Journal of Linguistics (2025), **61**, 369–396
doi:10.1017/S0022226724000203

Journal of Linguistics
CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Explaining microvariation using the Tolerance Principle: plugging the *amn't* gap

Gary Thoms[1], David Adger[2], Caroline Heycock[3], E. Jamieson[4] and Jennifer Smith[5]

[1]New York University, 10 Washington Place, New York, 10003, USA
[2]Queen Mary University of London, Mile End Road, London, E1, 4NS, United Kingdom
[3]University of Edinburgh, 3 Charles Street, Edinburgh, EH8 9AD, United Kingdom
[4]University of York, Heslington, York, YO10 5DD, United Kingdom
[5]University of Glasgow, 12 University Gardens, Glasgow, G12, 8QH, United Kingdom
Corresponding author: Gary Thoms; Email: gst1@nyu.edu

Received: 26 May 2023; **Revised:** 12 February 2024; **Accepted:** 18 March 2024

Keywords: dialect syntax; morphological gaps; negation; productivity; Scots

Abstract
In this article, we describe and explain patterns of variation in acceptance of *amn't* in varieties of Scots, drawing upon data from the Scots Syntax Atlas. Partly in line with findings from Bresnan (2001), we show that *amn't* is much more widely accepted in inversion environments (*amn't I*?) than in declaratives (*I amn't*), but nevertheless, *amn't* in declaratives is still accepted in certain regions of Scotland. We combine the productivity-based explanation of the *amn't* gap in Yang (2016, 2017) with new insights into the syntax of Scots negation from Thoms et al. (2023) to provide a predictive account of the attested variation.

1. Introduction

A well-known puzzle with English negation is that, in most varieties, it is not possible to attach the reduced form of negation *-n't* to *am* (Langendoen 1970, Zwicky & Pullum 1983). Examples such as (2b) with the form *amn't* are judged as unacceptable, and speakers typically report that there is no other appropriate form for attaching negation to the auxiliary. Only the uncontracted negation can be used, as in (2a). For some speakers, however, the form *aren't* is available in inversion contexts, as in (2d). This is unexpected, given that (2e) is consistently unavailable.

(1) (a) She's not your friend
(b) She isn't your friend.
(c) Isn't she your friend?

(2) (a) I'm not your friend
(b) *I *amn't* your friend.
(c) **Amn't* I your friend?
(d) *Aren't* I your friend?
(e) *I *aren't* your friend.

This absence of an *am+n't* form has come to be known as 'the *amn't* gap', in recognition of the fact that it represents a gap in what is generally thought of as a fairly consistent pattern of attachment of *-n't* to finite auxiliaries (Francis 1985, Bresnan 2001, Anderwald 2002, Broadbent 2009).

Although the judgments in (2) hold for most varieties spoken in North America and England, there is some degree of dialectal variation in the acceptance of (2b) and (2c). Hudson (2000) and Bresnan (2001) report that Irish English speakers accept *amn't* both in declarative, like (2b), and inverted orders, like (2c). In the varieties spoken in Scotland (herein, Scots), however, there is a curious asymmetry: *amn't* is acceptable in inverted orders like (2c), but not in declaratives like (2b). We will call this the 'inversion asymmetry'.

Bresnan (2001) proposes an account of the inversion asymmetry in terms of morphosyntactic competition, connecting it to the fact that Scots varieties have an additional form for contracted negation, *-nae.* This negative form can combine with all the auxiliaries that *-n't* can and, according to Bresnan (and, for example, Brown 1991), also with *am* to produce *amnae*; see (3).

(3) (a) She *didnae* see us
(b) I *amnae* your friend.

However, no combination of *-nae* and an auxiliary can invert with the subject; see (4). No such constraint obtains with *-n't*; see (5).

(4) (a) **Didnae* she see us
(b) **Amnae* I your friend?

(5) *Didn't* she see us?

Bresnan's (2001) proposal is that in Scots varieties, *amnae* outcompetes *amn't* in declaratives, but since *amnae* is not available in inversions, *amn't* becomes possible. In Irish English, however, since there is no *-nae* negation, *amn't* is used in both declaratives and inversions. Bresnan (2001) takes this account of the dialectal variation with *amn't* in inversions, and the accompanying account of the appearance of spurious *aren't*, to provide an empirical argument for optimality-driven morphosyntactic competition more generally (but see Embick & Marantz 2008).

However, the absence of *amn't* from varieties of English other than Irish English and Scots is ultimately taken by Bresnan (2001) to be an accidental lexical gap, and so although her account gives us a handle on the inversion asymmetry, it leaves the source of the *amn't* gap and its dialectal variability rather mysterious. More recently, Yang (2017) offers an analysis of the distribution of the *amn't* gap across varieties of English in terms of differences in the productivity of the rule for the contracted negative form *-n't.* He proposes that in American English, for example, there are, in fact, so many exceptions to the regularity of this 'rule' that

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

children do not treat it as productive but learn each combination of auxiliary and negation separately, thus leaving open the possibility of a learned 'gap'. For Scots and Irish English, however, he argues that there are fewer exceptions (mainly due to the absence of *ain't* in these varieties); hence, learners conclude that there is a productive rule of *-n't* suffixation, which will automatically result in the occurrence of *amn't.* Yang argues for a principle – his 'Tolerance Principle' – that determines the threshold at which a productive rule kicks in and shows how this principle predicts the empirical outcomes. In contrast to Bresnan's (2001) proposal, then, the Tolerance Principle offers an explanation of the presence or absence of the *amn't* gap across varieties.

Although the Tolerance Principle provides a new approach to the issue, it does not yet provide a satisfactory account of all the facts that Bresnan (2001) brought into the discussion. First, it offers no explanation for the inversion asymmetry in Scots and its absence from Irish English. Connected to this, as it stands, it has nothing to say about the potential role for the Scots-specific negation *-nae* in the conditioning of the behaviour of *amn't*, a possibility that was raised by Bresnan (2001). It is also unclear how the two negative attachment rules would be expected to interact in a system that adopts the Tolerance Principle, especially in the absence of an analysis of the morphosyntax of negation in the relevant varieties.

In this article, we build on Yang's productivity approach to the data but extend the theory, bringing in novel empirical data from microcomparative research into the syntax of negation, particularly in Scots varieties. We first describe the Scots empirical picture in detail, using fine-grained dialectal data from the Scots Syntax Atlas (Smith et al. 2019). We show that although the asymmetry Bresnan (2001) describes holds for some speakers, there are also many speakers who accept *amn't* in both declaratives and inversions, and that the variation observed is not tracked by acceptance/rejection of *amnae* forms in the way that Bresnan's (2001) account predicts. The atlas data allow us to see that there is a geographical coherence to where *amn't* is accepted in both syntactic contexts (inverted and uninverted). With this established, we argue that, given a fuller understanding of the syntax of Scots arising from previous microcomparative work on other aspects of its negation system, it is possible to build on Yang's productivity-based analysis in a way that can account for the richer pattern of data that the atlas has uncovered. In doing so, we stake out certain commitments to how the Tolerance Principle works in situations of morphosyntactic variability, extending the application of this set of ideas in a richer microcomparative domain.

2. Redrawing the Empirical Picture

This section introduces the Scots Syntax Atlas and shows how it provides a new perspective on the distribution of the *amn't* gap in Scots. With that in place, we show that the geographical distribution of inversion with negation across Scotland broadly follows Bresnan's (2001) description, but that the geographical distribution of two distinct negated forms of first person present *be* (*amn't* and *amnae*) does not.

2.1. The atlas

The Scots Syntax Atlas (SCOSYA, Smith et al. 2019) is a fully searchable online atlas, mapping morphosyntactic variation across Scots varieties. The atlas contains data gathered from over 140 locations across Scotland (Figure 1), with four speakers – two aged 18–25 and

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

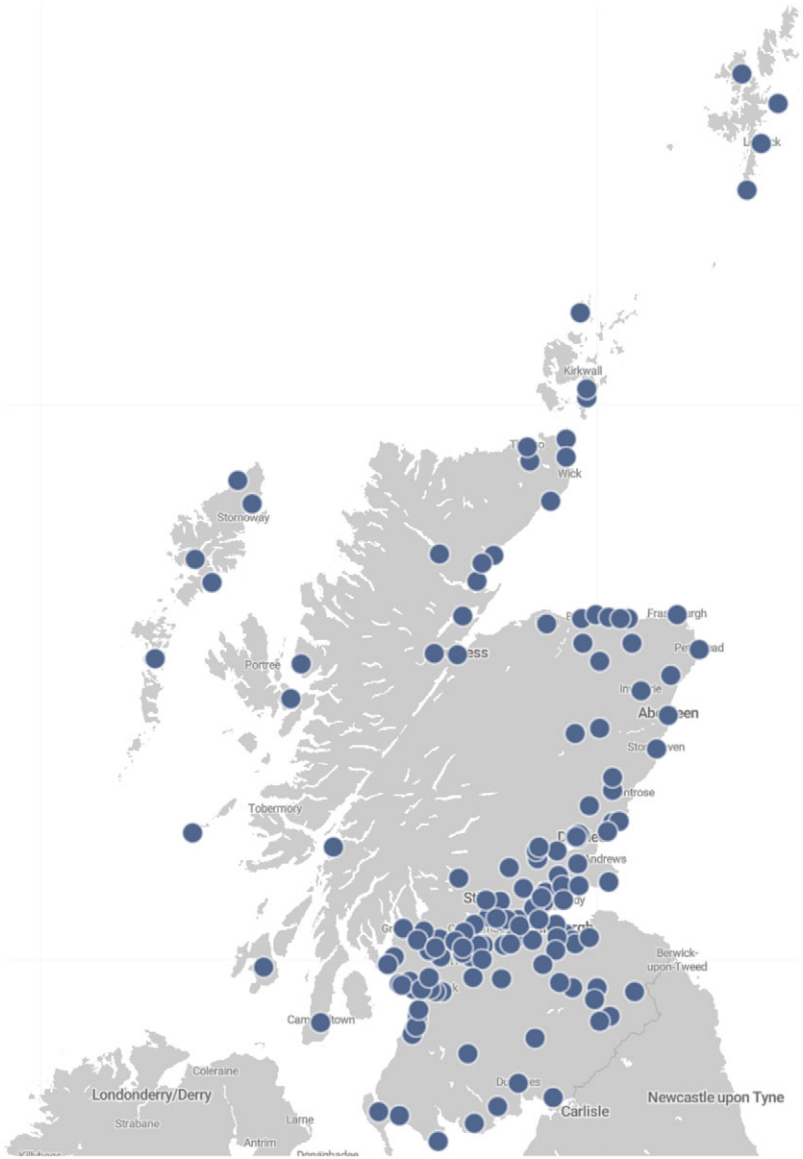


Figure 1. Locations around Scotland where data have been collected in the Scots Syntax Atlas. Each dot represents one community.

two aged 65+ – per location. Participants were recruited following standard sociolinguistic criteria for identifying speakers local to the community under study (e.g. Feagin 2013), and data collection was conducted by field-workers who were also members of the community in order to mitigate effects of the Observer's Paradox (Labov 1972).

Data were collected from each pair of participants in two ways. First, data were collected from sociolinguistic interviews (Labov 1984: 32), totalling ca. 275 hours of conversation. The interview recordings are text-to-sound transcribed using the open

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

source software TranscriberAG (https://transag.sourceforge.net), creating a rich corpus of searchable spoken data from Scots varieties. Second, participants completed an acceptability judgment task on 200+ examples of morphosyntactic phenomena, administered following the 'interview method' for syntactic data collection (Barbiers & Bennis 2007). Participants judged each example on a Likert scale from 1 to 5 (e.g. Schütze 2016). Each point on the scale was labelled, with 5 described as 'I would definitely say that'. At the other end of the scale, 1 was described as 'I would never say that'. Each participant was asked to provide around 200 judgments, with a total of over 110,000 judgments overall. These data were fed into the online atlas and allowed us to visualise patterns of acceptability across geographic space (the different locations) and time (the different age groups). The full dataset and more details about the project are available through the online atlas at https://scotssyntaxatlas.ac.uk.

In what follows, we will focus on the judgment data for the following sentences in the questionnaire: one testing *amn't* in an inversion environment (specifically a tag question[1]), one testing *amn't* in a declarative, and one testing *amnae* in a declarative.[2]

(6) I'm coming with you, *amn't* I?

(7) I *amn't* ready yet!

(8) I *amnae* ready!

We use these data to investigate three empirical issues. In Section 2.2, we consider whether the inversion asymmetry with respect to *amn't* holds more broadly across our larger dataset. In Section 2.3, we examine the relationship between *amn't* and *amnae* in declaratives, assessing the predictions of the proposal in Bresnan (2001) regarding the inversion asymmetry. In Section 2.4, we take a closer look at declarative *amn't* and identify some major regional differences in its acceptance.

Before moving onto the results, we should establish some methodological caveats regarding the data that we have gathered using these examples. One potential issue with the examples in (7)–(8) is that they both involve the use of negative contraction (*I amn't*/*amnae*) rather than auxiliary contraction (*I'm not*/*no*) in a context which strongly favours the use of the auxiliary contraction – namely, present tense *be* and a pronominal subject. The usage preference for auxiliary contraction in these contexts is well-documented for a number of British varieties, including varieties of Scots (Tagliamonte & Smith 2002), and this effect is visible in the SCOSYA corpus, which contains not a single example of *amn't* and only one example of *amnae*, given in (9); this is out of 1,800 potential contexts where negation and *am* occurred together, in the whole 3,000,000-word corpus.

(9) I'm no ticklish pal, genuinely *amnae.* (Lothian, Younger)

[1] We do not have data on whether there are any differences in judgments for different inversion environments. However, see Thoms et al. (2023) on differences between negative morphemes in different classes of inversion environments.

[2] There were some differences in the orthographic representations used for the example sentences for different areas; for example, *-nae* is represented as *-na* in questionnaires used in Dundee and the Northeast, in line with previously established orthographic representations of how these are pronounced. We will flag these differences up where relevant, but as far as we know, they do not impact the analyses we develop.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

We should stress that low frequency in use does not necessarily entail low acceptance or ungrammaticality; for example, *you aren't* is also vanishingly rare in our spoken corpus (there are only a handful of examples in the SCOSYA corpus), due to a preference for *you're not*, yet it seems unlikely to us that it is ungrammatical in any varieties in the United Kingdom. Nevertheless, it is possible that when naive informants are asked to rate *I amn't*/*amnae* (or indeed, *you aren't*) for acceptability, they may occasionally give it a low rating because of this preference, despite the best efforts of field-workers to explain the task – even if the relevant form is grammatical for the speaker and may show up (albeit rarely) in their usage. This might, in turn, lead to depression of the rate of acceptance of examples such as (7)–(8) – in particular, relative to examples like (6). We will return to this caveat at a few points below. Also, we note that there is a general role for age in the acceptability of *amn't*/*amnae* in the SCOSYA data, with younger speakers accepting both inverted and declarative *amn't* and declarative *amnae* at a higher rate than older speakers. A small benchmarking study investigating differences in acceptability between the age groups for seven known Scots-wide features showed an average acceptance rate of 72.2% for the older speakers and 90.2% for the younger speakers. It may be that this difference between the age groups is due to younger speakers having greater perceptual linguistic awareness than older speakers, evidenced in their accuracy in metalinguistic tasks (Drager 2011, Carrera-Sabaté 2014, Lawrence 2017, though see Jamieson 2020 for the opposite). It may also be the case that there is change in progress with *amn't*, moving towards greater acceptability. As instances of *amn't* and *amnae* are so rare in the SCOSYA corpus, we cannot use any additional existing data to test this hypothesis further. We therefore acknowledge that change may be a possibility, but we set it aside for future sociolinguistic investigation.

We will now present the geographical distribution of the *amn't* and *amnae* data from SCOSYA and highlight where it supports or provides counterarguments to Bresnan and Yang's proposals. Throughout, we report the acceptance rate: how many participants rated the example 4 or 5 in the judgment task. We use this as an operationalisation of acceptability, following, for example, Zanuttini et al. (2018), who use the same cut-off for similar types of acceptability judgment results in the Yale Grammatical Diversity Project, and Thoms et al. (2019), Jamieson et al. (to appear), who used these cutoffs when investigating other aspects of the SCOSYA data. We also report the mode throughout as a quick and clear way to establish how the majority of participants responded to a particular example or in a particular area.

2.2. *Geographical distribution of* amn't *and the inversion asymmetry*

To begin with, we show that there is indeed a strong asymmetry between declaratives and inversions with respect to the acceptability of *amn't* across Scots dialects. This broadly confirms one of the key empirical claims in Bresnan (2001). We can see this with maps representing the scores for the examples in 6 and 7. Figures 2 and 3 show acceptability maps, in which dark spots represent locations where the example was given a score of 4 or 5 by two or more speakers.

We can see from simply looking at the maps that *amn't* is accepted much more widely, and indeed much more uniformly, in inversions than in declaratives.[3] This can also be seen

[3] This pattern of wide acceptance for tags across the whole of Scotland is mirrored in a much larger number of crowdsourced judgments collected in the Speak for Yersel project, which can be viewed at http://www.speakforyersel.ac.uk.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

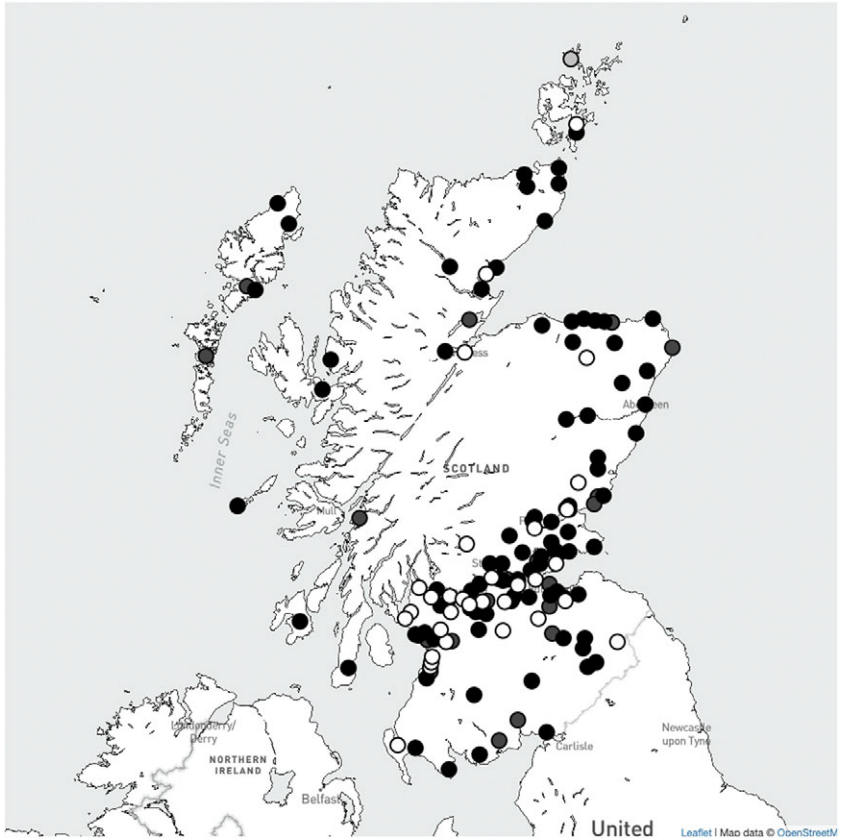


Figure 2. *I'm coming with you, amn't I?*, accepted by ≥ 2 speakers.

in the descriptive statistics: the inversion example in (6) had a mode of 5, with the individual-level acceptance rate at 60%. The numbers for declarative *amn't* are much lower: Example (7) had a mode of 1 and an acceptance rate of 34%.

Clearly, then, the example testing *amn't* in a declarative was less well accepted in the SCOSYA data than the example testing *amn't* in inversions, potentially indicating a difference in terms of the grammatical status of *amn't* in these two environments, in line with what was proposed by Bresnan 2001. While it is possible that some portion of the difference in acceptability between the two examples could be attributable to the preference for auxiliary contraction over negative contraction in declarative contexts, it seems unlikely that this would account for the full extent of the difference in scores. An additional consideration, which we expand on more in Section 2.4, is that whereas the scores for (6) in the map in Figure 2 are high more or less uniformly across the country, the scores for (7) in Figure 3 vary much more by region. This would be unexpected if the difference between inversions and declarative was due solely to methodological issues.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

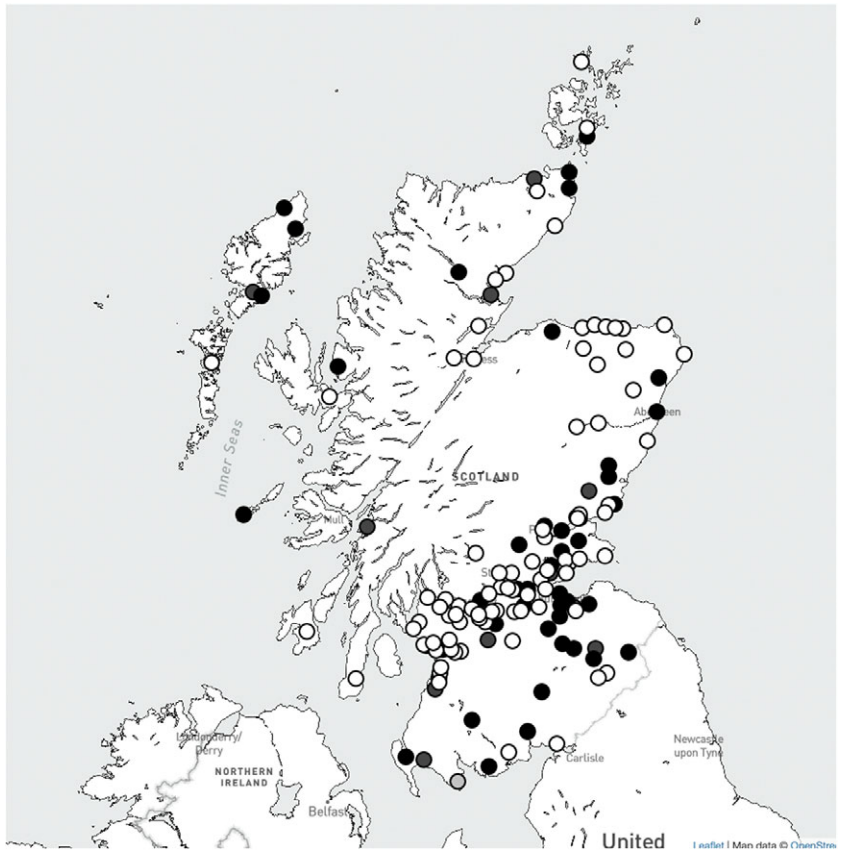


Figure 3. *I amn't ready yet*, accepted by ≥ 2 speakers.

2.3. *Geographical correlation between* amn't *and* amnae

We have seen that Bresnan's (2001) proposed inversion asymmetry holds in the SCOSYA data. Now let us consider the second major claim of her argument: that the *amn't* asymmetry is due to the fact that declarative *amnae* effectively blocks declarative *amn't.* In her account, then, the two forms should be in complementary distribution, at least in regions where *-nae* forms are used productively.

To assess this claim, we provide map data for declarative *amnae* in Figure 4, which can be compared with declarative *amn't* in Figure 3. In addition, to contextualise the results for *amnae*, we provide map data from another example of *-nae* negation: *havenae*, as tested in the sentence in (10).[4]

(10) I *havenae* been there before.

[4] We present data for *havenae* here; the same pattern holds across the SCOSYA data for other *-nae* examples, such as *isnae*, *wasnae* and *didnae.*

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

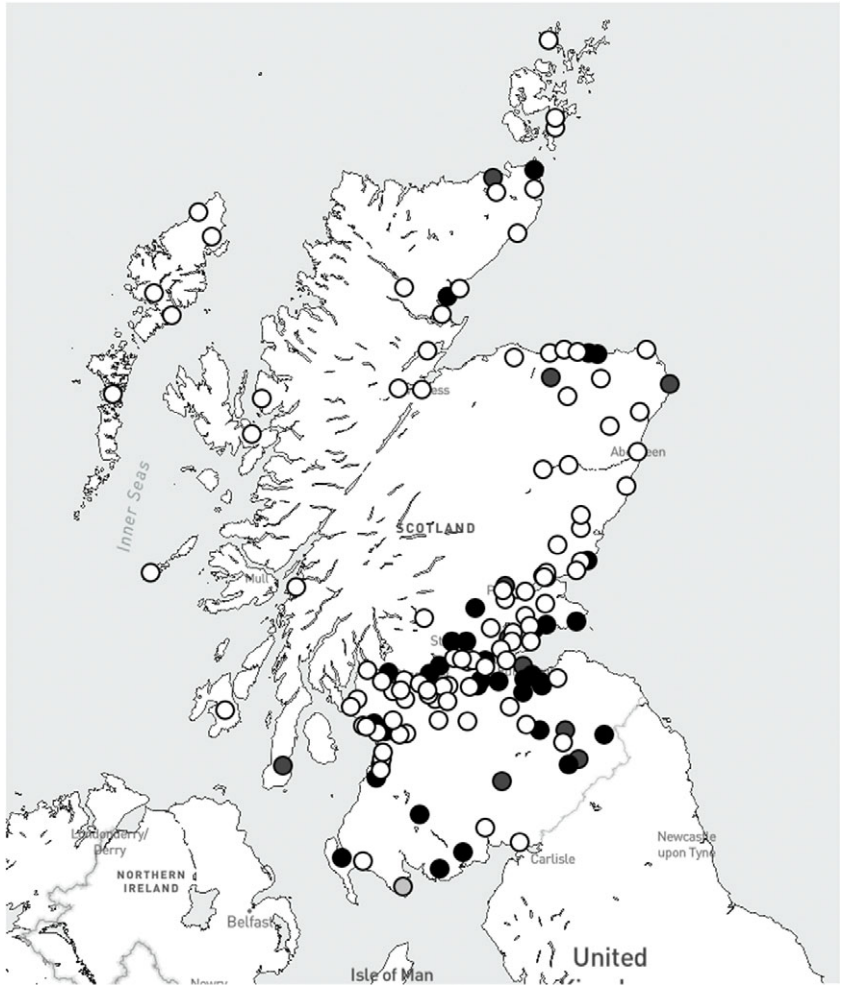


Figure 4. *I amnae ready*, accepted by ≥ 2 speakers.

It is easy to see from Figure 5 that *havenae* is generally accepted across Scotland, with the exception of the Highlands and the Western Isles, which exhibit standard English forms (e.g. Shuken 1984).

We can see in Figure 4 that *amnae* is also, unsurprisingly, rejected in these areas. However, what is surprising (in light of Bresnan 2001's claim) is that acceptance of *amnae* is generally low outside of the Highlands, too, especially when compared to other *-nae* forms, like *havenae.* This can also be seen in the descriptive statistics: the example in (10) with *havenae* has a mode of 5 and an acceptance rate of 65%, whereas (8) (*I amnae ready!*) has a mode of 1 and an acceptance rate of 32% – very similar to what we saw for declarative *amn't.* In fact, the maps for *amnae* and declarative *amn't* are quite similar once the Highlands and Western Isles are ignored, with patches of lower acceptance in the Northeast and various places in the central part of the country.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

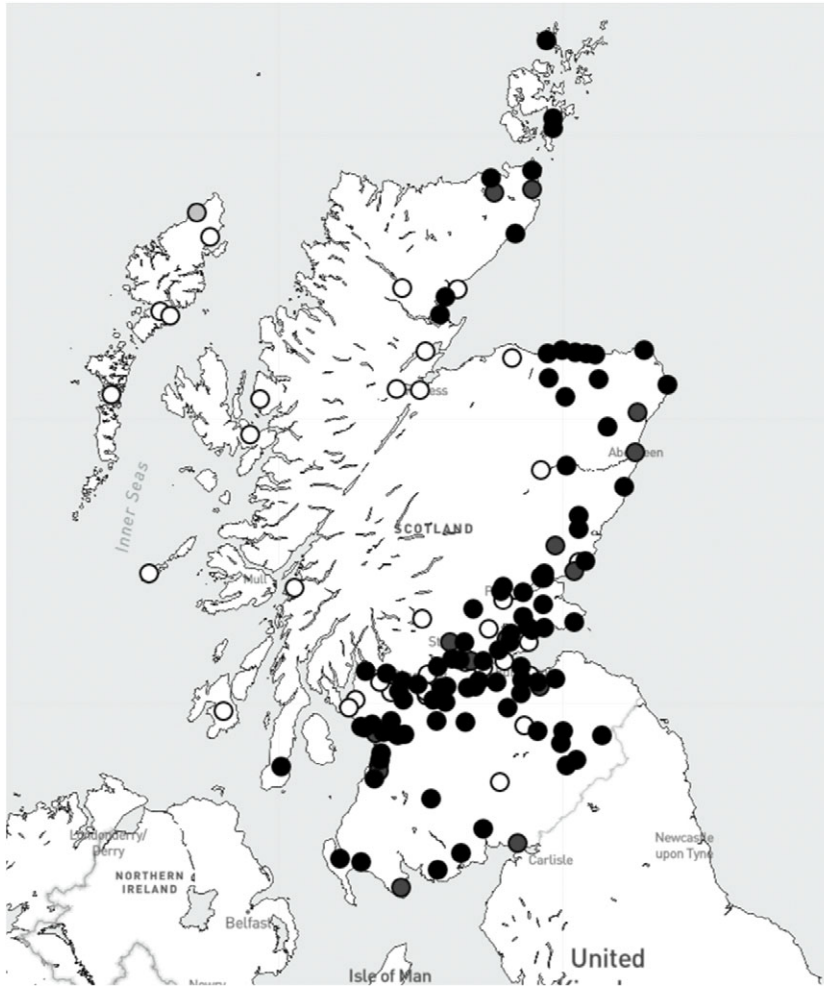


Figure 5. *I havenae been there before*, accepted by ≥ 2 speakers.

Looking at the data at the individual level,[5] it is remarkable that scores for *amnae* and declarative *amn't* are, in fact, moderately positively correlated (R = 0.52)– more so than *amnae* with other *-nae* forms such as *havenae* (R = 0.34). This is the opposite of what is expected on Bresnan (2001)'s account, and we believe it tells us that that account cannot be right. We suggest that any account of the *amn't* gap will have to also explain the fact that it tends to pattern together with *amnae* across Scots varieties, while still allowing them to come apart in some situations.

[5] In calculating these statistics, we excluded all locations from the Western Isles and the Highlands, as demarcated in Figure 6, since *-nae* is not used there in general (as noted above).

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

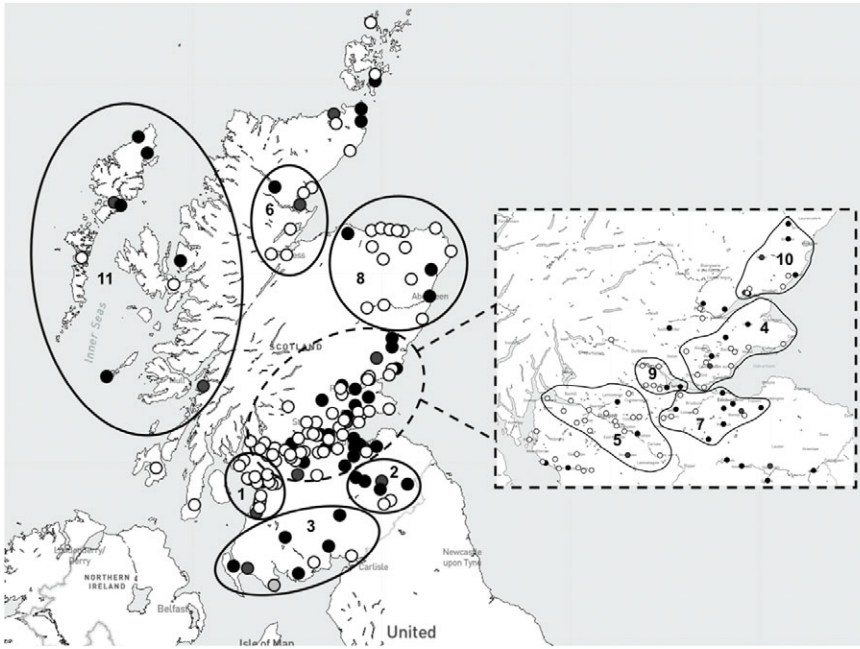


Figure 6. Labelled dialect regions of Scotland, corresponding to the regions in Table 1.

2.4. *Geographical distribution of acceptability of declarative* amn't

We have seen that the SCOSYA data do not support Bresnan's (2001) claim that declarative *amn't* is blocked by declarative *amnae*; instead, the presence of the two features is moderately positively correlated. We also saw in Figure 3 that there is regional variation in the patterns of acceptability of declarative *amn't.* We explore that further below, showing that although it is not predicted by Bresnan's (2001) account, Yang's (2016, 2017) approach opens up the possibility of a different explanation for the distribution of the forms. As noted in Section 2, the SCOSYA data come from over 140 research sites across Scotland. In the atlas, these are grouped into broader geographic regions based on areas demarcated in traditional dialectological work (e.g. Grant 1931, Aitken 1984, Johnston 1997, Millar 2007), resulting in 15 broad regions. Table 1 provides an overview of the descriptive statistics for declarative *amn't* for 11 of these dialect regions in the SCOSYA dataset.[6] The regions are labelled in Figure 6, which can be cross-referenced with Table 1.

The descriptive statistics clearly show regional variation in the scores for declarative *amn't.* Scores are particularly high in Lothian, the Western Isles and Tayside & Angus and particularly low in Ayrshire, Glasgow, the Northeast and Stirling & Falkirk. For the remaining regions (Borders, Dumfries, Fife and the Highlands), the picture is more mixed: in most of these regions, the mode was 1, indicating high levels of rejection. However, the number of participants accepting it was between 30%–50%, suggesting a more complex picture.[7]

[6] We have removed data from any dialect area where there are ≤ 20 observations to generalise over. This means we do not analyse or discuss data from Caithness, Kinross, Orkney or Shetland.

[7] It is these regions in which a change in progress with the acceptability of *amn't* seems more likely. However, as noted in Section 2.1, we do not have appropriate data to address the question of change in apparent time, and so we set this aside for future work.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Table 1. Declarative *amn't* across Scotland

		Accept %	*n*	Mode
1	**Ayrshire**	20	11/55	1
2	**Borders**	48	12/25	4
3	**Dumfries**	39	14/36	1
4	**Fife**	32	14/44	1
5	**Glasgow**	19	13/70	1
6	**Highlands**	35	8/23	1
7	**Lothian**	52	26/50	5
8	**Northeast**	20	13/64	1
9	**Stirling & Falkirk**	21	6/28	1
10	**Tayside & Angus**	55	18/33	5
11	**Western Isles**	53	23/43	5

There are two striking observations we can make from this data. The first is the high scores for declarative *amn't* in the Western Isles. The second is the regional variation on the east coast of Scotland. Acceptance is high in Lothian in the south but lower to the immediate north in Fife. Then it rises again as we go yet farther north to Tayside and Angus before plunging in the Northeast.

We can summarise the generalisations about the *amn't* gap and the inversion asymmetry that we have extracted from SCOSYA as follows:

(11) (a) There is indeed an asymmetry between declaratives and inversions with respect to acceptance of *amn't*, as noted by Bresnan (2001);
(b) There is, however, a positive correlation between acceptance of *amn't* and its counterpart *amnae* across Scotland, counter to the predictions of an account like Bresnan (2001)'s;
(c) There is regional variation with respect to acceptance of declarative *amn't*: it varies by dialect region across the rest of Scotland.

In what follows, we account for these generalisations by leveraging a view of the syntax of negation, proposed in Thoms et al. (2023), which takes negation in inversion constructions to be distinct from negation in declaratives. We combine this with a slight extension to Yang's Tolerance Principle to deal with cases of variation in exponence, and we show how the two together provide a way of understanding why *amn't* distributes the way it does in Scots (and potentially in varieties of English more generally).

3. On The Morphosyntax of Scots Negation

Speakers of Scots use two main forms of contracted negation: *-n't* and *-nae.* The use of *-nae* is illustrated in the following examples from the SCOSYA corpus.[8]

[8] The phonetic realisation of *-nae* differs across Scotland. In general, it is pronounced as [nə] in the Northeast, Tayside and the northern Isles, and it is often represented as *-na* orthographically. It is realised as [ne] elsewhere,

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

(12) (a) If you changed your mind, your wall *wouldnae* be bumpy forever. (Ayrshire, younger)
(b) You felt as though you *couldnae* get a breath out. (Caithness, older)
(c) I *didnae* ken at this point it was just a mouse. (Borders, younger)
(d) Would you have got a tree if I *hadnae* forced you to get a tree? (Northeast, younger)

Varieties in the Western Isles and Highlands use only standard contracted negation *-n't* (Thoms et al. 2023). The rest of Scotland is characterised by use of both *-n't* and *-nae* across communities and within individual speakers, with this variability subject to a number of sociolinguistic factors (e.g. Smith 2000a). This alternation between the forms can be seen in the following sentences from the SCOSYA corpus, where speakers switch from one form to the other within a single utterance.

(13) (a) *Didnae* have very much, you *didn't* have a fancy house, you certainly *didnae* have a big fancy house, but I was just content. (Ayrshire, older)
(b) I *couldn't* text anybody, I *couldnae* get on internet or that. (Stirling & Falkirk, younger)

As noted in Thoms et al. (2023), *-n't* and *-nae* show almost identical syntactic distributions. The examples in (12) and (13) all demonstrate cases of *-nae* attaching to finite auxiliaries in T, and all are contexts where *-n't*-based forms may occur as well. The examples in (14) (from Thoms et al. 2023) show that there are many other contexts that block contracted negation: note that both *-n't* and *-nae* are prohibited from occurring.

(14) (a) *You should really {*-n't* / *nae*} reply
(b) *You {should *N'T* / should *NAE*} reply.
(c) *Are you {*-n't* / *nae*} coming?
(d) *You should {*haven't* / *havenae*} bothered to reply.
(e) *We {*expectedn't* /*expectnae*} to like it.
(f) *We expected {*ton't* / *tonae*} like it.
(g) *We seem to {*haven't* / *havenae*} been given the right information.

The judgments reported in (14) hold across Scots varieties (modulo the absence of *-nae* forms in some dialects, as already mentioned), and as far as we know, the *-n't* judgments hold across all varieties of English.

However, there are differences, both between *-n't* and *-nae*, and between variants of *-nae* across different varieties of Scots.[9] A major difference between *-n't* and *-nae*, which we already discussed in the Introduction, and which is further demonstrated in (15) and (16), is

including Glasgow and Edinburgh, and is usually written *-nae.* Despite these different phonetic realisations, the morphosyntax of these forms is the same across these varieties. We use the *-nae* orthographic form by default in the rest of the paper since it reflects more transparently the pronunciation of the most commonly used form.

[9] There are few (if any) differences between varieties (in Scotland and elsewhere) with respect to the syntactic properties of *-n't*, as far as we know, although there are differences in the inventories of *-n't* forms, the variation with *amn't* being one prominent case.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

the asymmetry with inversion: auxiliaries negated with *-n't* but not *-nae* can invert with the subject in questions.

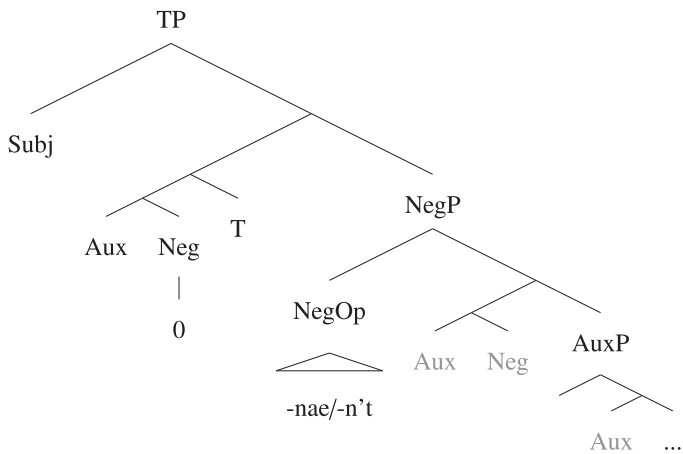
(15) (a) Why couldn't you read it
(b) Haven't you read it?
(c) They should read it, shouldn't they?

(16) (a) *Why *couldnae* you read it
(b) **Havenae* you read it?
(c) *They should read it, *shouldnae* they?

To account for this, Thoms et al. (2023) proposed an analysis of *-n't* and *-nae* as two distinct realisations of a single element, which they call NegOp. In declaratives, this occurs in the specifier of a NegP projection directly below finite T. NegOp attaches leftwards onto the adjacent finite auxiliary in T.

(17) provides a tree diagram for a simple example with the perfect auxiliary, and the associated spell-out rule is given in (18).[10] Leftward attachment of NegOp onto the finite auxiliary would produce the context for a variable rule in (18), which realises NegOp.

(17)



(18) #AUX$_{\text{NEG–OP}}$# → #AUX[nt]# *or* #AUX[ne]# / #T$_{\text{FIN}}$__#

This analysis captures the highly selective nature of *-n't*/*-nae* by stating a realisation rule that is specific to contexts where the NegOp has attached to the finite auxiliary. By positing the disjunctive realisation rule in (18), Thoms et al. (2023) analysed the variation between *-n't*

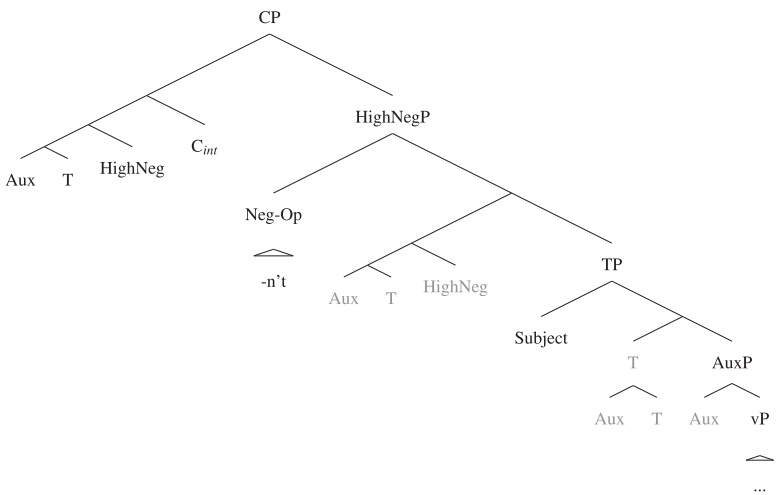
[10] The rule in (18) is an adjusted version of the one in Thoms et al. (2023), which is given in (i).
(i) NEG–OP → [nt] *or* [ne] / #T[fin]—#
The rule in (i) excludes the auxiliary from both the left and the right context, and so such a rule would have nothing to say about cases of stem allomorphy such as *mustn't*, where the form of *-n't* is the same as it usually is, but the stem of the auxiliary undergoes some change (the [t] is dropped; see Zwicky & Pullum 1983 and below). This distinction is important for the Tolerance calculation which is to follow, since cases of stem allomorphy must be counted as irregularities of the *-n't* attachment rule, so we adopt the adjusted rule in (18) in what follows. Thanks to Craig Sailor (personal communication) for pointing out this issue for us.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

and *-nae* as an instance of sociolinguistically conditioned allomorphy – of a kind that is familiar from a long tradition of work in variationist sociolinguistics (see Smith 2000b on Scots negation, in particular).

Thoms et al. (2023) treated the 'inverted' negation as an instance of NegOp appearing in the specifier of a distinct projection immediately below the C projection in the clause. They called this projection HighNegP. NegOp attaches onto the preceding raised auxiliary in C, much as its lower instantiation does. (19) illustrates a case like (15b), *haven't you read it?*: the auxiliary raises through T and other heads on the verbal spine to C (which bears an interrogative feature [int]), and then NegOp cliticises to C post-syntactically.

(19)



The differences in the form of negation in the different contexts – *-n't* or *-nae* in declaratives, only *-n't* in inversions – is then captured in terms of different spell-out rules that are sensitive to the features of their local contexts. The absence of *-nae* in most inversion contexts is captured by stating a spell-out rule for NegOp in the context of C_{INT}, which only realises NegOp as *-n't.* We extend our realisation rules then to the following:

(20) (a) #AUX$_{NEG–OP}$# → #AUX[nt]# *or* #AUX[ne]# / #T_{FIN}__#
(b) #AUX$_{NEG–OP}$# → #AUX[nt]# / #C_{INT}__#

We then have two realization rules for NegOp, each dependent on syntactic context.[11] The differences between *-nae* and *-n't*, on this account, ultimately rest in the inventory of spell-out rules rather than in syntactic parameters or features.

A prediction of this way of analysing contracted negation is that we should expect to see cases where the morphological paradigm for negated auxiliaries may differ depending on the syntactic context in which the negative occurs, since the distinct spell-out rules need not have

[11] Thoms et al. (2023) showed how the form of negation in negative imperatives is captured by a third realisation rule, while the form of negation in uncontracted contexts is the elsewhere case.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

the same inventories of forms. And, indeed, this is what we see. One case, mentioned in the Introduction, is the use of *aren't* to 'plug' the gap left by *amn't*, but only in inversion contexts.

(21) (a) I'm coming with you, aren't I?
(b) *I aren't coming with you.

This approach seems preferable to Bresnan's (2001) idea that the emergence of *aren't* in inversion contexts is the result of an optimality-driven retreat to the unmarked form. Bresnan's (2001) account predicts that speakers who accept *amn't* in inversion contexts will not accept (21a), but this is not the case, as at least some Scots speakers who accept *amn't* also accept and produce (21a) (see Embick & Marantz 2008 and Frampton 2001 for further problems with Bresnan's 2001 analysis).

In addition, Thoms et al. (2023) described a number of cases in varieties of Scots where the form of negation in certain inversion contexts is distinct from that of negation in T in declaratives. (22) and (23) illustrate two cases from Shetland Scots and Glasgow Scots, respectively (for more details, see Thoms et al. 2023 and Jamieson 2020).

(22) Shetland Scots
(a) You have a standing ticket, *do'n* you?
(b) *You *do'n* have a standing ticket.

(23) Glasgow Scots
(a) *Wint* it just a lovely day?
(b) *It *wint* a lovely day. ('it wasn't a lovely day')

Such facts are a challenge for the standard analysis where *-n't* in inversions is always just a pied-piped form which attaches to the auxiliary in T prior to auxiliary raising; however, if the forms of negation are allomorphs conditioned by adjacent syntactic features, the patterns fall into place.

4. The Limits of Productivity: The Tolerance Principle

Although the analysis of Scots negation just outlined provides a foundation for understanding how *-nae* could influence *amn't* and how an asymmetry between inversion and non-inversion contexts could come about, we are still in need of some explanation for why we see the variation we see in the acceptability of *amn't.* The second ingredient of our analysis of the *amn't* gap is the theory of productivity in Yang (2016) and its extension to the *amn't* gap in Yang (2017). In this section, we briefly outline the core components of Yang's approach to paradigm gaps and its application to *amn't.*[12]

The Tolerance Principle of Yang (2016) is a theory of what patterns in observed linguistic data will result in learners developing productive – generalising – rules. The underlying intuition, which is, as Yang remarks, a quite widely held one, is that rules only become (or remain) productive if they are applicable to a sufficiently large proportion of the

[12] See Belth et al. (2021), Liang et al. (2022) and Kodner (2023) for recent work that has put the Tolerance Principle to use in explaining acquisitional and diachronic linguistic data. For criticisms, see Goldberg (2019).

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Table 2. Auxiliaries and their negative forms in American English (adapted from Yang 2017: 7)

Auxiliary	Negative form	International Phonetic Alphabet (IPA) for negative form
am	**ain't**	**[ent]**
are	**aren't**	**[arnt]**
are	**ain't**	**[ent]**
can	**can't**	**[kænt]**
could	couldn't	[kudn̩t]
did	didn't	[dɪdn̩t]
do	**don't**	**[dont]**
does	doesn't	[dʌzn̩t]
had	hadn't	[hædn̩t]
has	hasn't	[hæzn̩t]
has	**ain't**	**[ent]**
have	haven't	[hævn̩t]
have	**ain't**	**[ent]**
is	isn't	[ɪzn̩t]
is	**ain't**	**[ent]**
must	**mustn't**	**[mʌsn̩t]**
need	needn't	[nidn̩t]
should	shouldn't	[ʃudn̩t]
was	wasn't	[wʌzn̩t]
were	**weren't**	**[wɹnt]**
will	**won't**	**[wont]**
would	wouldn't	[wudn̩t]

Note: Irregular forms in bold.

candidates they could apply to. That is, there is a limit to the number of exceptions a rule can tolerate.

The Tolerance Principle gives a specific quantification of the number of exceptions – the threshold θ_N – that a rule potentially applying to N items can 'tolerate' while being retained as a productive rule (for a detailed explanation of how this formula is arrived at, see Yang 2016, Chapter 3, especially pp. 60–66):

(24) Tolerance Principle:
Suppose a rule **R** is applicable to ***N*** items in a learner's vocabulary, of which **e** are exceptions that do not follow **R**. The sufficient and necessary condition for the productivity of **R** is

$$e \leq \theta_{N:} \text{ where } \theta_{N:} = [N/lnN]$$

This principle thus predicts whether or not a potential rule will be productive, on the basis of two numbers: N (the number of items in the learner's lexicon to which the rule is applicable) and e (the number of items in the learner's lexicon that are exceptions). In Yang (2017), it is

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Table 3. Auxiliaries and their negated *-n't* forms in Lothian Scots

Auxiliary	Negative form	IPA for negative form
am	amn't	[æmn̩t]
are	aren't	[arn̩t]
can	**can't**	**[kænt]**
could	couldn't	[kudn̩t]
did	didn't	[dɪdn̩t]
do	**don't**	**[dont]**
does	doesn't	[dʌzn̩t]
had	hadn't	[hædn̩t]
has	hasn't	[hæzn̩t]
have	haven't	[hævn̩t]
is	isn't	[ɪzn̩t]
must	**mustn't**	**[mʌsn̩t]**
need	needn't	[nidn̩t]
should	shouldn't	[ʃudn̩t]
was	wasn't	[wʌzn̩t]
were	weren't	[wɜrn̩t]
will	**won't**	**[wont]**
would	wouldn't	[wudn̩t]

Note: Irregular forms in bold.

Table 4. Auxiliaries and their *-nae*-negated forms in Lothian Scots

English auxiliary	IPA for Scots aux	*-nae* form	IPA for *-nae* form
am	[am]	amnae	[amne]
are	[ar]	arnae	[arne]
can	[kɪn] *or* [kan]	cannae	[kæne]
could	[kud]	couldnae	[kudne]
did	[dɪd]	didnae	[dɪdne]
do	[de]	dinnae	[dɪne]
do	[de]	**dae**	**[de:]**
does	[dɪz] *or* [dʌz]	doesnae	[dʌzne]
had	[had]	hadnae	[hædne]
has	[haz]	hasnae	[hæzne]
have	[hav]	havenae	[hævne]
is	[ɪz]	isnae	[ɪzne]
must	[must]	mustnae	[mʌstne]
should	[ʃud]	shouldnae	[ʃudne]
was	[wɔz] *or* [wɪz]	wasnae	[wʌzne]
were	[wur]	werenae	[wurne]
will	[wɪl]	willnae	[wɪlne]
would	[wud]	wouldnae	[wudne]

Note: Irregular forms in bold.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

argued that this approach to productivity can give an account of the dialectal distribution of the *amn't* gap, in the following way.

First, Yang proposes a general realisation rule for contracted negation in English that would plausibly be posited by learners, according to which a syllabic [n̩t] is attached to the finite auxiliary. (25) is an adjusted version of this rule (see footnote 10):

(25) $\#\text{AUX}_{\text{NEG–OP}}\# \rightarrow \#\text{AUX}[\text{n̩t}]\# \,/\, \#\text{T}_{\text{FIN}}\_\#$

This rule will produce forms such as *hasn't*, *couldn't* – and also *amn't.* If an acquirer has (25) as a productive rule, they will produce *amn't*, even if they have never heard this form in their input: a productive rule will simply 'fill in' the gap. Although children who have learned such a rule might at some stage overuse it, producing forms such as *willn't*, eventually they will learn the full range of irregular forms, and the irregulars will block these regularisations; that is, *won't* will block *willn't*, via the Elsewhere Principle.

Whether the rule in (25) actually does become productive will, by hypothesis, depend on the Tolerance calculation, and so the second step for Yang's account is to determine *N*: the number of items that the rule is applicable to. In this case, that is the number of finite auxiliaries that have a negative form sufficiently robustly attested in child-directed speech to plausibly be acquired. Surveying CHILDES and BNC data, and taking the minimum frequency for counting a form to be once per 1,000,000 words, Yang concludes that 18 finite auxiliaries have negated forms that are sufficiently frequent in the input. Table 2 lists the positive auxiliaries and the negative forms of these auxiliaries that are robustly attested in Yang's corpora, with his IPA transcriptions for the American English pronunciations of the negative forms. Given that $N = 18$, the formula in (24) tells us that the threshold for exceptions is six ($\theta_{18} = \frac{18}{ln(18)} = 6$). That is, for this rule to be productive, there can be at most six exceptions (auxiliaries with irregular contracted negative forms).

The third step, then, is to count the number of auxiliaries that actually do have irregular negated forms – that is, to calculate *e.* Based on the data in Table 2, Yang arrives at a count of 10 such forms in American English (bolded in the table).[13] Some of the irregulars are straightforwardly characterised as such (see the discussion in Zwicky & Pullum 1983): *won't* is suppletive, *don't* involves an unpredictable vowel change, *mustn't* involves an unpredictable [t] deletion, and *can't* is monosyllabic (whereas applying (25) would return a bisyllabic form). By similar reasoning, we would also count *weren't* and *aren't* as irregulars because they are also transcribed as monosyllabic in Yang's table, and nothing about the phonotactics of American English forces them to be monosyllabic.[14] The table also includes five entries for *ain't*, since it represents a robustly attested negated form of all of the finite forms of present tense *have* and *be* in Yang's corpora. Given that the number of exceptions to the

[13] The number would be higher if relic forms such as *shan't* were included (as they are in the discussion in Zwicky & Pullum 1983, who also make the point that the proportion of irregulars is quite high). But, as Yang notes, this form has fallen out of use and can scarcely be taken to be an influence on the acquisition of negation in any contemporary varieties. Yang's count does not include the use of *ain't* for *didn't*, which is attested in contemporary African American language varieties. We suspect it is underrepresented in these corpora because it is a fairly recent innovation, as observed by Fisher (2018). Adding this use of *ain't* to the count does not change any of the predictions discussed here.

[14] For some reason, Yang only counts *weren't* as an irregular in his final calculation. We assume this is an error of presentation.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

putative rule in (25) is above six – the threshold derived from the Tolerance Principle – the prediction is that (25) will not be acquired as a rule.

Irregular forms – those that are not the output of a productive rule – have to be learned individually. Such forms can be thought of as generated by a morpholexical rule (Yang 2002, 2017) keyed to the item or set of items. For this to happen, sufficient positive evidence is required (that is, frequent enough presence of the output form in the input to the learner). In the case at hand, this is how the attested negated forms must be acquired: both obviously irregular forms like *won't* and *don't*, but also forms like *didn't* and *couldn't.* However, in the absence of a productive rule, if there is not enough positive evidence in the input for an irregular form (and all attested forms are irregular by definition), we have the conditions for a paradigm gap. This is precisely what Yang argues is the explanation for the absence of *amn't* in American varieties: there is no productive rule to generate it, and it cannot be acquired from the input directly since it is not attested in the speech of American adults.[15]

5. The *amn't* Gap in Scots

The situation in Scots is different. As we saw in the atlas data in Section 2 above, there are many areas in Scotland where *amn't* is accepted. Thus, we might expect that even if there were no productive rule for contracted negation in Scots, this form could be acquired simply by learning it as an irregular form directly from the input, on a par with *don't.* It is possible that this is the case for some Scots speakers. However, recall from the discussion at the end of Section 2.1 that *amn't* is very rare in speech, even in Scotland. There was not a single example of *amn't* in the entire SCOSYA corpus of 3,000,000 words, and there was only one example of *amnae*, even though there were more than 1,800 contexts where one of them could have been used.[16] In consequence, we think that it is unlikely that direct acquisition of the *amn't* form could account for the whole range of data, given how rare it is in speech, and it would certainly leave unexplained the dialectal variation that we have documented.

Yang's account of the *amn't* gap makes strong, fine-grained predictions about dialectal variation: dialects are predicted to differ as to whether a gap may occur depending on the number of irregular forms that they include. For this reason alone, if these predictions are borne out empirically, it represents an advance on all existing accounts of the *amn't* gap, which typically have little to nothing to say about why dialects differ on this matter.[17]

[15] In fact, we are simplifying Yang's proposal here. Having suggested the possible rule in (25), he then goes on to propose that learners retreat from this rule by entertaining a narrower version that applies only to obstruent-final auxiliaries (Yang 2017: p. 222).

(i) NEG → [nt] \ #AUX$_{FIN}$[+obstruent]–NEG#

We adopt the simpler formulation here.

[16] Recall that this usage fact is arguably part of a broader tendency in more northern varieties of English, and particularly Scots (noted in the discussion in Section 2.1 above) for preference for the use of auxiliary contraction over negative contraction with present tense forms of be and a pronominal subject (*she's not*, as opposed to *she isn't*), as discussed by Tagliamonte & Smith (2002).

[17] For example, as mentioned briefly in the Introduction, Bresnan (2001) simply stipulates that United States varieties and Scots varieties differ with respect to whether they have stored an *amn't* form in their lexicons, in effect, an instance of lexical variation; the same is true for Hudson (2000).

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Our account is based on the three ingredients that we have established in the previous sections. The following is a summary:

- Ingredient 1: negation in inversion contexts does not have exactly the same syntactic components as negation in declaratives, and therefore, the two heads may be realised via distinct spell-out rules.
- Ingredient 2: the realisation rule for negation in declaratives in Scots is a variable rule that generates both *-n't* and *-nae* forms.
- Ingredient 3: for a given negative morpheme's realisation rule, the Tolerance Principle calculation will determine whether it is productive and therefore automatically generates *amn't.*

These ingredients conspire to derive the main results in the following manner. First, in Section 5.1, we show that, because negation in inversions can be the result of a distinct spell-out rule, the Tolerance calculation for negation in inversions can be different for the Tolerance calculation for negation in declaratives. In particular, this will be true when the inventories of negated auxiliaries in the two environments are distinct. Second, the use of a variable rule to realise negation in declaratives as either *-n't* or *-nae* will ensure that the two negatives are tied together when it comes to the Tolerance calculation. In particular, an increase in irregularity in one of the disjuncts of the rule will impact upon the rule's Tolerance calculation as a whole, with the consequence that irregularity in *-nae* will also impact upon the productivity of *-n't.* In Section 5.2, we will spell this out in more detail and show that it predicts both the correlation between *amn't* and *amnae* and the lower acceptability of declarative *amn't* in certain areas of Scotland.

5.1. *Accounting for* amn't *in inversion contexts*

In Section 3, we analysed negation in inversion versus declarative contexts as realisations of distinct morphosyntactic elements. That is, negation in declaratives is the realisation of a NegOP in Spec,NegP, which is just below T, and which cliticises to the finite auxiliary in T. Negation in inversion contexts, however, is the realisation of a NegOP in Spec,HighNegP, just below C in the left periphery. The realisation rules we posited above are repeated in (26):

(26) (a) #AUX$_{\textsc{neg–op}}$# → #AUX[nt]# *or* #AUX[ne]# / #T$_{\text{FIN}}$_#
(b) #AUX$_{\textsc{neg–op}}$# → #AUX[nt]# / #C$_{\text{INT}}$_#

Given that these are distinct realisation rules, it should be possible for each rule to be assessed separately with respect to Tolerance calculations. This provides the route to an account of the inversion asymmetry: if the inventories of negative auxiliary forms attested in declaratives and inversions are substantially different, then the outcome of the Tolerance calculations for (26a) and (26b) may differ, and with that come differences with respect to *amn't.*

Let us consider in detail how the Tolerance calculation would work out for (26b), the spell-out rule for *-n't* in inversion environments. Recall that the *N* for a given rule is the number of finite auxiliaries with reliably attested forms that correspond to an output of (26b), which would mean negated auxiliaries in the pre-subject position. Yang proposed the *N* for his general *-n't* attachment rule to be 18, since there were 18 finite auxiliaries that had reliably

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

attested forms in his COCA survey. We will assume that a similar *N* would apply for the rule in (26b).[18] This gives a threshold of six, as before.

What about the list of exceptions? In Scots, there are no *ain't* forms, and the reduced form of *don't*, [dɔ̃], does not occur in inversion contexts at all (Kaisse 1985). Unlike in American English, *aren't* and *weren't* are not stored as exceptions since they are realised as bisyllabic forms in Scots, as predicted by the rule. Ultimately, then, the list of exceptions only includes *can't*, *won't* and *don't*,[19] and, for some speakers, the use of *aren't* for *am* in *aren't I*? (See Section 3 above). This would amount to a minimum of four (maximally five) exceptions, below the threshold of six. In consequence, the theory predicts that (26b) will be productive, thus generating *amn't* in inversions. This is summarised below:

(27) Tolerance calculation for (26b) in, for example, Lothian Scots
  (a) Forms which it may apply to: up to 18 (as before)
  (b) $\theta_{18} = 6$
  (c) Exceptions: up to 5; *can't*, *won't*, *don't*, possibly *mustn't*, *aren't* for *am* in *aren't I*?
  (d) Productive?: yes

The calculation above will work out similarly for most varieties of Scots since the inventories of negated auxiliaries in inversion contexts are largely uniform.[20] The account, therefore, correctly predicts the fairly widespread acceptance of *amn't* in inversions in Scotland.

5.2. Amn't *in declaratives: the impact of the variable rule analysis*

What of *amn't* in declaratives? For this, we must consider the variable realisation rule in (26a). At first blush, one might expect that this rule is doomed to fail to become productive because any rule that realises negation as *-n't* would have the *-nae* forms as exceptions to contend with, and vice versa. However, it is essential to note that we are proposing that the different negative forms are generated by two distinct disjuncts of a single spell-out rule – a Labovian variable rule.

How can Yang's Tolerance calculations apply to such Labovian rules? There is no reason to expect a difference in how we should calculate *N* (the number of items to which the rule is applicable). The question is how we should calculate *e* (the number of exceptions). We propose that all that has to be done to make sure that the Tolerance calculation generalises to this case is to take seriously the idea that a variable rule is just that – a single rule. Hence, any realisation that cannot be generated by the rule counts as an exception.

[18] We acknowledge that the *N* might be lower than this for (20b), as it may be that certain negated auxiliaries (e.g. *needn't*, possibly also *mustn't*) are so infrequent in inversion contexts that they might not occur enough to impact upon the acquisition process. An *N* as low as 15 would still get the same threshold and thus the same results, so we do not dwell on the matter here.

[19] We might include *mustn't*, but *mustn't* is very high-register for most speakers of Scots, and it is rarely (if ever) used in questions.

[20] We did note in Section 2.4 that there are further forms of the negation in imperatives and in tags (and for non-information-seeking questions in general, see Thoms et al. 2023). However, each of these requires a separate realisation rule and hence a separate Tolerance Principle calculation.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

(28) Calculation of exceptions
The exception list for a rule R consists of all the input–output pairs for relevant inputs where the output cannot be generated by R.

That is, if the rule is not a variable rule, exceptions will be counted just as before. If there are multiple disjunctive parts (i.e. if the rule is variable), the exception list for the rule is the union of the exceptions to all the disjuncts. Thus, following the principle proposed in (28), *don't* and *can't* would constitute exceptions to (26a), and so would *daa* and *caa.* By contrast, *cannae*, *havenae* and other regular forms would not constitute an exception to this particular rule since they are generated by one of the disjuncts.

It follows that in Scots varieties where *-nae* 'contributes' more irregular forms, the entire rule in (26a) would become less likely to be productive because the irregular *-nae* forms would increase the number of exceptions. Crucially for the case at hand, this variable rule is the one that, if productive, generates *amn't.* That is, in a dialect where there are many irregular *-nae* forms, learners are less likely to retain a productive rule like (26a), and so they would be less likely to produce and accept *amn't.*

Let us demonstrate by considering two varieties: Lothian Scots and Northeastern Scots. First, consider the Lothian Scots forms in Tables 3 and 4, where again the irregular forms are in boldface. Following the principle in (28), the range of exceptions to (26a) in this variety is the union of the the irregular forms listed in these two tables. As shown in Table 3, the irregular *-n't* forms are *can't*, *don't*, *mustn't*, *won't* and the reduced *don't* form [dɔ̃]; recall that *aren't* and *weren't* are regular as pronounced in Scots, and *ain't* is absent altogether. In Table 4, we can see that there is only one irregular *-nae* form in Lothian Scots – *dae* for *don't* (and this only for a subset of speakers[21]). With this mix of forms, the rule in (26a) is below or at the Tolerance threshold, and so learners in that area ought to be likely to acquire (26a) as a productive rule.

(29) Applying (26a) to Lothian Scots
(a) Forms which it may apply to: 18
(b) Threshold: 6
(c) Exceptions: 5–6; *can't*, *won't*, *don't*, *mustn't*, [dɔ̃]; for some speakers *dae*
(d) Productive?: yes

As a result, the analysis predicts that *amn't* would commonly be accepted by speakers of this variety. This prediction is borne in our data: *amn't* was rated particularly highly in the Lothian region.

The calculation is quite different for Northeastern Scots, not because there are any differences in its *-n't* inventory compared to what we have just seen for Lothian Scots, but because there are more irregular *-nae* forms – namely, *caa* 'can't', *winna* 'won't'. This variety also has zero auxiliary forms for *have* and *do* when they are followed by negation (see

[21] Here and throughout, we base our claims about the distribution of irregular negative auxiliaries – in particular, the Scots-specific forms – on a combination of judgment data and corpus data. The judgment data are from informal follow-up consultations with speakers from the regions in question, as the SCOSYA questionnaire did not test for these forms systematically. The corpus data are from the SCOSYA corpus, broken up into subcorpora for the different regions. In the case of *dae* for *don't*, it is clear from the production data that this is a productive feature of some Lothian varieties, but by no means all of them.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Thoms et al. 2023), in addition to *daa* (the local variant of *dae*) 'don't'. This inventory puts the exception list for this variety well above the threshold, and so (26a) is predicted to typically fail to become productive here.

(30) Applying (26a) to Northeastern Scots
  (a) Forms which it may apply to: 18
  (b) Threshold: 6
  (c) Exceptions: 10 *can't*, *won't*, *don't*, *mustn't*, [dɔ̃], *daa*, *caa*, *winna*, ∅*–na* for *don't*, ∅*–na* for *haven't*
  (d) Productive?: no

Our analyses, therefore, predict that *amn't* should be widely rejected in the Northeastern Scots varieties – again, the right result.

Let us now consider how this account extends to other varieties of Scots. Recall that one of the major results that we sought to account for was the variation up the east coast, whereby declarative *amn't* is accepted very widely in Lothian, Tayside and Angus, very little in the Northeast, and to some degree in Fife. We have already shown that the account explains the substantial difference in scores between the Northeast and Lothian. The analysis extends readily to Tayside and Angus, which have the same inventory of negated auxiliary forms as Lothian (modulo some predictable vowel quality differences). Irregulars such as the *winna(e)* 'won't' and *caa* 'can't' that are present in the Northeast seem not to be a feature of these varieties, and the zero auxiliary forms are not used there at all. The Tolerance calculation is thus the same as for Lothian, and *amn't* is predicted to be productive in declaratives.

As for Fife, which occupies the peninsula between Lothian and Tayside & Angus, there are more irregular negative forms in the varieties spoken there, as *caa* and *winnae* are attested conservative features of the area, and *dae* is also widely used by young and old speakers. As a result, the number of irregulars in Fife is generally higher than in Lothian, Tayside and Angus, and almost as high among older or conservative varieties as in the Northeastern varieties. The picture with *amn't* in Fife, however, is quite mixed; in particular, there is an especially strong difference between the generations, as the acceptance rate among the old participants was particularly low, just 4.5% (mode and median 1), whereas the acceptance rate among younger participants was 50%. This can be compared with the situation in the other areas with generally low scores (Northeast, Glasgow and Ayrshire), where the rates of acceptance were much closer between the age groups (10% and 24%, 12% and 25%, and 19% and 19%, respectively). We suggest, then, that there is a true generational difference in the use of *amn't* in Fife, and we speculate that this could be related to change in the use of the irregular forms in this region. Although we cannot use corpus data to probe this analysis of the Fife data further, additional judgment data gathering in future work may be able to test it more rigorously. As it stands, we take this analysis to give us a plausible account of the Fife data, and thus a fairly complete understanding of the variation with *amn't* on the east coast.

Finally, let us comment on how the analysis handles the data from the rest of the country. Putting the varieties without *-nae* to one side, the account predicts two broad kinds of varieties: those with a larger number of exceptions to (26a), which should pattern with Northeastern Scots in disfavouring *amn't*, and those with fewer exceptions, which should pattern with Lothian Scots. Concerning the latter group, the analysis makes broadly the right predictions for varieties in Dumfries and the Borders, both of which have higher rates of acceptance of *amn't* (see Table 1 above). These varieties typically have the same inventory of

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

irregulars as Lothian Scots: although some speakers in these regions accept *daa*/*dae* (examples were attested in all areas), *caa* is largely unattested (except for in some varieties in the Borders), *winna(e)* is unattested, and the zero auxiliaries are completely unattested.

As for the remaining regions, the proposal's predictions are unclear in the cases of Stirling and Falkirk, and the Highlands (which encompasses Inverness and the area around it). In the case of Stirling and Falkirk, this is a somewhat unusual region: there are only seven locations in this grouping, and it is in a transitional area in central Scotland between several other, better-defined dialect regions. We generally have lower information about the dialects of this region compared to others (for instance, concerning the use of irregular auxiliaries[22]). We therefore tentatively put this region to one side, pending further empirical data collection.

Similar remarks apply to the Highlands region, which contains just seven locations (23 ratings in total for (7), *I amn't ready yet!*). Speakers in that region also gave *amn't* in inversions relatively low ratings. This suggests there is more to understand about the status of the data from there, and again, we put this region aside.

What about Glasgow and Ayrshire? The proposal seems to make the wrong prediction for Glasgow and Ayrshire, where there is little to no irregularity in the *-nae* paradigm, but *amn't* is rated quite low. However, the varieties of this region have at their disposal an additional form for the realisation of negated first-person singular present tense *be* – namely, *amurnae*:

(31) A: You're a liar
B: [amʌrne] 'I'm not!', lit. 'I'm aren't!'

Given the Elsewhere Principle, our account predicts that the availability of another, more specific form will result in blocking of the application of the general rule, and so *amurnae* will block application of (26a) to generate *amn't* for the relevant speakers, just like *won't* blocks generation of *willn't* for all other English speakers. Thus, even if (26a) is productive for speakers of these varieties, they are not expected to produce and accept *amn't* in declaratives if they have acquired *amurnae.* This accounts for the low acceptability of declarative *amn't* in this region. It is possible that *amurnae* is used beyond these regions – for instance, in parts of Stirling and Falkirk – in which case, this might explain some of the uncertainties surrounding what is going on in this region, but information on this feature is very sparse. We also leave this for future investigation.

The results of our analysis are summarised in Table 5, which shows that the predictions of our account are largely confirmed.

A final comment to make on the data from declarative *amn't* is the extent of interspeaker variation: in no single area was declarative *amn't* uniformly rejected or uniformly accepted, and in most areas there was a mix of responses. Thus, there was a degree of what we might call overacceptance, with some speakers accepting declarative *amn't* in some areas where our analysis predicts it should be absent (e.g. the Northeast). Similarly, there was a degree of what we might call underacceptance, where a number of speakers reject declarative *amn't* in areas where we would expect it to be accepted more uniformly (e.g. the Western Isles).

There may well be methodological reasons for this. It is, of course, reasonable to assume that some proportion of overacceptance and underacceptance can be attributed to the noisy

[22] We should note that for every region except Stirling and Falkirk and the Borders, we managed to consult with researchers or field-workers who grew up in that region (or had extensive research knowledge of it) and who had extensive training in linguistics.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Table 5. Our predictions regarding declarative *amn't* across Scotland

	Region	Median	Mode	Accepted	Exceptions	Productive?	Predicted?
1	Ayrshire	2	1	20%	5	No	Yes (other form)
2	Borders	3	4	48%	5–7	Probably	Yes
3	Dumfries	2	1	39%	5–6	Yes	Yes
4	Fife	2.5	1	32%	5–8	Variably	Yes (age split)
5	Glasgow	1	1	19%	5	Yes	Yes (other form)
6	Highlands	2	1	35%	5–6	Yes	Unclear
7	Lothian	4	5	52%	5–6	Yes	Yes
8	Northeast	1	1	20%	9–10	No	Yes
9	Stirling & Falkirk	2	1	21%	5–7?	Possibly	Unclear
10	Tayside & Angus	4	5	55%	5–6	Yes	Yes
11	Western Isles	4	5	53%	5	Yes	Yes

nature of large-scale acceptability judgment data gathering. However, previous work has shown (Jamieson et al. 2023) that the judgment data from the SCOSYA project largely lines up with patterns of production in the SCOSYA corpus – in particular, in the case of features that are not strongly stigmatised[23] and which are used commonly enough in the corpus for usage statistics to be observed.

We noted at the end of Section 2 that there might be some methodological issues with respect to gathering data on negative contraction examples specifically – speakers might rate negative contraction lower because of a preference for using auxiliary contraction *I'm no/not* – which would account for some amount of underacceptance of declarative *amn't.* This could potentially be a factor in explaining the fact that declarative *amn't* is rated lower than inversion *amn't* in general, even in regions where the two are predicted by our analysis to be more or less the same. For instance, in the Western Isles, where there are no *-nae* forms in the mix, declarative *amn't* is accepted 53% of the time, and inversion *amn't* 72% of the time.

In addition to methodological issues, our analysis makes room for a certain degree of individual-level variation, in particular, with respect to overacceptance of *amn't* in areas where the general *-n't*-attachment rules fail to become productive. Recall from the discussion in Section 4 that forms that may be the output of a productive rule in one variety may be learned as irregular forms in another where there are more exceptions and hence no productive rule. Indeed, on our account – following Yang – this is the situation for North American speakers learning the form *isn't* or *hasn't. Amn't* is likely to be extremely rare in the input to North American learners, so we would not expect it to be learned as an irregular in this way, but things could potentially be different in some Scots varieties. It is clear that *amn't* is rare in production – recall there was not a single clear example in the 3,000,000-word

[23] This seems to be the case with *amn't.* The lack of stigma with *amn't* in Scotland is perhaps in part due to the fact that it involves the use of the more standard form of negation rather than the more non-standard *-nae.*

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

SCOSYA corpus – and so usage of *amn't* is unlikely to be driving the whole set of dialect differences we have seen above, but it is possible that there will be individuals who have learned *amn't* on the basis of (scant) production data in neighbouring varieties, even in communities where the makeup of the negative auxiliary inventory is such that it would typically rule out a productive negative attachment rule.

Taken together, these factors give us a way to understand the variability in the *amn't* data. However, it is highly unlikely that these factors combine to explain the extensive and systematic regional differences that we have laid out above.

6. Conclusion

In this article, we have provided new data from SCOSYA that repaints the empirical picture of the distribution of the *amn't* gap, yielding a much richer landscape of microvariation, and one that challenges extant accounts. We have argued that these data can be accounted for on the basis of the analysis of Scots negation in Thoms et al. (2023), together with a new proposal for how the Tolerance Principle should be updated to apply also to Labovian variable rules. Thus, we hope to have both provided a new and much more adequate description of microvariation in Scots contracted negation and to have improved our theoretical understanding of the nature of productivity in linguistic rules.

Acknowledgments. We gratefully acknowledge the support of the Arts and Humanities Research Council for award No. AH/M005550/1. Material from this paper has been presented to audiences at Cambridge, Edinburgh, Konstanz, Yale, MIT, NYU and Stanford, and at Comparative Germanic Syntax Workshop (CGSW) 32 (Trondheim) and Generative Linguistics in the Old World (GLOW) 43 (Berlin/online). We thank those audiences for their feedback and input, as well as the anonymous reviewers for the *Journal of Linguistics* who gave us valuable feedback and advice. We also thank Brian Aitken, Derek Henderson, Laurel MacKenzie, Warren Maguire, Andrew Nevins, Craig Sailor, Devyani Sharma, Andrew Weir and Charles Yang.

References

Aitken, Jack. 1984. Scottish accents and dialects. In Peter Trudgill (ed.), *Language in the British Isles*, 94–114. Cambridge: Cambridge University Press.

Anderwald, Lieselotte. 2002. *Negation in non-standard British English: Gaps, regularizations and asymmetries*. London: Routledge.

Barbiers, Sjef & Hans Bennis. 2007. The syntactic atlas of the Dutch dialects. A discussion of choices in the SAND project. *Nordlyd.* 34. 53–72.

Belth, Caleb, Sarah Payne, Deniz Beser, Jordan Kodner & Charles Yang. 2021. The greedy and recursive search for morphological productivity. In *Proceedings of CogSci 2021*, https://arxiv.org/abs/2105.05790.

Bresnan, Joan. 2001. Explaining morphosyntactic competition. In Mark Baltin & Chris Collins (eds.), *Handbook of contemporary syntactic theory*, 1–44. Malden, MA: Blackwell.

Broadbent, Judith. 2009. The **amn't* gap: The view from West Yorkshire. *Journal of Linguistics.* 45(2). 251–284.

Brown, Keith. 1991. Double modals in Hawick Scots. In Peter Trudgill & J. K. Chambers (eds.), *Dialects of English: Studies in grammatical variation*, 74–103. London: Longman.

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.

Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech*. 54(1). 99–121.

Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry.* 39(1). 1–53.

Feagin, Crawford. 2013. Entering the community: Fieldwork. In J. K. Chambers & Natalie Schilling (eds.), *The handbook of language variation and change*, 17–37 (2nd edn). London: Wiley Blackwell.

Fisher, Sabriya. 2018. *Variation and change in past tense negation in African American English*, 1–13. Philadelphia, PA: University of Pennsylvania dissertation.

https://doi.org/10.1017/S0022226724000203 Published online by Cambridge University Press

Frampton, John. 2001. The *amn't* gap, ineffability, and anomalous *aren't*: Against morphosyntactic competition. In Mary Andronis, Christoper Ball, Heidi Elston & Sylvain Neuvel (eds.), *Proceedings of Chicago Linguistics Society 37*, 1–13. Chicago, IL: Chicago Linguistics Society.

Francis, Nelson. 1985. *Amn't I*, or the hole in the pattern. In Wolfgang Viereck (ed.), *Focus on: England and Wales varieties of English around the world*, 141–152. Amsterdam: John Benjamins.

Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.

Grant, William. 1931. *The Scottish national dictionary*. Edinburgh: Scottish National Dictionary Association.

Hudson, Richard. 2000. *I amn't. *Language.* 76(2). 297–323.

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.

Jamieson, E., David Adger, Caroline Heycock, Jennifer Smith & Gary Thoms. To appear. When intuitions (don't) fail: Combining syntax and sociolinguistics in the analysis of Scots. *English Language and Linguistics*.

Johnston, Paul. 1997. Regional variation. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 433–513. Edinburgh: Edinburgh University Press.

Kaisse, Ellen M. 1985. *Connected speech: The interaction of syntax and phonology*. London: Academic Press.

Kodner, Jordan. 2023. What learning Latin verbal morphology tells us about morphological theory. *Natural Language and Linguistic Theory.* 41. 733–792.

Labov, William. 1972. Negative attraction and negative concord in English grammar. *Language.* 48(4). 773–818.

Labov, William. 1984. Field methods of the project on linguistic change and variation. In John Baugh & Joel Sherzer (eds.), *Language in use: Readings in sociolinguistics*, 28–66. Englewood Cliffs, NJ: Prentice Hall.

Langendoen, D. Terence. 1970. *Essentials of English grammar*. New York: Rinehart and Winston.

Lawrence, Daniel. 2017. The social perception of a sound change. *Penn Working Papers in Linguistics*. 23(1). 111–120.

Liang, Kevin, Diana Marsala & Charles Yang. 2022. Distributional learning of syntactic categories. In Ying Gong & Felix Kpogo (eds.), *Proceedings of the 46th annual Boston University Conference on Language Development*, 442–455. Somerville, MA: Cascadilla Press.

Millar, Robert McColl. 2007. *Northern and Insular Scots*. Edinburgh: Edinburgh University Press.

Schütze, Carson T. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology.* Berlin: Language Sciences Press.

Shuken, Cynthia. 1984. Highland and Island English. In Peter Trudgill (ed.), *Language in the British Isles*, 152–166. Cambridge: Cambridge University Press.

Smith, Jennifer. 2000a. *Synchrony and diachrony in the evolution of English: Evidence from Scotland*. York: University of York dissertation.

Smith, Jennifer. 2000b. 'you 0 na hear o' that kind o' things': Negative *do* in Buckie Scots. *English World-Wide.* 21 (2). 231–259.

Smith, Jennifer, David Adger, Brian Aitken, Caroline Heycock, E. Jamieson & Gary Thoms. 2019. *The Scots syntax atlas*. University of Glasgow. https://scotssyntaxatlas.ac.uk.

Tagliamonte, Sali & Jennifer Smith. 2002. 'Either it isn't or it's not': NEG/AUX contraction in British dialects. *English World-Wide.* 23(2). 251–281.

Thoms, Gary, David Adger, Caroline Heycock & Jennifer Smith. 2019. Syntactic variation and auxiliary contraction: The surprising case of Scots. *Language.* 95(3). 421–455.

Thoms, Gary, David Adger, Caroline Heycock, E. Jamieson & Jennifer Smith. 2023. English contracted negation revisited: Evidence from varieties of Scots. *Language.* 99(4). 726–759.

Yang, Charles D. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yang, Charles. 2016. *The price of productivity.* Cambridge: MIT Press.

Yang, Charles. 2017. How to wake up irregular (and speechless). In Claire Bowern, Laurence Horn & Rafaella Zanuttini (eds.), *On looking into words (and beyond)*, 211–233. Berlin: Language Sciences Press.

Zanuttini, Rafaella, Jim Wood, Jason Zentz & Laurence Horn. 2018. The Yale grammatical diversity project: Morphosyntactic variation in North American English. *Linguistic Vanguard.* 4(1). 1–15.

Zwicky, Arnold & Geoffrey K. Pullum. 1983. Cliticization vs inflection: English n't. *Language.* 59(3). 502–513.

Cite this article: Thoms, Gary, Adger, David, Heycock, Caroline, Jamieson, E. and Smith, Jennifer. 2025. "Explaining microvariation using the Tolerance Principle: plugging the *amn't* gap." *Journal of Linguistics*. 61, 369–396. doi:10.1017/S0022226724000203