner. She concludes that reduced use of features applies "essentially to the individual speaker, not the community" (Dorian 1977: 30), with different speakers producing different patterns seemingly randomly. However, Adger (2017) argues that the levelling in East Sutherland Gaelic can be understood in terms of an abstract restructuring of the syntactic system, reconstructed by the semi-speakers on the basis of sparse input. Rather than acting at random, speakers are creating a new system within the confines of underlying constraints.

Is this flux that we see for speakers at the end of a change in both production and elicitation data reflected in acceptability judgment data, and if so, how? Do we see idiosyncratic behaviour, or evidence of a restructured system? In the next section I introduce the community that we will use to test potentially obsolescing speakers' behaviour with regards to the feature *-n* when presented with acceptability judgment tasks.

### 3 Shetland dialect

The Shetland Islands are the most northerly region of Scotland, located over 200 miles north of Aberdeen (see Figure 1). The islands have a relatively stable population of just over 23,000 (Shetland Islands Council 2017).

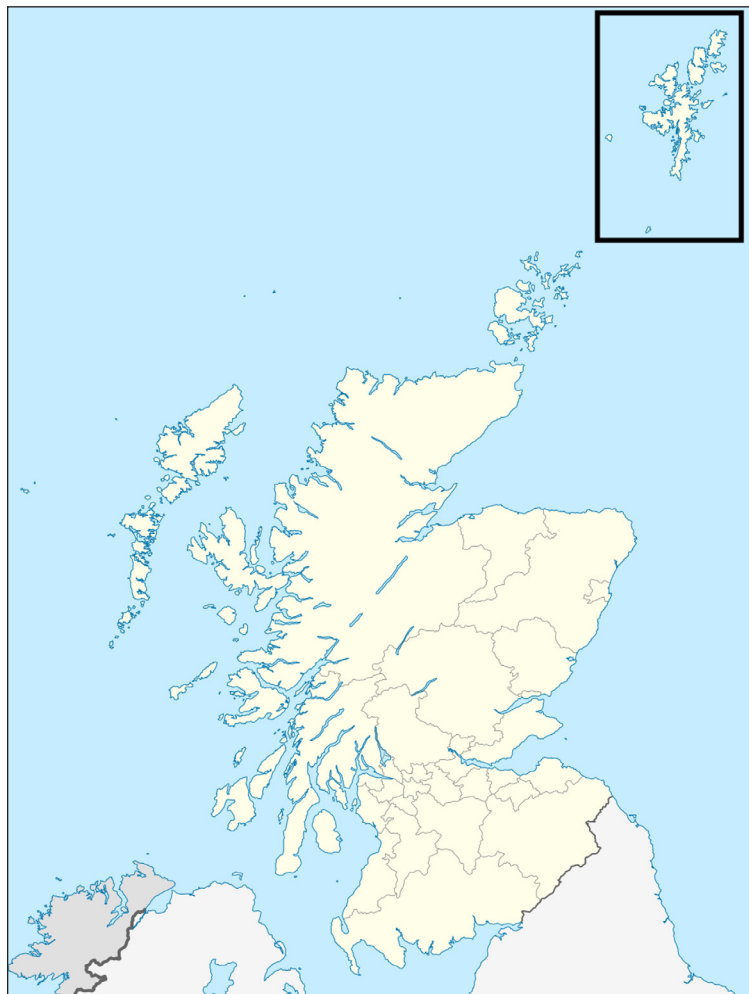
The dialect of Scots spoken in the Shetland Islands<sup>2</sup> in the present day is known for being especially distinctive among Scots varieties, retaining older Scots features presumably due in part to its remoteness (Melchers 2004) as well as features which are argued to have resulted from the period of language contact between Scots and Norn (Millar 2008), the west Norse language spoken in the isles until the mid-eighteenth century (e.g. Barnes 1998).

<sup>2</sup> Debate surrounds the nomenclature used for this variety (e.g. Tait 2007). I will refer to it as *Shetland dialect* throughout.

However, a number of recent works have noted that Shetland dialect appears to be obsolescing (van Leyden 2004; Melchers & Sundkvist 2010; Smith & Durham 2011; 2012; Sundkvist 2011; Durham 2013; 2014; 2017; Jamieson 2015). In particular, Smith & Durham's (2011; 2012) research demonstrates that young speakers of the dialect have reduced rates of both Shetland-specific and Scotland-wide lexical, phonological and syntactic variants as compared to both older and middle-aged speakers. Their research also shows some seemingly unpredictable behaviour among the younger speakers – for example, with demonstrative distal *yon*. *Yon* participates in a distal deictic system in Shetland dialect alongside *this* and *that*, and can either be used in a determiner-type role (1) or as a pronominal (2) (Smith & Durham 2011: 209).

- (1) Still actually have yon BMW.
- (2) I was just like “What’s yon?”

Smith & Durham investigated speakers’ usage rates of demonstratives in both determiner and pronominal positions, in singular distal contexts – contexts where there is variation between *yon* and *that*.<sup>3</sup> There were generally low rates of *yon* across the three age groups



**Figure 1:** Map of Scotland highlighting the Shetland Islands (CC-BY-SA-3.0).

<sup>3</sup> The literature disagrees as to whether *that* or *yon* indicates the further distance in Shetland dialect (Robertson & Graham 1952, Melchers 1997). There appears to be no clear system, at least in the present-day, though both are more distal than *this*.



that Smith & Durham studied but, surprisingly, there was a higher percentage of *yon* as compared to *that* in the younger group than the middle or older groups. Looking at this between speakers, however, only four of the eight younger speakers used any *yon* at all, with four speakers using *that* 100% of the time. Furthermore, two of the eight were especially prolific *yon* users, with around 36% *yon* usage – compared to < 20% *yon* across the older and middle groups – thus bumping up the overall younger speakers’ mean usage rates. The division in younger speakers’ usage rates did not correlate with any expected social factors, such as gender, social network, or attitude towards the islands, leading Smith & Durham to conclude that the variation seen is the kind of “personal pattern variation” found by Dorian (1994) in East Sutherland Gaelic, and that the dialect “may be facing rapid dialect attrition” (Smith & Durham 2011: 220).

Shetland dialect is therefore well placed as a community in which to investigate how speakers behave in acceptability judgment tasks when presented with a potentially obsolescing feature. In the next section, I introduce the acceptability judgment tasks that speakers completed in this research.

## 4 Methodology

### 4.1 The changing variable

The example sentences used in this research were designed to test the acceptability of a particle, *-n*, which can be used in a small subset of biased interrogative types in Shetland dialect – questions where the speaker already has a belief about what the answer to the question should be. The *-n* particle [ən] can be attached to any auxiliary verb (Robertson & Graham 1952: 10). There are no phonological changes to the root of the auxiliary when *-n* is suffixed, suggesting that it is not simply a reduction of an extant negative marker such as English *-n’t* or Scots negative marker *-na*. However, the particle is only distinguishable from a reduced form of *-n’t* in a limited number of phonological contexts (detailed in Table 1), such as (3). Cases like (4), where the example is *is’n*, are potentially confounded with standard English *isn’t* as they only require dropping a glottal stop in production. However, speakers of all ages judged examples like (5), with *is’n* in a regular declarative clause, to be unacceptable (Jamieson 2018), indicating that they do consider the *-n* particle different from a phonological reduction of standard English *-n’t*.

- (3) *Shetland dialect* (Robertson & Graham 1952: 10)

Can’n we no aa come in?  
 can.N we NEG all come in  
 “Can’t we all come in?”

- (4) *Shetland dialect* (Tait 1973: 13)

Tammy is’n yun Jeannie o Maanwil’s lass at’s gotten mairied dis week?  
 Tammy is.N that Jeannie of Maanwil’s daughter that’s got married this week  
 “Tammy, isn’t that Jeannie of Maanwil’s daughter who has got married this week?”

- (5) \*He is’n coming.

As can be seen in Table 1, with *can*, *do* and *will*, there is a clear difference between the form with *-n* and the Scots and English negative forms. With the rest of the verb forms (exemplified by *would* and *is*), however, there are syllabification differences in the *-na* cases, but only minor phonological changes (e.g. a vowel change in *would*, plus dropping the glottal stop) between the *-n* form and the standard English *-n’t* form. Therefore, where relevant, I discuss differences between *clearly local* forms of the *-n* marker (*can*, *do* and *will*) and those that could potentially be *confounded* with standard English variants (such as *would*, *is*, *did*, *should* etc).

**Table 1:** Sample list of verb forms and how they combine with *-n*, *-na* and *-n't* in Shetland dialect.

Verb	<i>-n</i> form		<i>-na</i> form		<i>-n't</i> form	
can	can'n	[kɪ.nən]	canna	[ka.nə]	can't	[ka.nʔ]
do	do'n	[du.ən]	dunna	[dʌ.nə]	don't	[dɔ.nʔ]
will	will'n	[wɪl.ən]	winna	[wɪ.nə]	won't	[wɔ.nʔ]
would	would'n	[wɪl.dən]	widna	[wɪd.nə]	wouldn't	[wu.dənʔ]
is	is'n	[ɪ.zən]	isna	[ɪz.nə]	isn't	[ɪ.zənʔ]

*-n* is undergoing some obsolescence in Shetland dialect, with older speakers accepting it in biased interrogatives (e.g. 4) to a greater extent than younger speakers. However, it is still acceptable for younger speakers in tag questions (6), exclamatives (7) and polar rhetorical questions (8) (Jamieson 2018).<sup>4</sup>

(6) You can come, can'n you?

(7) Is'n it a lovely day!

(8) Have'n you just got a new watch?!

These conclusions were reached on the basis of acceptability judgment tasks combined with qualitative comments. We can use this data to investigate how speakers at different stages of a change in a community treat the variables undergoing change in judgment tasks, and compare this to stable acceptable and unacceptable features to determine if, and if so how, the unpredictability in production associated with the later stage of a change is reflected in patterns of acceptability judgments.

#### 4.2 Experimental design

20 speakers of Shetland dialect (10 aged 18–30 and 10 aged 55+) took part in this study. Participants were recruited through the friend-of-a-friend approach (Milroy 1980). All were born and brought up in Shetland and had spent no significant time living away from the islands (no more than 1 year for the 18–30 group, and 3 years for the 55+ group). All had at least one parent who also met these criteria, and almost all had two parents who met these criteria. These participants were thus likely to be dialect users representative of their age group, having acquired the dialect from their parents (Payne 1980) and by continuing to be embedded in dense, multiplex networks in the community in the present day (Milroy 1980).

Participants met with a speaker of the local variety who acted as the interviewer. Using local fieldworkers mitigates against the Observer's Paradox (Henry 1995; Labov 2001; Adger & Trousdale 2007); the importance of this is heightened in Shetland where speaking to someone from outside the community has been shown to dramatically affect speakers' levels of dialect use (Smith & Durham 2012).

<sup>4</sup> On the surface, biased interrogatives (4) and polar rhetorical questions (8) look the same. The difference lies in the contexts that license them and the expectation of answer. While biased interrogatives are used in contexts where some confirmation is required from the addressee, polar rhetorical questions are used in contexts where the answer is known to all conversation participants, and do not require an answer. These constructions have received separate analyses in the literature (e.g. Romero & Han 2004 on biased interrogatives; Biezma & Rawlins 2017 on rhetorical questions) and so are treated separately here, but see Jamieson (2018) for an argument that they are the same type of construction.

Participants gave Likert scale judgments on 400 example sentences using an adaption of the *interview method* (Barbiers & Bennis 2007; Thoms 2014; Smith et al. 2019). The interviews took between 1.5 and 2 hours in total, and were split into two sessions of 200 examples, with a break in the middle. A written questionnaire was constructed which included examples of all the relevant phenomena. Each example was embedded in a short context, so as to make the example as naturalistic as possible. 236 of the 400 examples presented *-n* in a variety of interrogative and non-interrogative contexts of varying acceptability: assumed acceptable (9), assumed unacceptable (10) and undergoing change (11). Recall that examples like (9) are more clearly distinguished from a reduction of standard English *-n't* than examples like (11) – these will be discussed separately where this could confound judgments.

- (9) You are pretty sure that you have seen that I have a driving license. We agree to help one of our friends move house, and we're going to rent a van to do it. You say:  
**You can drive, can'n you?**
- (10) One of our friends loves dogs, but is very allergic to them. Despite this, she has always wanted one. I tell you that I recently saw her out walking a dog and you say:  
**She hasna gone and gotten a dog, has'n she?**
- (11) Last week, I went to a new pizza restaurant; I told you it was really good. My friend is in town and is looking for restaurant recommendations, so I am asking you for advice. You say:  
**Did'n you like that pizza place?**

The remaining 164 examples were fillers. Fillers were designed to give a broad variety of grammatical standard examples (12), grammatical examples using dialect features (13) and ungrammatical examples, again with some including dialect features (14) and some not (15).

- (12) One of our friends is always going on holiday, never to the same place. She's just back from somewhere again, but you're not sure where, so you say to me:  
**Where did she go this time?**
- (13) We are discussing Andy Murray's latest final. We were watching it together, but I had to go out. I ask you what happened and you say:  
**He never won in the end.**
- (14) I want to try out the new restaurant down the road. I ask you if it's good, but you say:  
**I no am been yet.**
- (15) We are trying to decided what to see at the cinema tonight. You say:  
**What see would like you?**

Normally, the interviewer reads out the relevant examples. However, due to the possibility of intonation affecting the interpretation of interrogatives (Pierrehumbert & Hirschberg 1990; Banuazizi & Creswell 1999; Wochner et al. 2015; Hedberg, Sosa, & Görgülü 2017), the interviewer read out the preceding context before participants pressed the space bar on a laptop to hear the relevant example sentence, which was recorded in advance by a



21-year old male speaker of the dialect. This allowed intonation to be controlled for while still retaining as much of the conversational feel of the interview method as possible.

After they had heard the example, participants were then asked to rate the example on a scale from 1 to 5. Only the end points of the scale were labelled in order to give participants flexibility in the remainder of the scale (Coward, p.c. to Schütze 1996: 186). A rating of 1 was described as “unusual; sounds weird; no one says that”, whereas 5 was described as “totally natural; I say that; people around me say that”. These labels attempted to access both the participant’s judgment of their own usage and of the rest of their community; this conflation creates potential ambiguity, which will be central to the discussion below.

Participants were then able to discuss their ratings with the interviewer, comment on particular aspects of the example in question or offer explicit alternatives that they would be likely to use instead. The interviewer noted specific comments, and all discussion was also recorded. In total, 1150 additional comments were coded across all participants. Both quantitative and qualitative results will be discussed in the analysis, below.

## 5 Quantitative results

### 5.1 Analysing Likert scale data

As described in Section 4.2 above, participants were asked to judge each item in the research on a 5-point Likert scale. 1 was described as something that was “unusual; sounds weird; no one says that”, while 5 was described as being “totally natural; I say that; people around me say that”. No particular descriptors were attached to the middle of the scale.

Although this is a numerical scale, we cannot assume that the distances between the points of the scale are equivalent – e.g. is the difference in acceptability between 4 and 5 of the same magnitude as the difference in acceptability between 3 and 4? The question thus arises as to whether scales like these can be interpreted as continuous, or if they must be analysed as discrete, ordinal points. The decision about how to categorise the data affects the choice of statistical test used.

If the data is interpreted as continuous, a t-test can be used to establish if there are significant differences in the mean ratings between two groups. If, however, the data is interpreted as ordinal, a non-parametric test e.g. the Mann-Whitney-Wilcoxon rank sum test would be more appropriate. As opposed to comparing means, the Mann-Whitney-Wilcoxon test ranks the judgment scores given by participants and calculates whether there is a different distribution of the data given the median of the two populations.

A test for normality can be conducted in order to establish whether the data resembles a normal distribution, and thus can be interpreted as continuous. Here, I conducted Shapiro-Wilk tests for normality. For each of the data cases presented below, the test returned a significant result, indicating that the distribution of scores did not align with a normal distribution. This is an initial indication that non-parametric tests would be better suited to the data. However, there is considerable research suggesting that even when continuous data are non-normal, t-tests are most appropriate when the sample size is large enough (e.g. Zimmerman & Zumbo 1993; Stonehouse & Forrester 1998; Skovlund & Fenstad 2001).

However, the research cited did not specifically include the interpretation of Likert scales. de Winter and Dodou (2010) conducted a comparison of levels of Type I (*false positive*) and Type II (*false negative*) errors between t-test results and Mann-Whitney-Wilcoxon results carried out on 5-point Likert scale data with different distributions. The authors found that the statistical power of both parametric and non-parametric tests was equivalent across most distributions, with two crucial exceptions. When the data was heavily skewed (e.g. almost all ratings were 5), non-parametric tests held more statistical power; on the other hand, when the distribution was strongly bimodal (e.g. ratings were fairly equally split between 1 and 5), parametric tests held more power.

Looking at the descriptive statistics for positive polar questions (e.g. Figure 2), we can see that the data is heavily skewed, with almost all cases being rated 5. Throughout the rest of the data, results are either skewed in this way, or have somewhat more flat distributions, where both t-tests and Mann-Whitney-Wilcoxon tests had equal power in de Winter and Dodou's (2010) work.

As non-parametric tests were shown to hold more power for the heavily skewed data, and other data were shown to have non-normal distributions through the Shapiro-Wilk test, I opt for non-parametric tests here. This will take the form of both Mann-Whitney-Wilcoxon rank sum tests and ordinal logistic regression models.

I will first present the results for a variety of examples expected to have different relationships to change. Firstly, I will present the results for a set of stable acceptable examples – positive polar questions, which in Shetland dialect are taken to be the same as in English,<sup>5</sup> e.g. (16).

(16) Do you have a dog?

I will then present results for two different sets of presumed unacceptable examples. Firstly, a set of unacceptable examples that don't contain any ostensibly dialectal features: *wh*-questions with incorrect word orders e.g. (17).

(17) What do want you?

Secondly, a set of unacceptable examples that include dialect features, but that are ungrammatical in the dialect – specifically, negative imperatives with *do no*,<sup>6</sup> e.g. (18). Unlike English negative imperatives where both *don't* and *do not* can be used (Potsdam 1996; Rupp 2003), negative imperatives in Scots varieties like Shetland dialect do not permit both options. Only *dunna/dinnae* (depending on location) is available; *do no* is unacceptable (Weir 2013: 9).

(18) Dinnae smoke in here!

(19) \*Do no smoke in here!

In these three contexts, I present the results from Mann-Whitney-Wilcoxon rank sum tests in order to test whether there were differences by age group.

I will then present results for the changing *-n* feature. Firstly, with tag questions like (9), above, where we see acceptability. Secondly, for tag questions like (10), above, where *-n* has never been attested in the dialect. Finally, for biased questions like (11), above, which appear to be undergoing change. The examples with *-n* will be compared with judgments for the standard Scots constructions that could be expected to be used in all three contexts (20–22).

(20) You can drive, can you no?

(21) She hasna gone and gotten a dog, has she?

(22) Did you no like that pizza place?

<sup>5</sup> Jamieson (2015) shows that there are remnants of verb raising in questions in present-day Shetland dialect; however, *do*-supported questions like (15) were preferred by speakers in all age groups in their research.

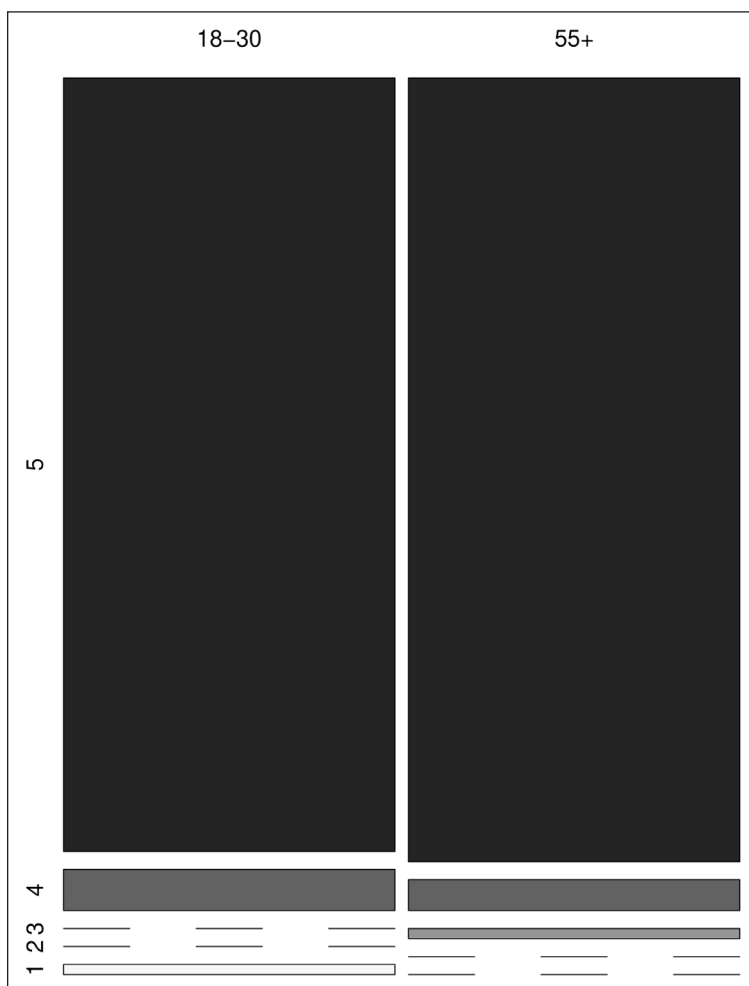
<sup>6</sup> The negative marker in Shetland dialect equivalent to standard English *not* is *no*, see e.g. Robertson & Graham (1952); Millar & Brown (1979).

To compare the  $-n$  variables with the relevant standard constructions, I use ordinal logistic regression models – a non-parametric variant of linear regression designed to deal with multiple, discrete response categories such as the Likert scale data presented here.

Regression models allow us to test to what extent a range of predictor variables (e.g. type of construction, age group, gender) affect a result (e.g. the judgment score given), and predict to what extent a change in a variable should affect the outcome for the predictor. For example, to what extent does being in the older age group affect the scores you give overall?

Here, I present results from cumulative link mixed models, run using the ordinal package (Christensen 2018) for R. Cumulative link mixed models are a form of ordinal logistic regression, which include mixed effects to allow potentially random effects (e.g. differences between individual participants, or responses to individual experimental stimuli) to be accounted for.

For each context, I ran models with random intercepts for participant and example, with construction type, age group and evidential context as predictor variables.<sup>7</sup> At no point was gender found to be a significant factor for any of the contexts, and so this was removed from the models.



**Figure 2:** Acceptability judgment results for positive polar questions, by age group.

<sup>7</sup> Evidential context has been shown to affect speakers’ choice of question form (Domaneschi, Romero, & Braun 2017). I leave evidence as a variable in the models presented here where relevant, but do not discuss the effects; see Jamieson (2018) for details.

## 5.2 Grammatical: Positive polar questions

Bader & Häussler's (2010) *ceiling effect* predicts that if constructions are acceptable to speakers, and are produced at more than 3% frequency in a corpus, they will be rated at ceiling level. To test whether this is true of participants undergoing change, I firstly present the results from positive polar questions, such as (23). Each participant judged eight positive polar questions, giving a total of 80 judgments per age group.

- (23) We're at the park. You decide to go and get an ice cream from the van. You say:  
**Do you want anything?**

As can be clearly seen in Figure 2, participants in both age groups behaved almost identically in rating positive polar question examples 5 in almost all cases. There is a very slight difference in the mean ratings (younger = 4.9, older = 4.938), but the median rating for both groups was 5; any difference between groups is not significant ( $W = 3240$ , n.s.). Bader & Häussler's ceiling effect seems to hold true for stable, acceptable examples even when speakers are undergoing change in their variety. Due to the fact that there may be variation in the numbers of different variants available in any given context that participants are judging, as well as potential additional effects of context or social pressures, I therefore take a combination of a median rating of 5 and a mean rating between 4 and 5 as the "ceiling" that indicates acceptability of a construction, and thus report both the median and mean throughout.

## 5.3 Ungrammatical: Wh-questions with incorrect word order

The second example we will look at is an example which is presumed unacceptable and does not include any notable "dialect" features: specifically, *wh*-questions with incorrect word orders, as in (24). Each participant judged 8 examples of questions like these, giving 80 judgments per age group.

- (24) You're going to the shop. I ask if you can get me something and you say:  
**What do want you?**

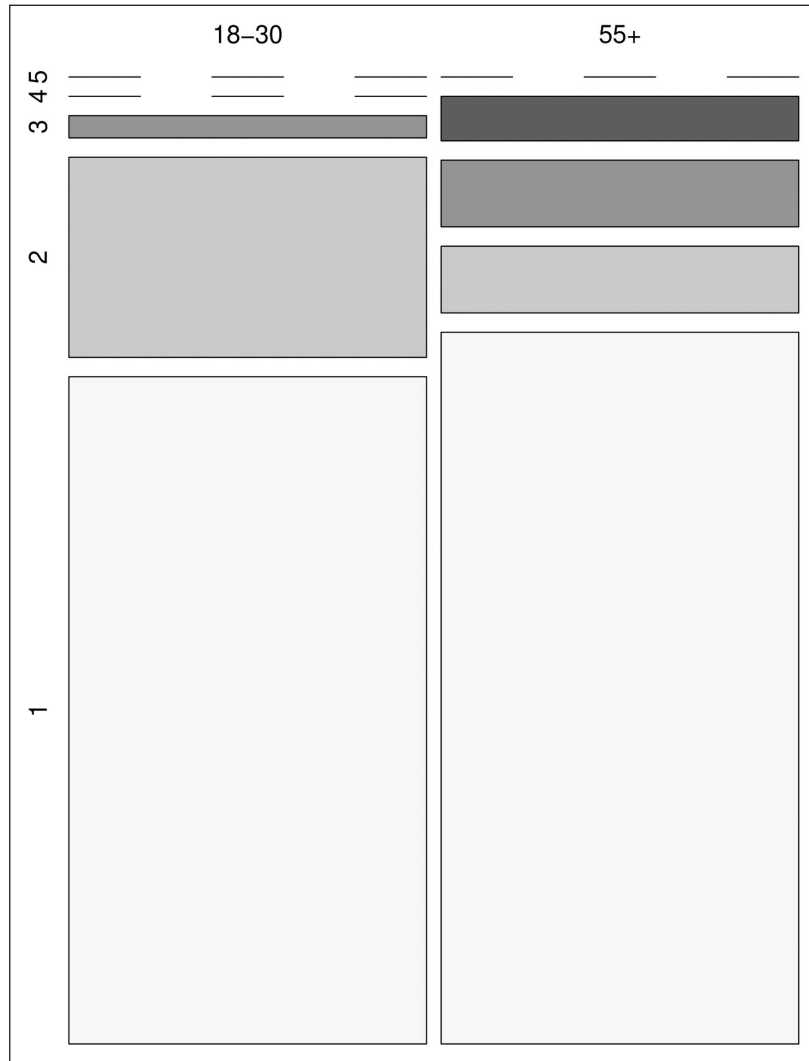
As can be seen in Figure 3, speakers in both age groups rate these constructions low: primarily giving ratings of 1. The median score for both age groups is 1; there is a slight difference in mean ratings (younger = 1.275, older = 1.375) but no statistically significant difference ( $W = 782$ , n.s.). Ungrammatical examples are given "floor" level ratings, as Bader & Häussler (2010) found for ungrammatical word orders in German. It therefore seems that judgments continue to be reliable even at the end stages of change. However, the picture is more complicated for ungrammatical examples that include dialect features.

## 5.4 Ungrammatical: Negative imperatives

Each participant judged four examples of negative imperatives which were constructed with *do no* in place of *dunna*, such as example (25), giving 40 judgments per age group. Recall that while *no* is the Scots form of *not*, Scots imperatives cannot be formed with *do no*, unlike English *do not*. Only *dunna* (equivalent of *don't*) is available.

- (25) I have just arrived at your house and I'm really hungry. There's some fresh scones sitting out, and I'm eyeing them up, but you say:  
**Do no touch them! They're for later.**

Speakers in both age groups did not rate these examples in the "grammatical" range, as set out above. In Figure 4, we see that there are very few 5 ratings in either age group.



**Figure 3:** Acceptability judgments for ungrammatical *wh*-questions, by age group.

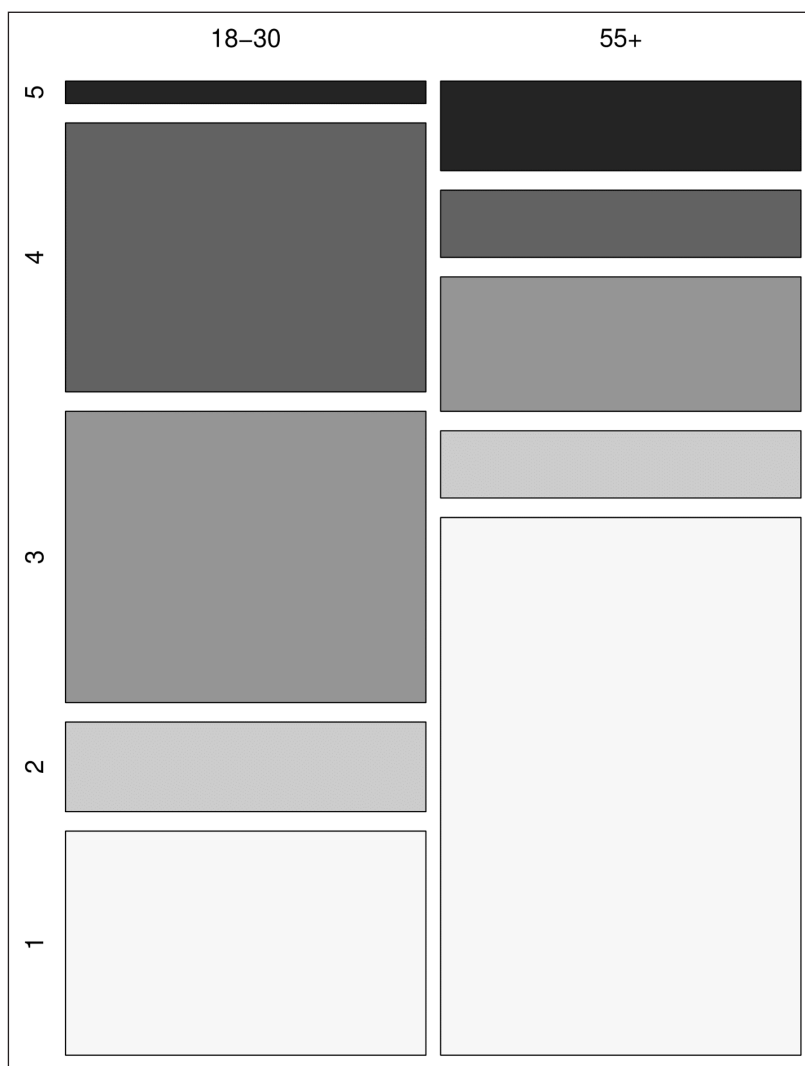
However, despite the fact that speakers did not consider these constructions acceptable, there were clear differences in the ratings between the age groups in both communities. In the 18–30 group, we see a higher number of 3 and 4 ratings, as compared to the higher numbers of 1 and 2 ratings seen in the 55+ group.

While neither of these groups rate the examples at ceiling, there are statistically significant differences with respect to how low the constructions are rated. Again, there were differences in mean ratings (younger = 2.75, older = 2.0) and in the median. The median for younger speakers was 3, while for older speakers it was 1. This difference was significant ( $W = 536, p = .008$ ). For younger speakers, there seems to be an unwillingness to always fully rule out an ungrammatical example: instead, sometimes a more mid-range rating is given.

This holds across participants quite generally: of the 10 1 ratings given here in the younger age group, 7 of them come from 2 participants, with the rest of the participants operating primarily across the mid-range. On the other hand, in the older age group, 6 of the 10 participants gave 1 ratings in at least 3 of the 4 test examples.

So, when it comes to dialect features, younger speakers seem to be doing something different. We can then turn to look at the changing variable in question, *-n*, and see how participants behave.





**Figure 4:** Acceptability judgment results for *do no* in negative imperatives, by age group.

### 5.5 *-n* in tag questions with positive anchors

In these cases, participants judged examples like (9), repeated here as (26), and general Scots examples like (27). There were 4 examples that were clearly local (40 judgments per age group), 12 examples that could be confounded with the standard (120 judgments per age group) and 16 general Scots examples (160 judgments per age group). These Ns are reflected in the width of the bars in Figure 5.

(26) You can drive, can'n you?

(27) You can drive, can you no?

As can be seen in Figure 5, in cases where the *-n* particle is potentially confounded with a standard variant, participants in both age groups give a high percentage of 5 ratings; the same is also true in the standard contexts. For speakers in the 55+ group, the clearly local contexts are also given a high percentage of 5 ratings; however, the younger group are more evenly split between ratings of 4 and ratings of 5.

I ran a cumulative link mixed model on the data, presented in Figure 6. The estimated coefficients in the left column of the main table indicate the inferred effect of the predictor

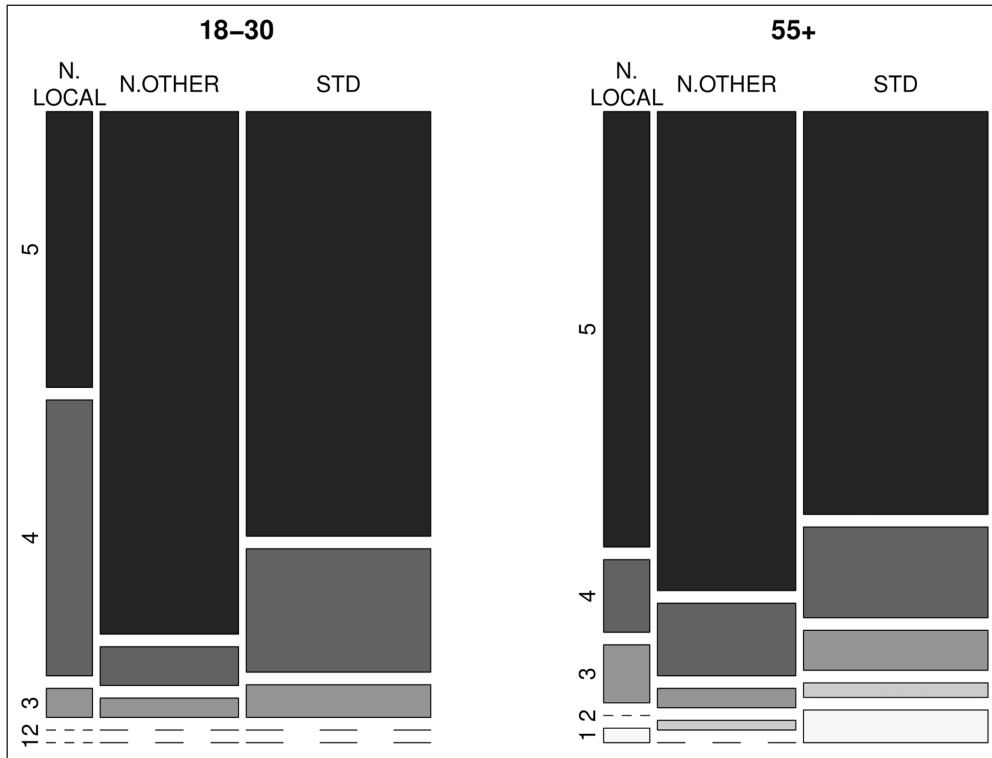


Figure 5: Acceptability judgment results for tag questions on positive anchors, by age group.

```

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 640 -430.31 888.63 1017(4072) 1.82e-04 1.5e+02

Random effects:
Groups Name Variance Std.Dev.
EXAMPLE.NO (Intercept) 0.2518 0.5018
PARTICIPANT (Intercept) 1.5381 1.2402
Number of groups: EXAMPLE.NO 32, PARTICIPANT 20

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
particleN.LOCAL      1.93176    0.72818   2.653 0.00798 **
particleN.OTHER      2.08940    0.53152   3.931 8.46e-05 ***
EVIDENCENEG          0.34602    0.36771   0.941 0.34670
AGEY                  0.56935    0.62527   0.911 0.36253
particleN.LOCAL:EVIDENCENEG -2.17685    0.82958  -2.624 0.00869 **
particleN.OTHER:EVIDENCENEG -1.36197    0.64222  -2.121 0.03395 *
particleN.LOCAL:AGEY   -1.78205    0.60057  -2.967 0.00300 **
particleN.OTHER:AGEY  -0.02663    0.50629  -0.053 0.95805
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
      Estimate Std. Error z value
112  -4.0350    0.5651  -7.140
213  -3.5524    0.5410  -6.567
314  -2.4331    0.5118  -4.754
415  -0.6304    0.4942  -1.276
    
```

Figure 6: Cumulative link mixed model results for tag questions on positive anchors.

variable on the participants' judgments on the Likert scale in log odds, with the z value (second rightmost column) indicating whether or not this is a statistically significant predictor. A z value over 2 generally indicates a significant predictor.

Figure 6 shows that there is a general effect of whether or not the example is a clearly local one (particleN.LOCAL), as well as an interaction between age group and the judgments given to clearly local instances of *-n* (particleN.LOCAL:AGEY). There is also a general effect of *-n* in confounded contexts (particleN.OTHER).<sup>8</sup> For older speakers, it appears that clearly local variants are as acceptable as the potentially confounded variants, while that is not quite the case for younger speakers. However, it is worth noting that both groups still rated *-n* highly, even in clearly local contexts.

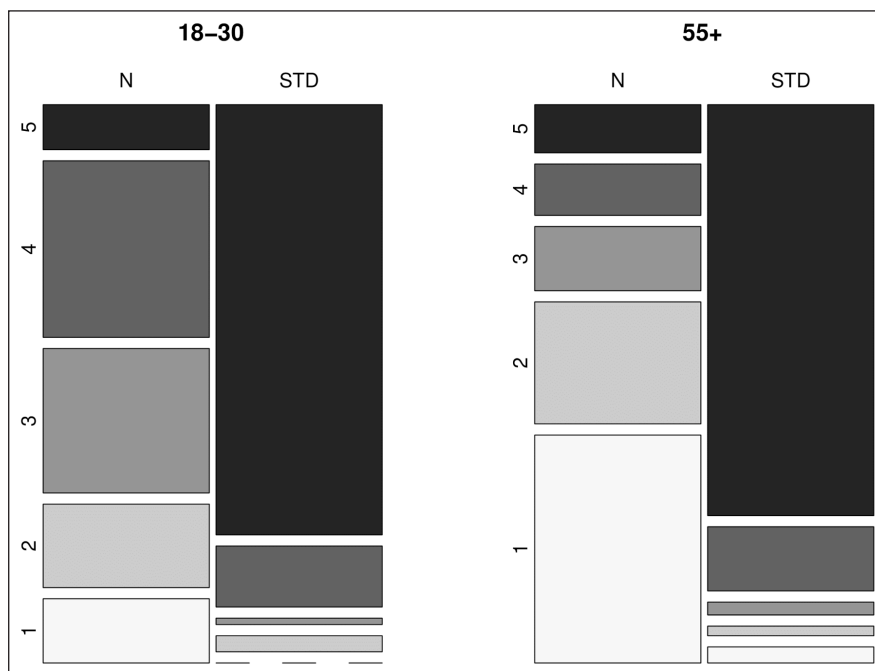
### 5.6 *-n* in tag questions with negative anchors

*-n* on tag questions with negative anchors, such as (28), was compared with standard constructions like (29). There were 16 examples of each type of construction (so 160 judgments per age group for each construction).

(28) She hasna gone and gotten a dog, has'n she?

(29) She hasna gone and gotten a dog, has she?

The picture is complicated in a similar way as it was for the clearly ungrammatical examples. From Figure 7, we can see that standard examples are rated highly by speakers in both age groups. For *-n*, on the other hand, younger speakers give more 3 and 4 ratings as compared to the older speakers' 1 and 2 ratings. Neither group is rating these examples in the "acceptable" range, but their behaviour regarding something being "unacceptable" seems to vary. This is backed up by the results of the cumulative link mixed model



**Figure 7:** Acceptability judgment results for tag questions on negative anchors, by age group.

<sup>8</sup> Potentially confounded examples such as *is'n* and *did'n* were in fact more acceptable than the standard Scots constructions like *is it no* or *did she no* in tag questions on positive anchors.

(Figure 8).<sup>9</sup> While there is an overall effect of the construction type (particleN), there is also an interaction between age and construction type (particleN:AGEY), indicating that younger speakers rated examples with *-n* in this unacceptable context significantly higher than older speakers did.

### 5.7 *-n* in matrix biased questions

With examples of *-n* in matrix biased questions, like (30), the picture is complicated. Speakers also judged standard examples like (31). There were 4 examples that were clearly local (40 judgments per age group), 12 examples that could be confounded with the standard (120 judgments per age group), and 16 general Scots examples (160 judgments per age group). These Ns are reflected in the width of the bars in Figure 9.

(30) Did'n you like that pizza place?

(31) Did you no like that pizza place?

The results can be seen in Figure 9. As expected, standard examples are rated highly. There are also very high ratings for *-n* in potentially confounded contexts. From the results of the cumulative link mixed model (Figure 10), we can see that there are overall effects of the different types of *-n* as compared to the standard construction (particleN.LOCAL and particleN.OTHER); there is also an interaction between the clearly local cases of *-n* and the age group.

Looking at the descriptive statistics, we see that for older speakers, the median for the N.LOCAL cases is 5, with over 50% of examples rated 5. The mean is mid-range (3.73), seemingly from a collection of 1–2 ratings (27.5% of the total) bringing down the average.

```

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 640 -633.51 1285.02 611(2356) 3.43e-05 1.8e+02

Random effects:
Groups Name Variance Std.Dev.
EXAMPLE.NO (Intercept) 0.3777 0.6146
PARTICIPANT (Intercept) 1.2440 1.1154
Number of groups: EXAMPLE.NO 32, PARTICIPANT 20

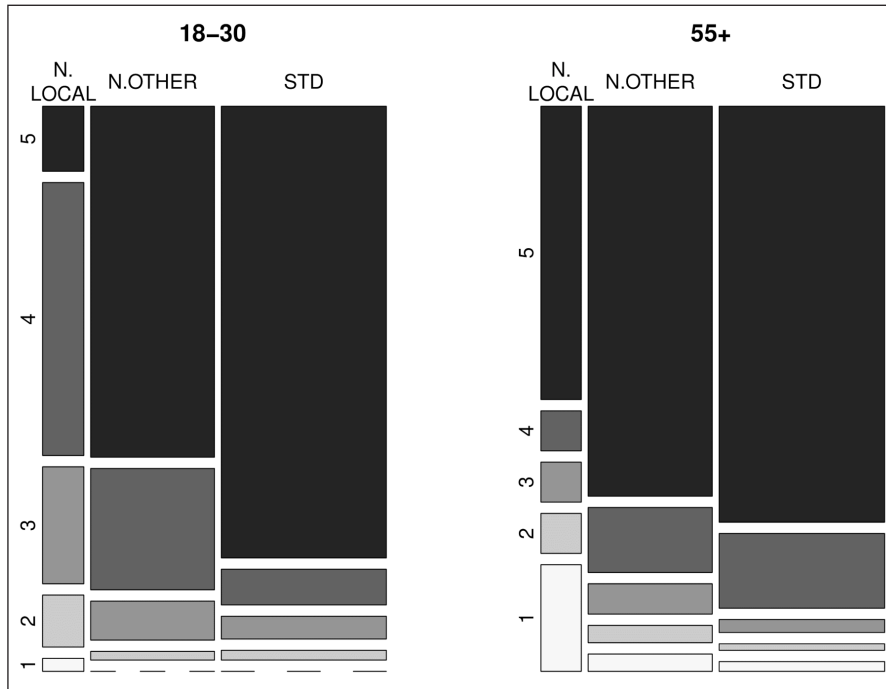
Coefficients:
Estimate Std. Error z value Pr(>|z|)
particleN -5.5790 0.4176 -13.360 < 2e-16 ***
AGEY 0.2804 0.5994 0.468 0.639962
particleN:AGEY 1.4407 0.4043 3.563 0.000366 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
Estimate Std. Error z value
112 -6.0262 0.5208 -11.571
213 -4.6909 0.4946 -9.484
314 -3.5096 0.4744 -7.398
415 -1.7579 0.4481 -3.923

```

**Figure 8:** Cumulative link mixed models for tag questions on negative anchors.

<sup>9</sup> Evidence was removed from this model as there were no effects or interactions.



**Figure 9:** Acceptability judgment results for biased questions, by age group.

```

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 640 -471.20 970.40 1046(4973) 1.14e-04 2.7e+02

Random effects:
Groups Name Variance Std.Dev.
EXAMPLE (Intercept) 0.4399 0.6633
PARTICIPANT (Intercept) 1.2426 1.1147
Number of groups: EXAMPLE 32, PARTICIPANT 20

Coefficients:
Estimate Std. Error z value Pr(>|z|)
particleN.LOCAL -2.5967 0.7476 -3.473 0.000514 ***
particleN.OTHER -1.3848 0.5542 -2.499 0.012466 *
AGEY 0.2457 0.6143 0.400 0.689171
EVIDENCENEUT -1.0069 0.4852 -2.075 0.037986 *
particleN.LOCAL:AGEY -1.1772 0.5674 -2.075 0.038010 *
particleN.OTHER:AGEY -0.7123 0.4621 -1.541 0.123207
particleN.LOCAL:EVIDENCENEUT 0.6351 0.9284 0.684 0.493886
particleN.OTHER:EVIDENCENEUT 1.6531 0.6966 2.373 0.017646 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
Estimate Std. Error z value
1|2 -6.0594 0.6251 -9.693
2|3 -5.1911 0.5930 -8.754
3|4 -4.1767 0.5703 -7.324
4|5 -2.5812 0.5450 -4.736
    
```

**Figure 10:** Cumulative link mixed model results for biased questions.

On the other hand, the 18–30 group’s median is 4, with over 50% of examples rated 4. The mean is still, however, mid-range (3.63), with only 12.5% of judgments rated 1 or 2.

If we were simply to take the mean judgments of the older and younger speakers, then, we could quite easily conclude that the participants were behaving in the same way.



However, the way in which the two groups reached this result was quite different, as shown by the model, and further illuminated by the descriptive statistics.

### 5.8 Summary

In the quantitative results from the Shetland participants, we see patterns of behaviour clearly based on age. For sentences that are known to be grammatical in both the dialect and the standard variety, we see ceiling level ratings from speakers in both age groups; for ungrammatical examples with no clearly dialectal features, we see floor level ratings from speakers in both age groups. However, with ungrammatical examples that contain “dialectal” features and with examples undergoing change, the results are more complicated. In ungrammatical cases, younger speakers rate the sentences as more acceptable than older speakers do; in cases of ongoing change, the mean ratings for older and younger speakers are very similar, but reached in very different ways, with younger speakers’ judgments clumping in the middle as opposed to the more polarised judgments of the older speakers. How can we interpret these findings? Do these findings correspond to usage patterns? In the next section I consider these speakers’ behaviour through the lenses of *hyperdialectalism* (Labov 1972) and *linguistic insecurity* as defined by Preston (2013), using the qualitative comments that accompanied the judgments speakers gave.

## 6 Discussion

A first possible analysis for the data worth entertaining is that in the face of ongoing linguistic change, young speakers in Shetland are simply unsure about giving judgments, hedging their bets in case they get something “wrong”. This would account for, for example, the data on tags on negative anchors presented in Section 5.6: younger speakers are unsure of what other dialect speakers would do. However, this would be surprising as younger speakers have been shown to have greater perceptual awareness and metalinguistic understanding than older speakers in metalinguistic tasks (e.g. Drager 2011; Carrera-Sabaté 2014; Lawrence 2017). Furthermore, in the data, younger speakers *do* appear to be certain about particular judgments, such as the tags on positive anchors shown in Section 5.5, and treat ungrammatical examples with dialect features, as shown in Section 5.4, differently to ungrammatical examples with no dialect features (as shown in Section 5.3). The younger speakers do not seem to be simply uncertain about their judgments.

A second possible analysis is that we are dealing with what Labov (1972) terms *hyperdialectalism*. Hyperdialectalism is a form of *hyperadaptation* in which speakers become “hyperlocal” (Labov 1972), increasing their use of particular features of the traditional dialect – especially in cases where there is ongoing language attrition. Speakers may extend a feature to contexts where it did not originally appear – *structural* hyperdialectalism (e.g. Trudgill 1986; Britain 2009). For example, loss of rhoticity along the English/Welsh border led to speakers in traditional rhotic areas developing rhotic forms in words with no etymological <r>, such as *last* [laʰst] (Britain 2009: 135). There is also the potential for *statistical* hyperdialectalism, where speakers do not extend use across contexts but simply increase the usage rates of a particularly salient feature (Wolfram & Schilling-Estes 2006). Smith & Durham (2011) propose that statistical hyperdialectalism is what we see with the younger speakers in Shetland who used higher percentages of *yon* as compared to older speakers.

We could conclude from the quantitative results that *-n* exhibits hyperdialectal usage patterns in Shetland dialect for younger speakers in the present day – both structural, via extension to tags on negative anchors (Section 5.6) and statistical, via usage in matrix biased questions at a rate higher than expected (Section 5.7). However, qualitative data indicates that it is not the case that younger speakers believe they produce *-n* in contexts that older speakers do not. Discussion with participants makes it clear that younger

participants are giving mid-range ratings, as we saw for *-n* on negative anchors and in biased questions, based on the perception that other people in the community might say something, even though they themselves wouldn't. This is shown in example (32), in which a younger speaker rates an example of *-n* in a rhetorical *wh*-question, another context in which *-n* is not attested in the dialect, nor acceptable for older speakers.

- (32) We are organising a birthday party. You have spoken to all of our friends to see if they can come, and everyone was able to come. When I ask you who is able to come to the party, you say: **Who *can'n* come?!**
- SY07: I wouldn't use *can'n*, I'd say "who canna come?"
- INT: Mmm. So what would you give it?
- SY07: I'll give it a four.
- INT: Do you think some folk might say it?
- SY07: Yeah cos well my auntie she's– they stay up in North Roe and she's *really* broad so she definitely says it and I've heard it from her and like folk around there.

Here, SY07 explicitly states that she would not use the example given, and gives the alternative that she would use, with standard Scots *-na* negation: *canna*. However, she still gives the example a rating of 4, the second highest rating. SY07 justifies this in a number of ways, drawing a clear distinction between herself and speakers who are "*really* broad". By stating that (older) members of her own family would use a variant, SY07 implicitly defends her position as an in-group member. Furthermore, in mentioning a specific area (North Roe, a rural area at the northmost point of mainland Shetland) where a particular feature could be used, SY07 is marking out her perceived understanding of the dialect. However, examples like the one in (32) received a mean rating of 2.23 from the older speakers, with a mode of 1. The younger speaker's understanding of the variation in the community, therefore, does not seem to be an accurate perception and would extend the potential for *-n* to a context where it is not accepted for users of a more traditional form of the dialect.

This type of justification given by the participant also appeared to extend to other ungrammatical examples, such as the instance in (33). Participant SY05 was one of the youngest participants at age 18, but was also from Whalsay, a distinctive dialect area which is reputedly particularly "broad" (Cohen 1978; Melchers 1985; Durham 2017).

- (33) I want to try out this new restaurant down the road. I ask you if it's any good, but you say: **I no am been yet.**
- SY05: mmm no so much, maybe like a three
- INT: three, so yeah, so like—
- SY05: like I've heard it, but I wouldn't say it myself

Despite the fact that the example shown in (33) is entirely ungrammatical in terms of its word order, both in standard English and in Shetland dialect, SY05 states that this is something she has "heard". However, she clearly goes on to distance herself from it – "I wouldn't say it myself". The participant here may have been commenting on the use of the *be*-perfect in this example,<sup>10</sup> a traditional feature of Shetland dialect that SY05 consistently gave ratings of 5 to in assumed grammatical filler examples (e.g. *Are you been to Majorca?*). It may be that the participant has an awareness of the *be*-perfect and is suggesting that other speakers use it in different ways (e.g. with this unusual word order).

<sup>10</sup> Thanks to an anonymous reviewer for observing the potential relevance of the *be*-perfect in this example.

Younger participants also made similar comments about imperatives with *do no*, which was discussed in Section 5.4. For example, participant SY02 said that the example presented in (24) was something that “older” speakers would use, and that while she herself would use *dunna*, she nonetheless gave it a rating of 4.

On the other hand, older participants did not attempt to attribute ungrammatical examples to either younger speakers or to a previous generation of dialect speakers.<sup>11</sup> They were also generally more disparaging in their comments about ungrammatical examples, describing *-n* in contexts like (32) as “doesn’t make any sense” and “daft”, and tags on negative anchors as in (28) as “bad” or “contradictory”.

The results of the younger speakers’ acceptability judgments therefore do not seem to reflect hyperdialectal *usage* patterns, despite the fact that these are claimed to be part of the variety. However, I interpret these results as an instance of hyperdialectalism within the realm of perceptual dialectology. While the younger speakers use the *-n* particle in certain contexts (e.g. tags on positive anchors), they are aware that, traditionally, *-n* can be used in more contexts than this. However, they aren’t accurately able to pinpoint what those contexts are. They therefore extend the possibility of a variant to new contexts (e.g. tags on negative anchors) and lose the defined original environments for use (e.g. matrix biased questions) – just as we see in hyperdialectalism in usage patterns. This then extends to slightly higher ratings for plausibly dialectal ungrammatical examples (e.g. imperatives with *do no*), through a general willingness to accept that there may be dialect features that they as younger speakers do not use.

This phenomenon does not appear to be limited to the *-n* particle. Earlier judgment work on Shetland dialect also shows younger speakers exhibiting these patterns with respect to a different set of variables. In work investigating the loss of verb raising in imperatives and questions in Shetland, Jamieson (2015) finds that speakers in their younger (18–30) age bracket give mid-range ratings to examples of verb raised questions, as in (34). Participants gave the same kinds of comments about who exactly would use these examples – for example, participant 01A cites an older family member as someone who would use the verb raised example.

- (34) Jamieson (2015: 56)  
 INT: [...] “What says you?”  
 02A: You do hear that, yeah, “what says you?” but not very often. Three maybe.  
 01A: It’s what my grandad would say.  
 02A: Yeah.

Speakers in Jamieson’s older age bracket (60 +) also did not rate these questions at ceiling level. However, these speakers’ comments indicated they would use these variants *sometimes*, rather than attributing use of them to other community members (35–36).

- (35) Jamieson (2015: 54)  
 04J: The quantity of questions with the verb coming first [...] which sometimes we use, but most time we *dunna* [...] but they are– they are times when it would be appropriate.

<sup>11</sup> One older speaker, SO06, claimed that a tag like “you can come, can’n you no?” was possibly used in more rural areas. Constructions like these, with both *-n* and *no*, appear to be a mid-point in the changes taking place from use of e.g. *can’n* to the standard Scots *can you no* question in Shetland dialect (Jamieson 2018). Participant SO06 was from Lerwick, Shetland’s main town, where levelling is particularly rapid (e.g. Sundkvist 2011). SO06 actually gave examples like *can’n* in biased questions, the context where loss is occurring, a rating of 1. It therefore seems as though SO06’s observation here is an accurate acknowledgment of the fact that this is a variant available for some speakers, and that her own (urban) variety of the dialect may be further ahead in the change.

- (36) Jamieson (2015: 55)
- 01I: I'm very aware that where I'm hearing it and where I'm saying it has got a lot to do with my age.
- INT: Right OK, and so is there certain kind of contexts where you would be more likely to use this kind of thing?
- 01I: yeah [...] with mam or with [sister] or [...] with my own age group

This is similar to what we see with *-n* in matrix biased question contexts in this research: though there was no statistical difference in younger and older speakers' judgments (see Section 5.6), the information conveyed by the numerical judgment was different, as reflected by the qualitative comments.

Taking the qualitative comments into account, younger speakers of the obsolescing Shetland dialect often give judgments that do not seem to reflect their own grammars. Beyond what they know they definitely say, they rely on their hypotheses about older or more rural speakers' grammars when giving judgments. This is in sharp contrast to older speakers of the same variety, whose judgments are more definitive and whose comments appeared to reflect their own usage, rather than hypothesising about other speakers' dialects. I term the phenomenon seen here in the Shetland results *perceptual hyperdialectalism* – exhibiting the same sorts of structural and statistical effects as in hyperdialectal usage, but through the perspective of a metalinguistic task.

The behaviour of the younger speakers in this research is reminiscent of Henry's (1995) speakers of Dialect A, a subset grammar of Dialect B, who formed "incorrect hypotheses" about what Dialect B speakers could do. The results here support Henry's point that in dialect syntax research, in order to access a grammar it is crucial to get judgments from speakers for whom that *is* their grammar. However, the results here from Shetland dialect also show younger speakers forming incorrect hypotheses about features that are *not* undergoing change (i.e. *do no*), indicating that there is something broader at play.

This final point also suggests that it is not that younger speakers are reorganising the system following the underlying syntactic ecology of the language, as Adger (2017) argues for East Sutherland Gaelic. Younger speakers do not claim to produce *-n* in new contexts; rather, they seem to be attempting to establish their knowledge of traditional dialect more generally, claiming awareness of other speakers' dialect use.

I therefore propose that the effects of perceptual hyperdialectalism seen in the Shetland dialect results here is due to the sort of *linguistic insecurity* defined by Preston (2013). Traditionally, linguistic insecurity (Labov 1966) has been defined as occurring when speakers feel "that the variety they use is somehow inferior, ugly or bad" (Meyerhoff 2006: 292). However, Preston (2013) argues that it can be seen in terms of an individual speaker's relationship to their regional standard, "when one feels that they are not able to perform the linguistic job at hand" (Preston 2013: 324). More specifically, speakers may believe that their own variety is "correct", but feel they are lacking in their ability to use it. I propose this is what is happening with the younger Shetland dialect speakers in this research.<sup>12</sup>

<sup>12</sup> There are obvious parallels to be drawn here with Dorian's (1977; 1994) "semi-speakers". Older and/or more fluent members of the community would recognise these people as "semi-speakers", and openly make judgments that they didn't speak Gaelic "right". In turn, the semi-speakers were aware of their status in the community, aware of their linguistic errors, and embarrassed about their lack of the language (Dorian 1994) – in effect, they were insecure. Within the Shetland community, speakers of all ages play a part in the identification of (local) people who don't speak the dialect – this happened during this research, and is evidenced in Smith & Durham's work (2011: 217). However, I do not wish to label the speakers in this research as "semi-speakers". For one, it is not clear how this terminology would apply in a community where use of the dialect exists on a continuum (e.g. Smith & Durham 2012), rather than competition between two clearly distinct languages (as seen in Dorian's work, where Gaelic was competing with English). Instead, I frame the younger speakers in the broader terminology of linguistic insecurity.



Wolfram & Schilling-Estes (1995: 698) note that “nonmainstream dialects... play a large role in the shaping of cultural identity” and that obsolescence of a dialect threatens this cultural identity (1995: 699). Children of Shetland heritage have a generally positive attitude towards the dialect and consider it to be “an important facet of their identity” (Durham 2014: 303). Furthermore, there is a strong desire for them to distance themselves from “soothmothers” (Karam 2017), people who are not members of the community in-group. Despite these positive attitudes and the clear importance of the dialect as an tool for identity construction, however, Durham (2014, 2017) shows that younger Shetland speakers’ self-reported usage estimates lean more strongly towards English than dialect when compared to the children in Melchers’ (1985) 1983 corpus, even in informal, local contexts such as speaking to a friend who is also from Shetland. Stadler (2016) and Stadler et al. (2016) show that both younger and older speakers in Shetland were able to identify the direction of a morphosyntactic change and place themselves in relation to the change curve, with younger speakers consistently positioning themselves ahead of the curve to a greater extent than older speakers. Taken together, these findings indicate that younger speakers are aware that they are using fewer traditional dialect features, while still maintaining that dialect use is a core part of their identity as Shetlanders.

How does this relate to acceptability judgment tasks? When given acceptability judgment tasks, speakers are asked explicitly to act as “gatekeepers” for the variety (Jørgensen 2010) and are given an opportunity to make conscious choices about their linguistic resources in order to signal in-group membership (Coupland 2007). Following Le Page & Tabouret-Keller (1985: 181):

“The individual creates for themselves the patterns of their linguistic behaviour so as to resemble those of the group or groups with which from time to time they wish to be identified, or so as to be unlike those from whom they wish to be distinguished.”

Aware that their grammars are different to those of older Shetland dialect speakers, and insecure about how this fact relates to their identity as Shetlanders, younger participants turn to what they “know”, or think they know, about the linguistic features of a more traditional grammar – distinguishing their own usage (majority 5 ratings, such as *-n* in tags on positive anchors) with what they perceive to be possible, or overheard (majority 3–4 ratings, such as *-n* in matrix biased questions or in tags on negative anchors). Participants back this up by making explicit their connections to the community, legitimizing their right to be making judgments, as seen in examples like (32–34). Though in the task participants are not employing the resources in conversation, they are doing as much as they can to preserve their in-group status through “knowledge” of a more traditional form of the dialect. However, we can compare with older speakers’ behaviour and see that younger speakers’ “knowledge” of older speakers’ grammars is not necessarily accurate. Their desire to signal their in-group membership leads to judgments that do not reflect any current grammar of the dialect, and which must be carefully interpreted.

## 7 Conclusions

As highlighted at the beginning of this paper, acceptability judgment methodologies have been shown to be reliable sources for understanding stable variation within communities. However, the reliability of this methodology for dealing with linguistic change, particularly the unpredictable sort of variation attested at the end of a change, had not been explored. I thus investigated the reliability of acceptability judgment methodologies within a variety that is undergoing obsolescence: the Shetland dialect of Scots.



Older and younger speakers of the dialect patterned together when it came to filler examples that did not include any dialect features: grammatical examples were rated at “ceiling”, while ungrammatical examples were rated at “floor”. However, when dialect features were introduced, older and younger speakers diverged in their patterns of judgment. Examples that were truly acceptable were rated at ceiling for both age groups. For unacceptable examples, younger speakers gave mid-range ratings as compared to the older speakers’ low ratings; for examples undergoing change, younger speakers generally rated examples mid-range, while older speakers were more polarised in their responses.

It didn’t seem that participants were restructuring a system as Adger’s (2017) work found in East Sutherland Gaelic, as this rating pattern held across four different examples (ungrammatical *do no* imperatives, *-n* in tag questions on negative anchors and matrix biased questions, and the verb raising in questions found in Jamieson 2015). Instead, these results appeared to show unpredictable behaviour on behalf of the younger participants, perhaps retaining variables at surprising rates or extending contexts for use.

Incorporating speakers’ qualitative comments into the discussion, however, I argued that these younger speakers were exhibiting *perceptual hyperdialectalism*, following the principles of both structural and statistical hyperdialectalism in the metalinguistic domain. I argued that this is a reflex of Preston’s (2013) definition of linguistic insecurity. These speakers, who hold dialect as an important part of their cultural identity but whose grammar is different from that of the older speakers, attempt to distinguish themselves from community outsiders through their “knowledge” of that traditional dialect.

This research strongly supports Henry’s (1995) point that getting judgments from speakers of the grammar you are interested in is crucial for dialect syntax research, and that failing to do this can make the judgments “unreliable”. However, the “unreliability” of the judgments made by speakers undergoing significant linguistic change can also be enlightening, if carefully interpreted alongside qualitative (metalinguistic) data.

## Abbreviations

N = biased question marker in Shetland, NEG = negation

## Acknowledgements

This research was supported by funding from the Arts and Humanities Research Council. Thanks to Jennifer Smith and Caroline Heycock for a great deal of valuable discussion and comments. Thanks also to the audience at the workshop on Language, Place and Periphery (University of Copenhagen, January 2018), to David Adger, Gary Thoms and Clara Cohen for additional discussion, and to three anonymous reviewers for their helpful comments and suggestions.

## Competing Interests

The author has no competing interests to declare.

## References

- Adger, David. 2017. Structure, use and syntactic ecology in language obsolescence. *Canadian Journal of Linguistics* 62(4). 614–638. DOI: <https://doi.org/10.1017/cnj.2017.32>
- Bader, Markus & Jana Häussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics*. 46(2). 273–330. DOI: <https://doi.org/10.1016/j.lingua.2009.05.007>
- Banuazizi, Atissa & Cassandre Creswell. 1999. Is that a real question? Final rises, final falls and discourse function in question intonation. *Chicago Linguistic Society (CLS)* 35(1). 1–13.

- Barbiers, Sjef & Hans Bennis. 2007. The Syntactic Atlas of the Dutch Dialects: A discussion of choices in the SAND project. *Nordlyd* 34. 53–72. DOI: <https://doi.org/10.7557/12.89>
- Bard, Ellen, Dan Robertson & Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1). 32–68. DOI: <https://doi.org/10.2307/416793>
- Biezma, Maria & Kyle Rawlins. 2017. Rhetorical questions: Severing questioning from asking. In Dan Burgdorf, Jacob Collard, Sireemas Maspong & Brynhildur Stefánsdóttir (eds.), *Proceedings of the 27<sup>th</sup> Semantics and Linguistic Theory Conference*, 302–322. University of Maryland: LSA. DOI: <https://doi.org/10.3765/salt.v27i0.4155>
- Biggs, Alison. 2015. A new case for A-movement in Northwest British English. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O’Hara & Iris Chuoying Ouyang (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 218–227. Somerville, MA: Cascadilla Proceedings Project.
- Britain, David. 2009. One foot in the grave? Dialect death, dialect contact and dialect birth in England. *International Journal of the Sociology of Language* 196/197. 121–155. DOI: <https://doi.org/10.1515/IJSL.2009.019>
- Campbell, Lyle, & Martha C. Muntzel. 1989. The structural consequences of language death. In Nancy Dorian (ed.), *Investigating obsolescence*, 181–196. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620997.016>
- Carrera-Sabaté, Josefina. 2014. Does meta-linguistic awareness play any role at the beginning of an ongoing sound change? The case of some vowel-ended verbs in Catalan. *Sociolinguistic Studies* 8(2). 193–221. DOI: <https://doi.org/10.1558/sols.v8i2.193>
- Chomsky, Noam. 1966. Topics in the theory of generative grammar. In Thomas Sebeock (ed.), *Current trends in linguistics 3: Linguistic theory*. Bloomington: Indiana University Press.
- Christian, Donna, Walt Wolfram & Najo Dube. 1988. *Variation and change in geographically isolate communities: Appalachian English and Ozark English*. Tuscaloosa: University of Alabama Press.
- Christensen, Rune Haubo Bojesen. 2018. ordinal – Regression models for ordinal data. *R package*. <https://CRAN.R-project.org/package=ordinal>.
- Cohen, Anthony. 1978. The same – but different: The allocation of identity in Whalsay, Shetland. *The Sociological Review* 26(3). 449–470. DOI: <https://doi.org/10.1111/j.1467-954X.1978.tb00141.x>
- Cook, Eung-Do. 1989. Is phonology going haywire in dying languages? Phonological variations in Chipewyan and Sarcee. *Language in Society* 18. 235–255. DOI: <https://doi.org/10.1017/S0047404500013488>
- Cornips, Leonie & Cecilia Poletto. 2005. On standardizing syntactic elicitation techniques. *Lingua* 115(7). 939–957. DOI: <https://doi.org/10.1016/j.lingua.2003.11.004>
- Coupland, Nikolas. 2007. *Style: Language variation and identity*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511755064>
- de Winter, Joost & Dimitra Dodou. 2010. Five-point Likert items: t-test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation* 15(11). 1–16.
- Domaneschi, Filippo, Maribel Romero & Bettina Braun. 2017. Bias in polar questions: Evidence from English and German production experiments. *Glossa* 2(1). 23. 1–28. DOI: <https://doi.org/10.5334/gjgl.27>
- Dorian, Nancy. 1973. Grammatical change in a dying dialect. *Language* 49. 413–438. DOI: <https://doi.org/10.2307/412461>
- Dorian, Nancy. 1977. The problem of the semi-speaker in language death. *International Journal of the Sociology of Language* 12. 22–32. DOI: <https://doi.org/10.1515/ling.1977.15.191.23>

- Dorian, Nancy. 1994. Varieties of variation in a very small place: Social homogeneity, prestige norms and linguistic variation. *Language* 70(4). 631–696. DOI: <https://doi.org/10.2307/416324>
- Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech* 54(1). 99–121. DOI: <https://doi.org/10.1177/0023830910388017>
- Durham, Mercedes. 2013. Was/were alternation in Shetland English. *World Englishes* 32(1). 108–128. DOI: <https://doi.org/10.1111/weng.12009>
- Durham, Mercedes. 2014. Thirty years late: Real-time change and stability in attitudes towards dialect in Shetland. In Robert Lawson (ed.), *Sociolinguistics in Scotland*, 296–318. Basingstoke: Palgrave Macmillan. DOI: [https://doi.org/10.1057/9781137034717\\_15](https://doi.org/10.1057/9781137034717_15)
- Durham, Mercedes. 2017. Changing domains of dialect use: A real-time study of Shetland schoolchildren. In Sylvie Hancil & Joan Beal (eds.), *Perspectives on Northern Englishes*, 245–270. Berlin: Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110450903-011>
- Hedberg, Nancy, Juan M. Sosa & Emrah Görgülü. 2017. The meaning of intonation in yes-no questions in American English: A corpus study. *Corpus Linguistics and Linguistic Theory* 13(2). 321–368. DOI: <https://doi.org/10.1515/cllt-2014-0020>
- Henry, Alison. 1995. *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford: Oxford University Press.
- Heycock, Caroline, Antonella Sorace, Zakaris Svabo Hansen, Frances Wilson & Sten Vikner. 2012. Detecting the late stages of syntactic change: The loss of V-to-T in Faroese. *Language* 88(3). 558–600. DOI: <https://doi.org/10.1353/lan.2012.0053>
- Hill, Jane. 1989. The social functions of relativization in obsolescent and nonobsolescent languages. In Nancy Dorian (ed.), *Investigating obsolescence*, 149–164. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620997.014>
- Horn, Laurence. 2013. I love me some datives: Expressive meaning, free datives and F-implicature. In Daniel Gutzmann & Hans-Martin Gaertner (eds.), *Beyond expressives: Explorations in use-conditional meaning*, 151–199. Leiden: Brill. DOI: [https://doi.org/10.1163/9789004183988\\_006](https://doi.org/10.1163/9789004183988_006)
- Ingason, Anton. 2010. *Explaining case variation: A death rattle hypothesis for minority rules*. Presented at the ScanDiaSyn/N’CLAV Grand Meeting, Sommarøya, Trømsø. 8 June 2010.
- Jamieson, E. 2015. *An investigation of verb raising in the Shetland dialect of Scots*. Edinburgh: University of Edinburgh Masters thesis.
- Jamieson, E. 2018. *Questions, biases and ‘negation’: Evidence from Scots varieties*. Edinburgh: University of Edinburgh PhD thesis.
- Jørgensen, J. N. 2010. *Languaging – Nine years of poly-lingual development of young Turkish-Danish grade school students*. Copenhagen: University of Copenhagen.
- Karam, Kerry. 2017. *“Knappin”: Standard vs dialect speech modification in Shetland*. Aberdeen: University of Aberdeen PhD thesis.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics. DOI: <https://doi.org/10.1017/CBO9780511618208>
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Langsford, Steven, Amy Perfors, Andrew Hendrickson, Lauren Kennedy & Danielle Navarro. 2018. Quantifying sentence acceptability measures: Reliability, bias and variability. *Glossa* 3(1). 37. 1–34. DOI: <https://doi.org/10.5334/gjgl.396>

- Lawrence, Daniel. 2017. The social perception of a sound change. *Penn Working Papers in Linguistics* 23(1). 111–120.
- Le Page, Robert & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Linzen, Tal & Yohei Oseki. 2018. The reliability of acceptability judgments across languages. *Glossa* 3(1). 100. 1–25. DOI: <https://doi.org/10.5334/gjgl.528>
- Maling, Joan & Sigríður Sigurjónsdóttir. 2002. The “new impersonal” construction in Icelandic. *Journal of Comparative Germanic Linguistics* 5. 97–142. DOI: <https://doi.org/10.1023/A:1021224923742>
- Millar, Robert McColl. 2008. The origins and development of Shetland dialect in light of dialect contact theories. *English World-Wide* 29(3). 237–267. DOI: <https://doi.org/10.1075/eww.29.3.02mil>
- Melchers, Gunnel. 1985. “Knappin”, “Proper English”, “Modified Scottish”: Some language attitudes in the Shetland Isles. In Manfred Görlach (ed.), *Focus on: Scotland*, 87–100. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/veaw.g5.07mel>
- Melchers, Gunnel. 1997. This, that, yon: on “three-dimensional” deictic systems. In Jenny Cheshire & Dieter Stein (eds.), *Taming the vernacular*, 83–92. London: Longman. DOI: <https://doi.org/10.4324/9781315841410>
- Melchers, Gunnel. 2004. English spoken in Orkney and Shetland: Morphology, syntax and lexicon. In Bernd Kortmann, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie & Clive Upton (eds.), *A handbook of varieties of English (vol. 2: morphology and syntax)*, 34–46. Berlin: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110208399.2.285>
- Melchers, Gunnel & Peter Sundkvist. 2010. Orkney and Shetland. In Daniel Schrier, Peter Trudgill, Edgar W. Schneider, & Jeffrey P. Williams (eds.), *The lesser known varieties of English: An introduction*, 17–34. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511676529.003>
- Meyerhoff, Miriam. 2006. *Introducing sociolinguistics*. London: Routledge. DOI: <https://doi.org/10.4324/9780203966709>
- Millar, Martin & Keith Brown. 1979. Tag questions in Edinburgh speech. *Linguistische Berichte* 60. 24–45.
- Milroy, Lesley. 1980. *Language and social networks*. Baltimore, MD: University Park Press.
- Myler, Neil. 2013. On “coming the pub” in the North West of England: accusative unaccusatives, dependent case and preposition incorporation. *Journal of Comparative Germanic Linguistics* 16(2–3). 189–207. DOI: <https://doi.org/10.1007/s10828-013-9055-1>
- Nagata, Hiroshi. 1992. Anchoring effects in judging grammaticality of sentences. *Perceptual and Motor Skills* 75(1). 159–164. DOI: <https://doi.org/10.2466/pms.1992.75.1.159>
- Newmeyer, Frederick. 1983. *Grammatical theory: Its limits and its possibilities*. Chicago: University of Chicago Press.
- Payne, Arvilla. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.), *Locating language in time and space*, 143–178. New York: Academic Press.
- Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonation contours in the interpretations of discourse. In Philip Cohen, Jerry Morgan & Martha Pollack (eds.), *Intentions in communications*, 271–311. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/3839.003.0016>
- Potsdam, Eric. 1996. *Syntactic issues in the English imperative*. Santa Cruz: University of California PhD thesis. DOI: <https://doi.org/10.4324/9781315052830>



- Preston, Dennis. 2013. Linguistic insecurity forty years later. *Journal of English Linguistics*, 41(4). 304–331. DOI: <https://doi.org/10.1177/0075424213502810>
- Quirk, Randolph & Sidney Greenbaum. 1970. *Elicitation experiments in English: Linguistic studies in use and attitude*. Harlow: Longman.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Robertson, Thomas Alexander & John J. Graham. 1952. *Grammar and usage of the Shetland dialect*. Lerwick: The Shetland Times.
- Romero, Maribel & Chung-Hye Han. 2004. On negative yes-no questions. *Linguistics & Philosophy* 27(5). 609–658. DOI: <https://doi.org/10.1023/B:LING.0000033850.15705.94>
- Rupp, Laura. 2003. *Syntax of imperatives in English and Germanic: Word order variation in the minimalist framework*. Hampshire: Palgrave Macmillan. DOI: <https://doi.org/10.1057/9780230505179>
- Schilling-Estes, Natalie & Walt Wolfram. 1999. Alternative models of dialect death: Dissipation vs concentration. *Language* 75(3). 486–521. DOI: <https://doi.org/10.2307/417058>
- Schütze, Carson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press. DOI: [https://doi.org/10.26530/OAPEN\\_603356](https://doi.org/10.26530/OAPEN_603356)
- Shetland Islands Council. 2017. *Shetland In Statistics 42*. Lerwick, Shetland. [https://www.shetland.gov.uk/economic\\_development/documents/ShetlandInStatistics\\_2017\\_241117.pdf](https://www.shetland.gov.uk/economic_development/documents/ShetlandInStatistics_2017_241117.pdf).
- Skovlund, Eva & Grete Usterud Fenstad. 2001. Should we always choose a nonparametric test when comparing two apparently nonnormal distributions? *Journal of Clinical Epidemiology* 54. 86–92. DOI: [https://doi.org/10.1016/S0895-4356\(00\)00264-X](https://doi.org/10.1016/S0895-4356(00)00264-X)
- Smith, Jennifer, David Adger, Brian Aitken, Caroline Heycock, E. Jamieson & Gary Thoms. 2019. The Scots Syntax Atlas. University of Glasgow. <https://scotssyntaxatlas.ac.uk>.
- Smith, Jennifer & Mercedes Durham. 2011. A tipping point in dialect obsolescence? Change across the generations in Lerwick, Shetland. *Journal of Sociolinguistics* 15(2). 197–225. DOI: <https://doi.org/10.1111/j.1467-9841.2011.00479.x>
- Smith, Jennifer & Mercedes Durham. 2012. Bidialectalism or dialect death: Explaining generation change in the Shetland Islands, Scotland. *American Speech* 87(1). 57–88. DOI: <https://doi.org/10.1215/00031283-1599959>
- Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality and experimental syntax. *Biolinguistics* 1. 123–134.
- Sprouse, Jon & Diogo Almeida. 2017. Setting the empirical record straight: Acceptability judgments appear to be reliable, robust and replicable. *Behavioural and Brain Sciences*, 40(E311). DOI: <https://doi.org/10.1017/S0140525X17000590>
- Sprouse, Jon, Carson Schütze & Diogo Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua* 134. 219–248. DOI: <https://doi.org/10.1016/j.lingua.2013.07.002>
- Stadler, Kevin. 2016. *Direction and directedness in language change: An evolutionary model of selection by trend-amplification*. Edinburgh: University of Edinburgh PhD thesis.
- Stadler, Kevin, E. Jamieson, Kenny Smith & Simon Kirby. 2016. Metalinguistic awareness of trends as a driving force in linguistic evolution: An empirical study. In Sean Roberts, Christine Cuskley, Luke McCrohon, Lluís Barceló-Coblijn, Olga Feher & Tessa Verhoeft (eds.), *The evolution of language: Proceedings of the 11th international conference (EVOLANG11)*. <http://evolang.org/neworleans/papers/145.html>.
- Stonehouse, John & Guy Forrester. 1998. Robustness of the t and U tests under combined assumption violations. *Journal of Applied Statistics* 25. 63–74. DOI: <https://doi.org/10.1080/02664769823304>



- Sundkvist, Peter. 2011. The Shetland Islands: globalisation and the changing status of Standard English. *English Today* 27(4). 19–25. DOI: <https://doi.org/10.1017/S0266078411000496>
- Tait, James. 1973. Tammy Stays Put. *The New Shetlander*, 103. 13.
- Tait, James. 2007. Shetland Identity and (the) (Sh(a)etlan(d(ic))) (D/dialect). *Shetland Life*, 321. 10–11.
- Thoms, Gary. 2014. Report for “Refining the questionnaire method of judgment data collection: a Buckie pilot study” pilot project. Ms. Edinburgh: University of Edinburgh.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.
- Tsoulas, George & Rebecca Woods. 2019. Predicative possessives, relational nouns and floating quantifiers. *Linguistic Inquiry* 50(4). 825–846. DOI: [https://doi.org/10.1162/ling\\_a\\_00320](https://doi.org/10.1162/ling_a_00320)
- van Leyden, Klaske. 2004. *Prosodic characteristics of Orkney and Shetland dialects: An experimental approach*. Leiden: Leiden University PhD thesis.
- Weir, Andrew. 2013. The syntax of imperatives in Scots. In Janet Cruickshank & Robert McColl Millar (eds.), *After the storm: Papers from the Forum for Research on the Languages of Scotland and Ulster triennial meeting, Aberdeen 2012*, 261–285. Aberdeen: University of Aberdeen.
- Wochner, Daniela, Jana Schlegel, Nicole Dehé & Bettina Braun. 2015. The prosodic marking of rhetorical questions in German. *Interspeech 2015*, 987–991. ISCA Archive.
- Wolfram, Walt & Natalie Schilling-Estes. 1995. Moribund dialects and the endangerment canon: The case of Ocracoke brogue. *Language* 71(4). 696–721. DOI: <https://doi.org/10.2307/415741>
- Wolfram, Walt & Natalie Schilling-Estes. 2006. *American English*. Oxford: Blackwell.
- Wood, Jim & Raffaella Zanuttini. 2018. Datives, data and dialect syntax in American English. *Glossa* 3(1). 87. 1–22. DOI: <https://doi.org/10.5334/gjgl.527>
- Zanuttini, Raffaella, Jim Wood, Jason Zentz & Laurence Horn. 2018. The Yale Grammatical Diversity Project: Morphosyntactic variation in North American English. *Linguistics Vanguard* 4(1). 1–15. DOI: <https://doi.org/10.1515/lingvan-2016-0070>
- Zimmerman, Donald & Bruno Zumbo. 1993. The relative power of parametric and non-parametric statistical methods. In Gideon Keren & Charles Lewis (eds.), *A handbook for data analysis in the behavioural sciences: Methodological issues*, 481–517. Hillsdale, NJ: Erlbaum.

**How to cite this article:** Jamieson, E. 2020. Viewing dialect change through acceptability judgments: A case study in Shetland dialect. *Glossa: a journal of general linguistics* 5(1): 19.1–28. DOI: <https://doi.org/10.5334/gjgl.979>

**Submitted:** 24 April 2019    **Accepted:** 03 December 2019    **Published:** 02 March 2020

**Copyright:** © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS