Syntactic variation and auxiliary contraction: the surprising case of Scots

Gary Thoms, David Adger, Caroline Heycock and Jennifer Smith*

April 18th 2019

Abstract

This article is concerned with the distribution of contracted auxiliaries in English, in particular the restriction against their occurrence in the immediate context of a gap created by movement or ellipsis. We document apparent exceptions to this restriction in varieties of Scots, all occurring in what we call the LOCATIVE DISCOVERY EXPRESSION. We analyse these as mirative constructions, and using new data from the Scots Syntax Atlas we describe patterns of variation in the acceptance of auxiliary contractions in locative discovery expressions which provide clues as to the role of syntax in conditioning auxiliary contraction. Adapting the proposal in Wilder (1997), where contracted auxiliaries are prosodically incorporated into the following predicate, we provide an account in which the differences across dialects with respect to contraction are explained in terms of the availability of different abstract structures.

Keywords: auxiliary contraction, microcomparative syntax, Scots

1. Introduction. This article is concerned with AUXILIARY CONTRACTION in English, whereby finite auxiliary verbs appear in a reduced or contracted form attached to a preceding element. This is illustrated in 1: here we see reduced forms of *will*, *have*, *has* and *is* attaching onto preceding elements of various kinds (pronominal subjects, a relative pronoun and a *wh*-phrase).[1]

(1)   a.   I**'ll** be there.
        b.   He**'s** a fool.
        c.   You**'ve** made it.
        d.   The dog**'s** been out already.
        e.   Those who**'ve** heard this talk before can zone out.
        f.   How high**'s** the water momma?

[1]Here and throughout we rely on standard orthographic means of representing contractions, leaving fuller phonetic transcriptions only for where it is relevant.

Interest in auxiliary contraction among syntacticians flourished in the 1970s when a number of syntactic restrictions on such contractions were noticed.[2] Perhaps the best-known of these was that noted by King (1970): auxiliary contraction is impossible when the immediately following context is an ellipsis site, 2b, or a gap created by movement, 3b. We will call this the GAP RESTRICTION. King and others since have demonstrated that the gap restriction does not follow from purely phonological conditions such as utterance-finality (cf. 3b).

(2)   Ellipsis
       a.   Mary's usually there when Kim is ____ .
       b.   *Mary's usually there when Kim's ____ .
       c.   Mary's usually there when Kim's there.

(3)   Movement
       a.   I don't know where the bathroom is ____ in this building.
       b.   *I don't know where the bathroom's ____ in this building.
       c.   The bathroom's where you met me earlier.

The gap restriction has proven to be a particularly strong generalization, with no counterexamples arising in the syntax literature to date (see Close 2004 for a recent summary) or from corpus analysis (MacKenzie 2012), and various analysts have taken it to motivate articulated models of the syntax-phonology interface: see Bresnan (1978), Selkirk (1984), Wilder (1997), Inkelas & Zec (1993). All derive strong constraints which lead us to expect that the gap restriction should be exceptionless.

     With this in mind, we examine a curious corner of some varieties of English spoken in Scotland (herein 'Scots'), where we seem to see exceptions to the gap restriction. This corner involves examples like those in 4-5, which we call LOCATIVE DISCOVERY EXPRESSIONS (LDEs).

(4)   a.   Here it is!
       b.   There it is!

(5)   a.   Here he is!
       b.   There he is!

Our observation is that, for some speakers of Scots, this construction exceptionally allows for what look like violations of the gap restriction. Below are two examples that have been overheard by the authors in different parts of the country. In both cases the auxiliary was contracted onto the pronoun, with auxiliary and pronoun being phonologically reduced and fused to form a monosyllable.

(6)   [A is looking for a bin bag and she asks B about its whereabouts.]
       A: Did you steal our bin bag?
       B: No!
       A: [spots the bin bag hanging on a door handle] There it's!

(7)   [A and B are in a museum looking for a child who has wandered to another exhibit, and A spots the child and says]
       A: There he's.

Examples such as these are widely accepted by speakers of certain varieties of Scots, but they are rejected by speakers from elsewhere in Scotland and by speakers of any other English variety we have managed to test. The fact that such examples are rejected by all other speakers of English[2] can plausibly be explained in

terms of the gap restriction, since the configuration is broadly similar to what we see in 3, with the position following *be* lying empty. But if this is right, then the question is why speakers of some varieties of Scots accept and produce examples of this kind. This question is particularly pressing since the same speakers share the judgments reported in 2-3, in that they behave like all other speakers in rejecting auxiliary contraction before a gap in all other syntactic contexts. This makes 6-7 a curiously specific exception to what is otherwise a very strong generalization.

In this article we seek to explain this curious exception, and we do so by taking into account a fuller range of microvariation observed with variants of the LDE construction. We argue that the gap restriction can be maintained, but it requires a theoretically sophisticated understanding of what 'gap' means. To get there, we first develop an explicit analysis of the construction in question in section 2 and demonstrate the tightly constrained distribution of this exceptional contraction phenomenon. In section 3 we then describe the microvariation picture in more detail and show that distinct but related systems exist across dialects. We use two datasets to establish this picture: (i) grammaticality judgment data gathered from in-depth interviews with non-naive speakers from ten settlements across Scotland, and (ii) grammaticality judgment data from the Scots Syntax Atlas, a major dialect survey of varieties of Scots which comprises data from 200+ dialect phenomena from 140 locations. In section 4 we provide an analysis of the full range of exceptional contraction data in terms of prosodic incorporation, where differences in abstract structures and pronoun inventories determine when a clitic auxiliary can be contracted leftwards onto a preceding host.

2. Locative discovery expressions.

A remarkable feature of the exceptional contractions in varieties of Scots is that they are only possible in the environments in 6-7: the same speakers who accept these examples reject auxiliary contraction with gaps created by ellipsis 2 and *wh*-movement 3. They also reject auxiliary contraction with the other constructions which show sensitivity to contraction, such as pseudogapping 8 and comparative subdeletion 9, and other A′-extraction rules like relativization 10 and topicalization 11.

(8) a. Joan's taken more from you than Bill has ⎽⎽⎽ from me.
  b. *Joan's taken more from you than Bill's ⎽⎽⎽ from me.

(9) a. Sandy's a bigger bibliophile than Kim is ⎽⎽⎽ a stamp collector.
  b. *Sandy's a bigger bibliophile than Kim's ⎽⎽⎽ a stamp collector.

(10) a. This isn't the land of plenty that everyone says it is ⎽⎽⎽ .
  b. *This isn't the land of plenty that everyone says it's ⎽⎽⎽ .

(11) a. His own worst enemy, I don't think he is ⎽⎽⎽ .
  b. *His own worst enemy, I don't think he's ⎽⎽⎽ .

Something specific to these *here/there*-constructions opens up the possibility of variation with auxiliary contraction, so we first establish a baseline analysis for them.

The examples in question are all of the format in 12, where an initial locative *here* or *there* is followed by a nominative subject pronoun and then a finite form of *be*. Structures of this form are of course possible for all speakers when *be* appears in its uncontracted form. Some simple examples are given in 13.[3]

---

speakers of English'. In such cases we report on data which has been gathered in informal acceptability judgment consultaitons with speakers of various standard and nonstandard Englishes, or which have been presented in the literature as uncontroversial, and so we refer to these speakers and do not make overstrong claims about the universality of our data points. We ask that the reader bear this qualification in mind.

[3]We have found little discussion of the syntax of this construction in the previous literature. Huddleston & Pullum (2002, :1389-1390) contains a brief discussion of some related constructions. A more thorough discussion of these construction (where they are called *presentatives*) has recently appeared in Zanuttini (2017) and Wood & Zanuttini (2018).

(12)   *t/here* pronoun$_{nom}$ be$_{fin}$

(13)   a.   Here it is!
        b.   Here you are!
        c.   There she is!
        d.   There he is!
        e.   There they are!

In all cases, the initial locative bears focal stress, with the rest of the sentence deaccented. The use of these examples requires a context where some interlocutor, typically the speaker, has just discovered the exact location of the entity referred to by the subject; we call this aspect of the meaning of these constructions a DISCOVERY INFERENCE, where by discovery, we mean that the speaker is signalling a sharp change in their epistemic state, usually the discovery of new information about location, although this may be weakened in certain pragmatic contexts.

This informational change is an obligatory component of the meaning of such sentences. Consider 14: a sentence of the form in 12 cannot be used felicitously in a context where the exact location of the subject is not known, even though a simple locative is perfectly acceptable in the same context.

(14)   [context: B asks A where his coat is]
        A: #Here it is, it's somewhere in this room.
        A′: It's here, it's somewhere in this room.

Knowing the exact location of the subject is not sufficient; rather, a discovery must be made, meaning that the relevant information must be new and unexpected to some extent. Sentences of the form of 12 cannot be used to report knowledge which is not new to any interlocutor in the context.

(15)   [contexts: A and B arrive in a bar, where there is supposed to be a party, but no one's there.]
        A: I thought you knew where the party is?
        B: #I told you already, here it is!
        B′: I told you already, it's here!

Typically the discovery in question is attributed to the speaker, but it can also be attributed to another participant in the discourse. Consider 16: this could be uttered in a context when the speaker is raising their hand to grab the attention of an addressee, and so in this case it is the addressee that has made the discovery. Again this construction is impossible in a context where no discovery is made, even though the simple predication is possible, as shown by 17.

(16)   [context: A waves to B, who has just spotted him across the room]
        Here I am!

(17)   [context: A is on the phone to B, and B asks if A is in the departure lounge, where they are supposed to be meeting.]
        A: #Here I am ... I'm at the bar.
        A′: I'm here ... I'm at the bar.

We therefore call these constructions LOCATIVE DISCOVERY EXPRESSIONS (LDEs): 'locative' due to the initial locative, and 'discovery' to acknowledge the obligatory discovery inference.

Typical examples of the LDE are in the present tense, but past tense forms are also possible, so long as the discovery is anchored to the event time. The discovery component is still crucial in the past tense, as 18b shows.

(18)   a.   I pushed the door open and lo and behold, there he was!
        b.   #I've told you already, there he was!

They are also possible in habitual contexts, where the clause is formally nonpast, but the discovery being described is not at the utterance time. Once more, the discovery component (the sharp change in epistemic state) is obligatory, as the contrast with 19b, which is pragmatically odd, shows.

(19)   a.   Whenever I open the door, there he is.
        b.   #Whenever he pretends not to be at home, there he is.

One thing to note about the examples in 18a and 19a is that they differ from the previous examples with respect to their stress patterns. In all of the examples in 13, the main stress falls on the initial locative, and both the auxiliary and the subject pronoun are destressed monosyllables (*thére he ìs*); it is possible for there to be a minor secondary stress on the auxiliary, but this does not seem to be obligatory. However with 18a and 19a the main stress falls on the auxiliary, with the locative bearing secondary stress (*thère he ís*). We suggest that this difference is due to the fact that in 13, in the here and now, the location is completely new information and its discovery is often accompanied by a pointing gesture of some kind, whereas in the past and habitual cases the location is at least partly discourse-given and so is not in focus in the same way that it is in the here-and-now cases.

The LDE has a few other significant properties which distinguish it from other similar constructions. First, for the most part the initial locative of the LDE must be a minimal proform, and so it cannot be modified or replaced by a more complex PP with lexical content.

(20)   a.   *Over there he is!
        b.   *Right here he is!
        c.   *In the corner he is!

The same restriction is not seen with locative inversion, the LDE's much better-studied cousin (see e.g. Coopmans 1989, Bresnan 1994, Levin & Hovav 1995 ch.6). Locative inversion involves initial placement of a PP and postposing of the subject. As the following examples show, locative inversion is not subject to a proform restriction, although it is compatible with proforms (Culicover & Levine 2001).

(21)   a.   Into the room walked Mary.
        b.   Over there is Mary Smith.
        c.   Here is the book.
        d.   In the room was HIM!

When sentences with the locative inversion-type word order are used in discovery contexts, there is a strong preference to use a simple proform PP. Thus 21b is acceptable only in a context where *over there* is one of a number of salient locations which the speaker is discussing, and they are not reporting new information. It may be used in a context where the speaker is pointing out a number of people in a room and telling the listener who is who. But when we set up a discovery context, where the speaker is happening upon new information and the locative bears pitch accent, the modified version is odd.

(22)   a.   #Over thére's Mary Smith! I didn't expect to see her today.
        b.   Thére's Mary Smith! I didn't expect to see her today.

If we take the discovery inference to be the crucial component of the LDE, then we would say that 22b is an LDE with a nonpronominal postverbal subject, while 21b is an instance of locative inversion. This is the position we adopt, so we account for the variable position of the subject in LDEs in section 4; however, given

that contraction is generally not restricted when the subject appears after the auxiliary (see Kaisse 1985), LDEs with postverbal subjects are not directly relevant to our analysis of the unexpected contraction data, so our discussion of these cases will be brief. We therefore restrict our attention to LDEs of the form in 12.

In general, LDEs tend to be informationally light, in that they do not convey any new information other than the exact location of the subject and the fact that this is a novel discovery. When the subject is in Spec-TP (namely, the canonical subject position in a sentence), it is always discourse-old, and so it is typically an unstressed pronoun, although epithets are also possible, as in 23a, and some speakers allow a discourse-old full DP in the subject position too, as in 23b. Indefinites and focussed DPs do not occur in this position, as shown by 24; typically, an LDE with such subjects involves use of the postverbal subject position (*here's a dog*).

(23)   a.   Here the little bugger is!
       b.   %I'm trying to find the information about the concert. Ah, here the listing is!

(24)   a.   *Here a dog is!
       b.   *Here HE is!

Little elaboration of the lower part of the clause is possible. Adverbs with substantial semantic content like *definitely* or *predictably* cannot be included, and expanding the clause structure to include a further layer of embedding with a predicate like *seem* is also disallowed.

(25)   a.   *Here he definitely is!
       b.   *Here he predictably is!
       c.   *Here he seems to be!

Adverbs are not excluded entirely though, as the following examples from a Google Books search show that *suddenly* is possible when it describes a discovery which was made in the past. This suggests that the relevant restriction is tied to the meaning of the LDE, and not some syntactic restriction.

(26)   In the high alpine forest, however, there he suddenly was, looking at us![4]

(27)   She must have heard the commotion we were making, because there she suddenly was, saying, 'What is the meaning of this?'[5]

Typically the verb in an LDE is *be*, but some speakers allow these constructions with lexical verbs with locative semantics. 28 is accepted by some speakers if *there* receives particularly strong stress, *put* is deaccented and the context is one where the interlocutors are looking for whatever it is that referred to by the object.

(28)   There you put it!

LDEs also seem to be possible with *come* and *go*, as in 29. However these do not seem to have corresponding unfronted forms, as the meanings of the sentences in 30 are not transparently related to those in 29 in the way that *here he is* is related to *he is here*, and 29b is also unusual in that it does not necessarily imply motion. This suggests that the structures in 29 may be a kind of light verb construction rather than lexical verbs with semantic content relating to motion.[6] Most other verbs of motion do not work at all, as can be seen by 31.

---

[4]From 'Brian's Legacy: As Shared by His Father Siegfried Other' by Siegfried Othmer and Brian Othmer, accessed at http://bit.ly/2tQ5Vy1 on 11/7/17.

[5]From 'Stories from Blue Latitudes: Caribbean Women Writers at Home and Abroad', edited by Elizabeth Nunez and Jennifer Sparrow, accessed at http://bit.ly/2sXq6qx on 11/7/17.

[6]The uses of *go* in this construction may be part of a more general phenomenon in English where a semantically bleached *go* is used much like the copula, for instance in expressions like *here you go* or *there he goes talking rubbish about you again*.

(29) a. Here he comes!
      b. There he goes!

(30) a. #He comes here!
      b. #He goes there!

(31) a. *Here he runs!
      b. *There he arrives!

Finally, it is possible for there to be some additional locative material in the postverbal position, where the locative provides more explicit information about the location involved. Typically such examples would be accompanied by some gesture, and the main stress would fall on the second locative.

(32) a. Here it is on the shelf.
      b. There it is in the corner.

There is no necessary prosodic boundary between the copula and the following prepositional phrase in these examples, in contrast to examples where a prepositional phrase has been extraposed (cf. *We put the book once again on the shelf*). Non-Scots speakers do not allow the auxiliary to contract onto the subject verb in these cases.

(33) a. *Here it's on the shelf.
      b. *There it's in the corner.   e.g. Standard English

The effect seen here looks similar to what we saw with pseudogapping and comparative subdeletion in 8-9 above, where the gap restriction shows its force despite there being additional material in the postverbal position.

    To summarize, we have seen that sentences of the format in 12 are restricted in a number of ways. Their context of use is restricted by an obligatory discovery inference. They are also restricted in their form, as indicated by the basic template in 12, although we saw that some restrictions are stricter than others: while the restriction on elaborating the lower clause seems to have some give, the restriction on the initial locative does not. In section 4 we propose an analysis of LDEs as mirative constructions, where the initial *here/there* is realized in a mirative complementizer position which encodes the discovery inference. But first we extend the description of the exceptional contraction data in Scots, considering data from a wide variety of constructions and dialect regions.

3. The microcomparative picture with Scots LDEs.   Recall that our main goal is to explain the fact that some speakers of Scots allow auxiliary contraction in LDEs, in apparent violation of the gap restriction. So far we have only seen a handful of examples of this phenomenon in the introduction. In this section, we expand the empirical picture by considering data from a number of related constructions from locations all over Scotland.

3.1. Kinds of nonstandard LDEs. The examples of unexpected auxiliary contraction that we opened with involved a single locative expression followed by pronoun-auxiliary cluster. In all of these cases, the initial *here/there* and the pronoun-auxiliary cluster form a single prosodic unit, with no major intonation boundaries.

(34) a. There you're!
      b. Here she's!
      c. There they're!   Scots

These are not the only kinds of nonstandard contraction that we find in LDE constructions in Scots, as cases

also exist where there is additional structure following the pronoun-auxiliary cluster. Some examples are given in 35-36, with various different pronominal or pronoun-like subjects (including epithets, cf. 23a above). There are no major intonation breaks on either side of the pronoun-auxiliary cluster in any of these cases.

(35)   a.   Here it's here!
       b.   There she's there!
       c.   Here I'm here!
       d.   There you're there!
       e.   Here the wee bugger's here!
       f.   There the bastard's there.   Scots

(36)   a.   Here it's right here!
       b.   Here I'm over here!
       c.   There they're under the table!
       d.   There it's at the top of the hill!
       e.   Here he's behind me!   Scots

In 35, the element that follows the pronoun-auxiliary is identical to the initial locative, and in 36 there is also a second locative, but it is not identical to the initial one. Versions of 35-36 without auxiliary contraction (*here it is here*) are acceptable for many speakers of non-Scots varieties. Therefore it is the contraction which marks 35-36 out as nonstandard, and seemingly specific to certain Scots varieties. As such, these examples can be considered of a piece with 34 as instances of unexpected contraction. But since the cases in 34, 35 and 36 turn out not to have an identical distribution across Scots varieties, we need terminology to distinguish them. We call cases such as 34 MINIMAL LDEs, to reflect their compact form. We call cases such as 35 DOUBLED LDEs and cases such as 36 ELABORATED LDEs; we group doubled and elaborated LDEs together as one with the term REPEATING LDEs. Finally, LDEs with no auxiliary contraction at all will be called simply STANDARD LDEs. Table 1 summarizes the different types of LDE.

Type	Subtype	Example
Standard LDE		Here it is
Minimal LDE		Here it's
Repeating LDE	Doubled LDE	Here it's here
Repeating LDE	Elaborated LDE	Here it's under the table

Table 1: Summary of LDE types

The nonstandard LDEs share the core properties of the standard LDEs discussed above. The initial *here/there* cannot be modified, even if the second one is.

(37)   a.   (*Right) here it's!   minimal LDE
       b.   (*Right) here it's here!   doubled LDE
       c.   (*Right) here it's right here!   elaborated LDE

All these nonstandard LDEs necessarily generate the same discovery inference we have seen for the standard case.

(38)   [contexts: A and B arrive in a bar, where there is supposed to be a party, but no one is there.]
       A: I thought you knew where the party is?
       B1: #I told you already, here it is!
       B2: #I told you already, here it's!

B3: #I told you already, here it's here!
B4: #I told you already, here it's in this bar!
B5: I told you already, it's here!
B6: I told you already, it's here in this bar!

Examples such as B3-B4 are acceptable for both Scots speakers and other English speakers if there is a major intonational break following the first locative (*here, it's here!*). A major intonation break also makes it possible to modify the initial locative.

(39)   Right here, it's right here!

Since they show quite distinct properties, we take cases with a major intonational break following the initial locative to be distinct from LDEs, and so we put them to one side in what follows.

3.2. The data. As noted in section 1, there is considerable variation in acceptance of the different kinds of LDEs in Scotland. We now describe that variation in the rest of this section, drawing upon data from two sources.

The first dataset arises from in-depth questionnaires with native speakers from ten locations across Scotland. All were non-naive consultants, and the majority of them are linguists. We call this the in-depth questionnaire data. The ten locations are distributed across Scotland, covering most of the major dialect regions of the country, as shown in Figure 1.[7] These data give us a preliminary understanding of the variation that we find in minimal LDEs. We tested all combinations of minimal and doubled LDEs with the different subject pronouns and locatives, and we also tested a handful of elaborated LDEs with different locatives and subject pronouns. While most of these are the result of interviews with just one informant, a number of the data points have been confirmed with multiple informants.[8]

---

[7]The Shetland Islands are located more than 100km to the northeast of the mainland and are inserted in a box for space reasons.
[8]All maps were generated using Mapbox: https://www.mapbox.com/about/maps/.
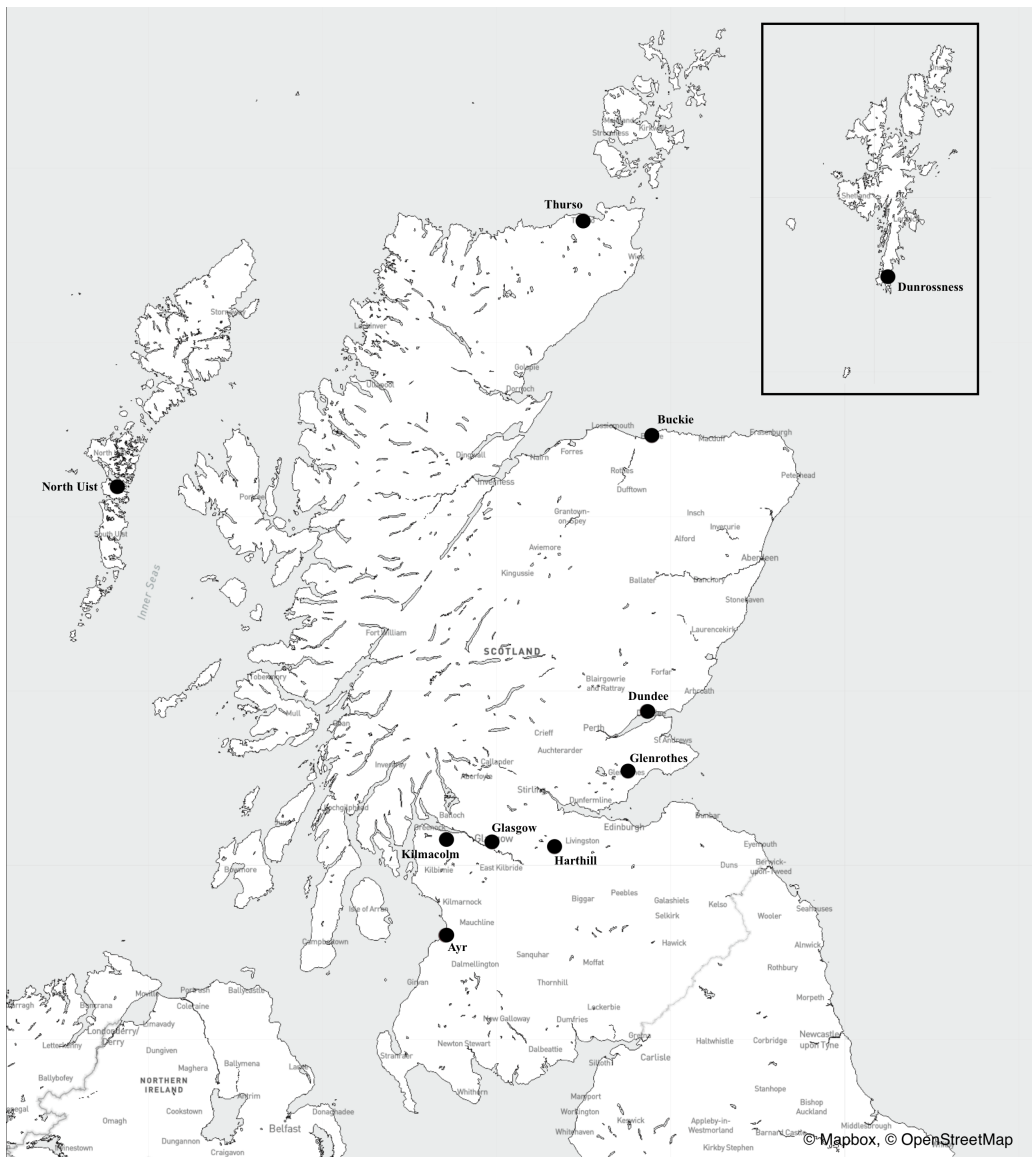
Figure 1: map of locations surveyed for in-depth questionnaires

The second dataset comes from the Scots Syntax Atlas (SCOSYA; Smith 2015-2019), a fully searchable online atlas of the non-standard morphosyntactic features found in the varieties of Scots spoken across Scotland. To date, 140 locations have been targeted, with four speakers in each area – two from the 18-25 age bracket, and two from the 65+ age bracket. As the focus of the study is the documentation of non-standard norms, participants were recruited on the basis of a set of sociolinguistic criteria characteristic of vernacular speakers (see Labov 1984). Sociolinguistic research shows that the fieldworker collecting the data has a significant impact on language norms (Smith & Durham 2012), therefore in order to mitigate the Observer's Paradox (Labov 1972) and thus gain access to systematic vernacular patterns, we recruited fieldworkers who came from the communities that we investigated.

Figure 2 displays the locations where the full datasets have been gathered to date, where the concentration of sampling reflects population density across Scotland.
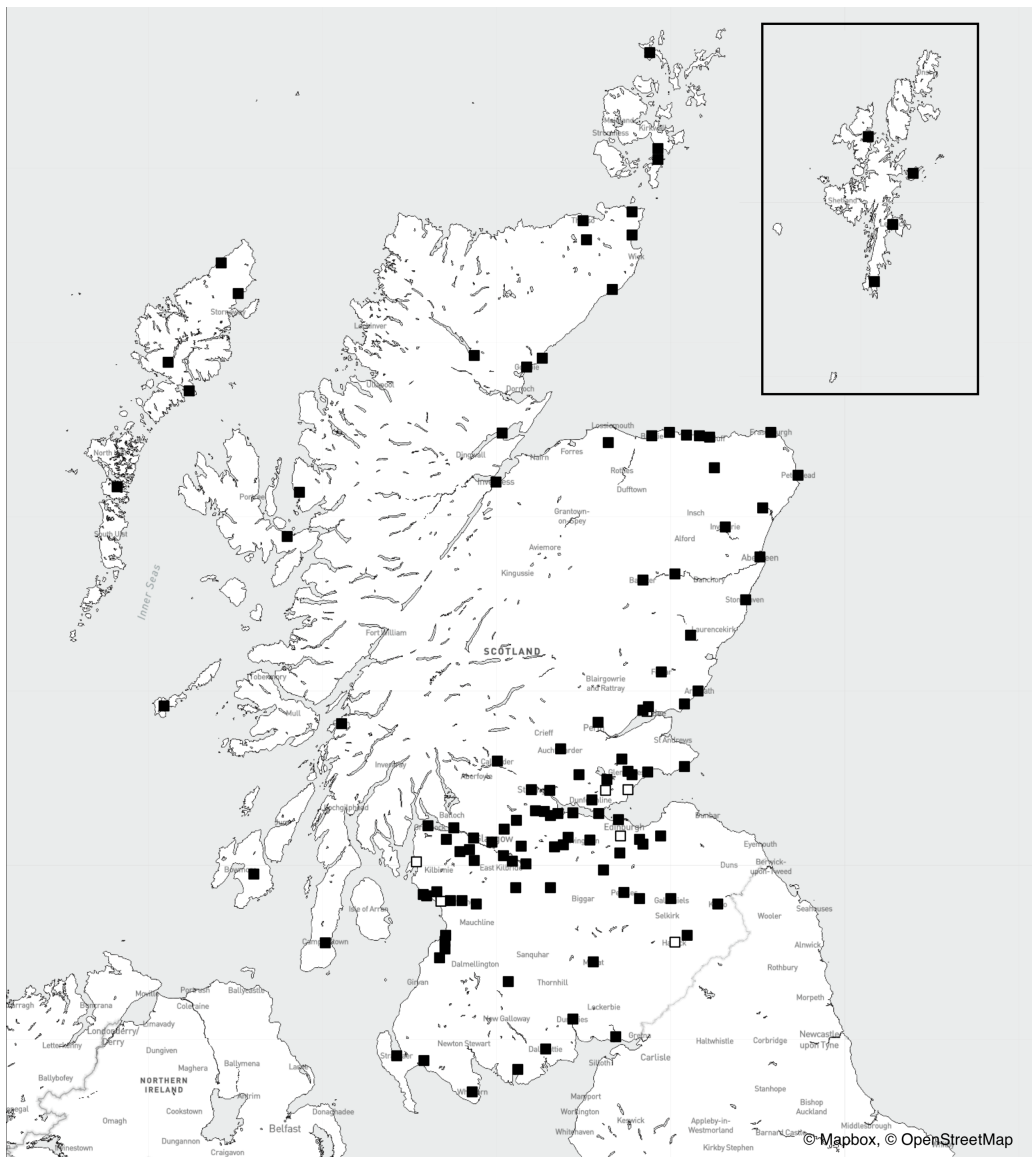
Figure 2: locations sampled in SCOSYA fieldwork

Two kinds of data were gathered in the SCOSYA interviews: (i) spontaneous conversation data between community dyads (ii) grammaticality judgment data. The spoken data was gathered through sociolinguistic interviews (Labov 1984), with full text-to-sound aligned transcription of each recording using Transcriber (Boudahmane et al. 2008). The grammaticality judgment data was gathered in face-to-face interviews using a questionnaire containing 200+ questions testing a variety of grammatical phenomena. Judgments were elicited using the face-to-face interview method described in Barbiers & Bennis (2007). Questionnaires were tailored to different dialect areas with minor changes in lexis to reflect localized use, and certain examples were added for some areas and excluded for others in order to drive down questionnaire length; these exclusions were guided by initial consultations with linguist informants from the different regions. Participants were asked to rate sentences on a 5-point Likert scale (Schütze & Sprouse 2013), with 1 being the bottom end of the scale corresponding to unacceptability and 5 being the top end corresponding to full acceptability. Each point on the scale was given four different descriptions, with the top line describing the extent to which the speaker would say the sentence – 1 as 'I would never say that' and 5 as 'I would definitely say that' – and other lines restating this in different terms. The data was uploaded to a database which feeds into the online atlas used to generate a

set of maps. These data provide an archive of vernacular speech from communities from all over Scotland, the majority of which have not been investigated before.

As examples of LDEs are very rare in the sociolinguistic interview data,[9] the primary data source for the current analysis is the grammaticality judgment data. With regard to LDEs, in the SCOSYA interviews there were examples testing minimal LDEs and doubled LDEs, with both *here* and *there*, and with *I'm* and *it's* pronoun clusters; 8 examples in all. Elaborated LDEs were not tested in the SCOSYA interviews. The full range of pronoun clusters was tested in the in-depth questionnaires (i.e. the work with non-naive consultants outside the SCOSYA interviews), which were more focussed and less time-pressured. Elaborated LDEs were only tested in the in-depth questionnaires.

3.3. Key observations. Tables 2 and 3 present the in-depth questionnaire data. We can make a number of observations regarding the patterns observed in the tables, and we bring in the atlas data where it supplements the points made by the in-depth questionnaire data.

	Ayr	Glasgow	Harthill	Kilmacolm	Dundee	Glenrothes	Buckie	Thurso	N. Uist	Dunrossness
Here I'm	✓	✓	✓	*	*	*	*	*	*	*
There I'm	✓	✓	✓	*	*	*	*	*	*	*
Here you're	✓	✓	✓	*	*	*	*	*	*	*
There you're	✓	✓	✓	*	*	*	*	*	*	*
Here he's	✓	✓	✓	*	*	*	*	*	*	*
There he's	✓	✓	✓	*	*	*	*	*	*	*
Here she's	✓	✓	✓	*	*	*	*	*	*	*
There she's	✓	✓	✓	*	*	*	*	*	*	*
Here it's	✓	✓	✓	✓	*	*	*	*	*	*
There it's	✓	✓	✓	✓	*	*	*	*	*	*
Here we're	✓	✓	✓	*	*	*	*	*	*	*
There we're	✓	✓	✓	*	*	*	*	*	*	*
Here they're	✓	✓	✓	*	*	*	*	*	*	*
There they're	✓	✓	✓	*	*	*	*	*	*	*

Table 2: Minimal LDEs

---

[9]There are just eight examples of LDEs in the 3m word SCOSYA corpus, six of which are found in a single passage of one interview, when the interviewees are discussing a photograph. We suspect that the rarity of LDEs is an artefact of the sociolinguistic interview format, which typically tends towards personal narratives and conversation rather than discussion of where things are in the immediate context.

	Ayr	Glasgow	Harthill	Kilmacolm	Dundee	Glenrothes	Buckie	Thurso	N. Uist	Dunrossness
Here I'm here	✓	✓	✓	✓	✓	*	*	*	*	*
There I'm there	✓	✓	✓	✓	✓	*	*	*	*	*
Here you're here	✓	✓	✓	✓	✓	*	*	*	*	*
There you're there	✓	✓	✓	✓	✓	*	*	*	*	*
Here he's here	✓	✓	✓	✓	✓	*	*	*	*	*
There he's there	✓	✓	✓	✓	✓	*	*	*	*	*
Here she's here	✓	✓	✓	✓	✓	*	*	*	*	*
There she's there	✓	✓	✓	✓	✓	*	*	*	*	*
Here it's here	✓	✓	✓	✓	✓	✓	✓	✓	*	*
There it's there	✓	✓	✓	✓	✓	✓	✓	✓	*	*
Here they're here	✓	✓	✓	✓	✓	*	*	*	*	*
There they're there	✓	✓	✓	✓	✓	*	*	*	*	*
There it's over there	✓	✓	✓	✓	✓	*	*	*	*	*
Here it's right here	✓	✓	✓	✓	✓	*	*	*	*	*
There it's in the corner	✓	✓	✓	✓	✓	*	*	*	*	*

Table 3: Doubled and elaborated LDEs

OBSERVATION 1. Speakers accept both the *here* and *there*-based form of a given LDE type. That is, we do not find speakers accepting *here I'm here* but not *there I'm there*. This holds of all of the in-depth questionnaire data, and it also plays out in the atlas data: acceptance of a *here*-variant generally goes together with acceptance of a *there*-variant, with only a very small amount of variation which may be attributable to noise.

OBSERVATION 2. There is an effect of subject pronoun type: speakers either accept LDEs of a given kind with all subject types, or they only accept them with *it*. This is seen both with minimal LDEs and doubled LDEs: Kilmacolm only allows *it's*-based minimal LDEs, and both Buckie and Glenrothes allow *it's*-based doubled LDEs. Note that the effect is not an effect of person, since examples based on *he's* and *she's* pattern with others based on *you're* and so on, and not with *it's*. There is only atlas data for *it's* and *I'm* based LDEs, but these allow us to see a general trend for *it's*-based LDEs being more widely accepted than *I'm*-based LDEs. Compare Maps 3 and 4, which represent data for *here it's* and *here I'm* respectively. These use the search criteria 'rated 4 or 5 by two or more people', a setting which we typically take to pick out locations where the feature in question is generally accepted; black dots mark locations which meet these criteria, and white dots mark those which don't. With these criteria (and excluding incomplete datasets), 20% accepted both *here it's* and *here I'm*, 23% accepted only *here it's* (as in Kilmacolm), and 54% accepted neither. Just 2% of respondents (8 speakers) accepted *I'm* but not *it's*; in two cases *here it's* was given a middle score of 3, and four of the others come from two locations. We take this number to be small enough to be attributed to noise or error.
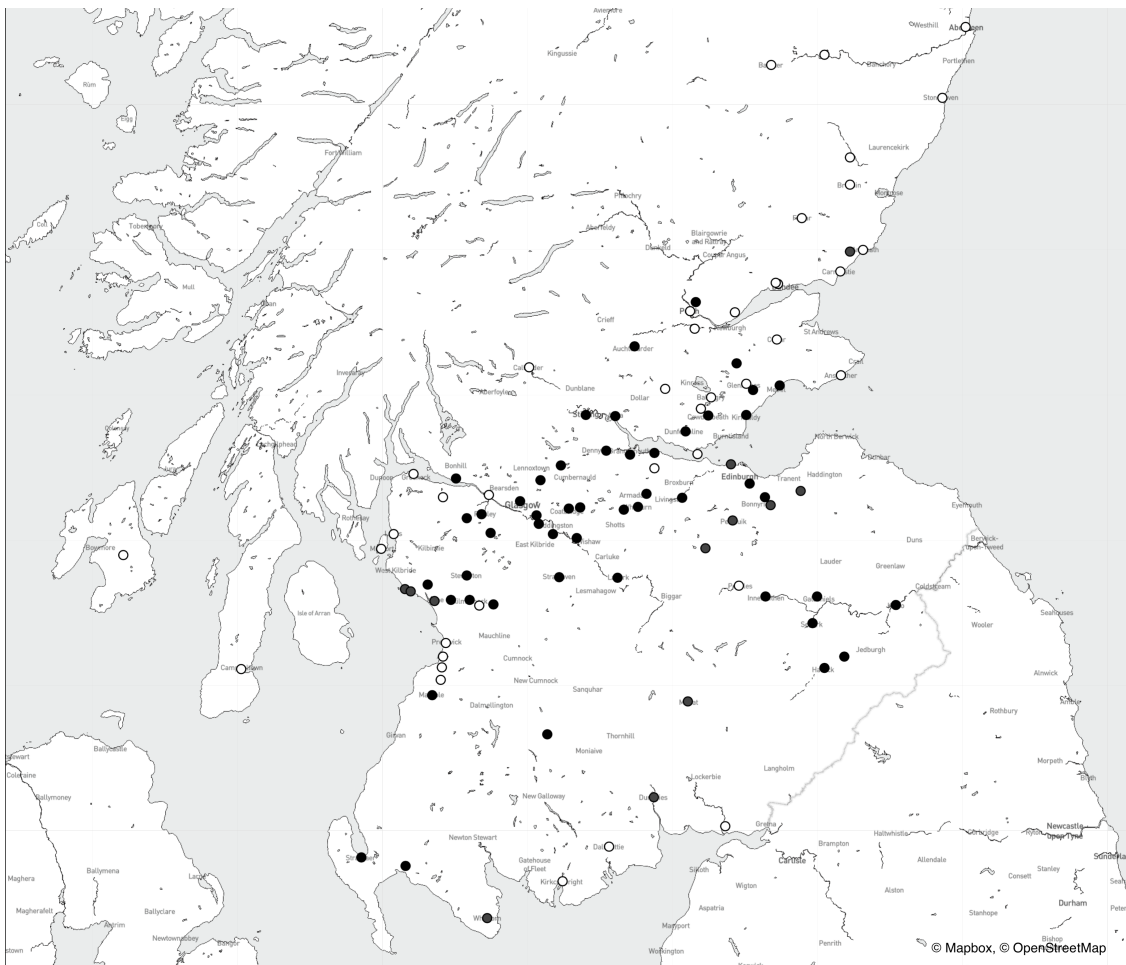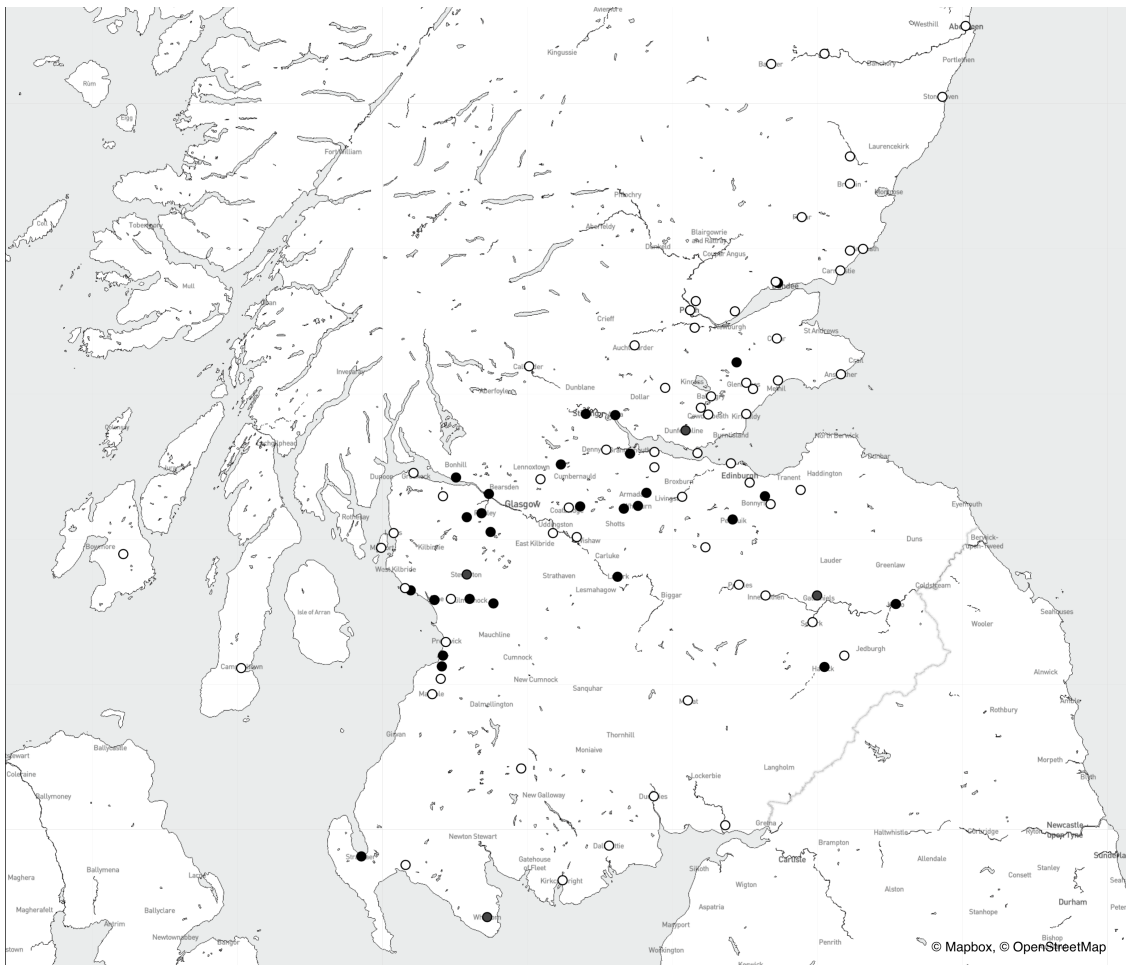

Figure 3: *here it's* rated 4/5 by 2 or more

Figure 4: *here I'm* **rated 4/5 by 2 or more**

OBSERVATION 3. There is a one-way entailment relation between the different kinds of LDEs, such that if a speaker accepts minimal LDEs such as *here it's*, they accept doubled and elaborated LDEs. We see this in the in-depth questionnaire data: Ayr, Glasgow, Harthill and Kilmacolm accept at least some minimal LDEs, and they all accept all of the doubled and elaborated LDEs, with any subject pronoun. We also see this in the atlas data. Of the participants who provided scores for both *here it's* and *here it's here*, 32% accepted neither, 40% accepted both, 25% accepted *here it's here* only and just 3%, 13 participants, accepted *here it's* only. Of the latter group, all were outside the main dialect area, and 9 gave *here it's here* a score of 3 ('I might say that').[10] This number is small enough to consider as noise.

An important aspect of this entailment relationship between minimal LDEs and the other kinds that we should draw attention to is that it operates independent of the pronoun-based restriction. This is shown by the Kilmacolm in-depth questionnaire data, where minimal LDEs are possible only with *it's*, but a full range of repeating LDEs are accepted. This is also observable to some degree in the atlas data: there are several locations (in particular in the Borders) where doubled LDEs are accepted with *it's* and *I'm*, but minimal LDEs are only accepted with *it's*. Thus we can say that accepting *I'm*-based minimal LDEs entails accepting all doubled and elaborated LDEs.

OBSERVATION 4. Acceptance of the nonstandard LDEs is highest in and around the greater Glasgow area, and scores are generally lower the further one travels from there. This is clearest with minimal LDEs, as we

---

[10]This can be compared with the group who accepted *here it's here* but not *here it's*, which were much more decisive: 84% gave *here it's* a score of 1 or 2. In general 3 was used infrequently and typically in cases where the speakers were hesitant or unsure.

can see in Figures 3 and 4. In Figure 3, we see that *here it's* is accepted in almost all locations around Glasgow, and there are more grey dots as we go further from the city (e.g. the southwest, the south coast). In Figure 4, we see that most of the locations where *here I'm* is accepted are around Glasgow (west of the central belt area), with only a few in the Fife area (north of Edinburgh). We can also see this in Figure 5, which combines the average scores for all of the nonstandard LDEs (minimal LDEs and doubled LDEs with *it's* and *I'm*). The darker spots which indicate higher average scores are centred mainly on Glasgow.
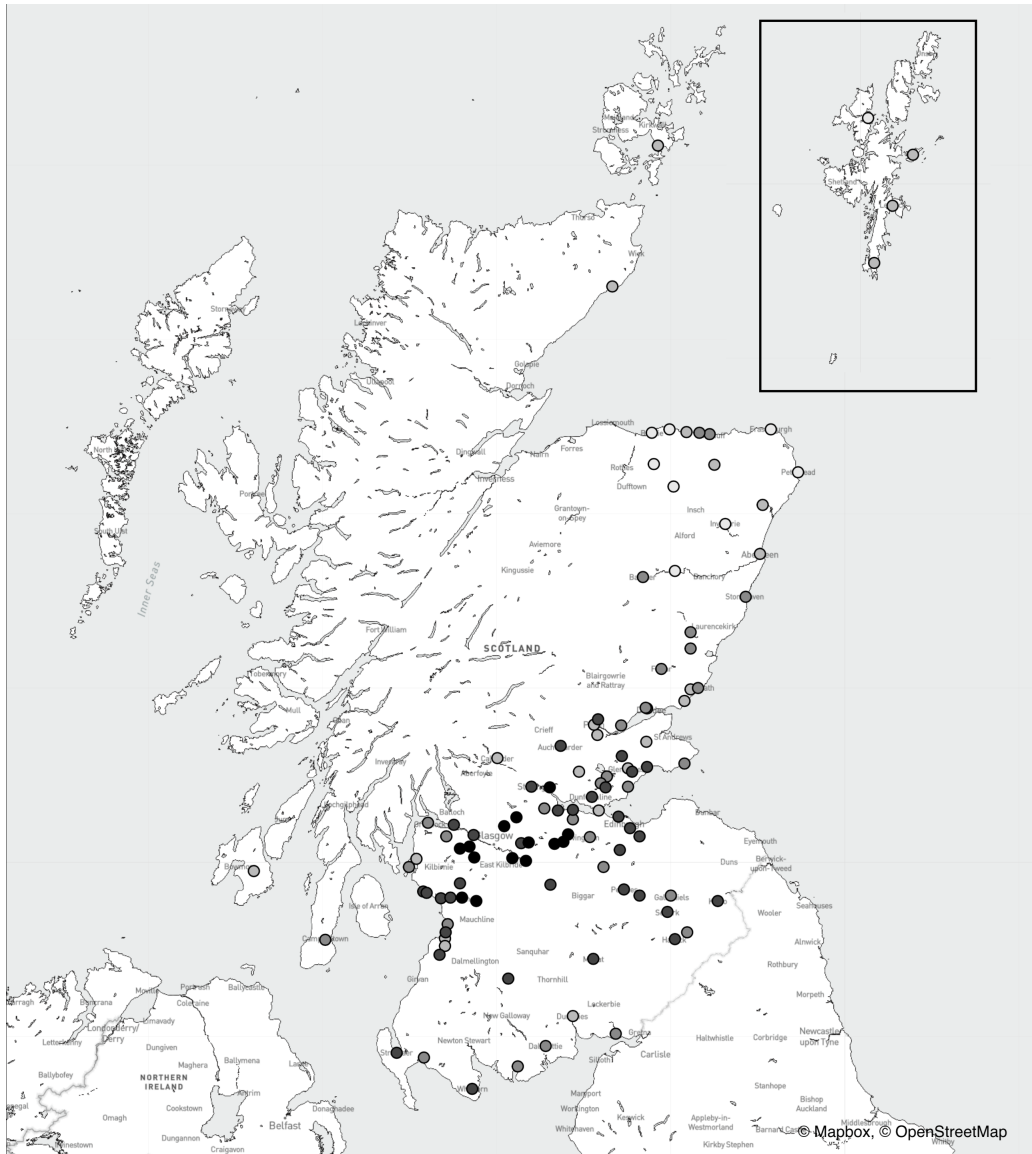

Figure 5: av. scores for all nonstandard LDEs

3.4. Discussion. With the facts laid out, let us now outline the general characteristics of the variation we see with the contraction data and how it impacts upon our analytical options. First, the data is generally very structured. While there are very many *a priori* possible systems, ultimately we can describe what we find in terms of four distinct systems. There are two fully productive systems which allow for some minimal LDEs and all repeating LDEs: one which allows LDEs with all pronouns, which we find in Glasgow, and another which only allows *it's*-based minimal LDEs, which we find in Kilmacolm. We also find a semi-productive system in

Dundee, where all repeating LDEs are accepted but no minimal LDEs are. Finally, the system in Buckie can be described as an unproductive one, since it only allows for doubled LDEs with *it's*.[11] Table 4 summarizes the results in a more condensed form.

	Glasgow	Kilmacolm	Dundee	Buckie	Dunrossness
productive minimal LDEs	✓	*	*	*	*
limited minimal LDEs (*it's* only)	✓	✓	*	*	*
productive doubled LDEs	✓	✓	✓	*	*
elaborated LDEs	✓	✓	✓	*	*
limited doubled LDEs (*it's* only)	✓	✓	✓	✓	*

Table 4: Summary of LDE systems

There are numerous types of variation that we might in principle have observed but which we do not find in the data. For instance, we do not find variation with respect to the initial element, as noted in observation 1: acceptance of a *here* form entails acceptance of the *there* form and vice versa. Likewise we might also expect to see many different systems with respect to acceptance of the different LDEs with the pronoun-auxiliary pairings, for instance ones where contraction is accepted with 3S forms generally (including *he's* and *she's*), or only with singular forms, but we do not find this. Finally, we may note that there is also a degree of geographic order to the distribution of systems across the map, as those locations where the variation is less well-behaved tend to be at the edges of the broad region where the core examples are accepted. For instance, the handful of locations where it is not the case that both *here* and *there*-type LDEs are accepted (the potential noise mentioned in the discussion of observation 1) are located at the periphery of the nonstandard constructions' "core region," i.e. the area where the high scores are concentrated. These peripheral areas can be understood as transition zones with respect to the relevant linguistic phenomena (see Chambers & Trudgill 1998, ch.8), and so it is unsurprising that in these locations we would find that their systems are more variable and less regular. Indeed this is what we found in the Fife area, which shows quite mixed results in general: while minimal LDEs are attested there, we also find that some speakers (including our non-naive informant from Glenrothes) had only quite restricted doubled LDE systems.

4. Explaining the gap restriction and its apparent exceptions in Scots. In this section we provide an analysis of the gap restriction and its apparent exceptions in the Scots data. First we analyse standard LDEs as mirative fronting constructions. Second, we develop an analysis of the most productive system of nonstandard LDEs, the one found in the Glasgow dialect, and we show how it predicts the exceptional contractions but still excludes contraction in the other gap contexts; for this we work with a version of the approach to the gap restriction proposed by Bresnan (1978) and later revived by Wilder (1997). Third, we show that the other productive Scots systems can be modelled as minimally different versions of the Glasgow system. Finally, we outline an analysis of the unproductive system found in peripheral areas, which is distinct from the productive one in important ways.

4.1. Standard LDEs. As noted at the end of section 2, we propose that LDEs are instances of *mirative fronting*. Mirativity was first introduced in the typological literature by DeLancey (1997). DeLancey defined mirativity as 'the grammatical marking of unexpected information', and argued that typical mirative constructions convey an inference that the speaker's expectations have been exceeded in some way or another. De-Lancey argues that mirativity is a grammatical category in its own right, distinct from evidentiality and other

---

[11]We include Dunrossness, Shetland Isles, as an indication that not all of Scotland's varieties allow for nonstandard LDEs. The Dunrossness dialect differs from other non-Scots varieties in many other ways.

such means of encoding epistemic information, and since his seminal paper various grammatical reflexes of mirativity have been identified in a number of different languages (see Aikhenvald 2004, 2012) and refinements of the exact nature of the mirativity inference have been proposed in the theoretical literature. In this paper, we follow Mexas (2016), who, building on Aikhenvald (2012), argues that the mirative category is best understood as marking realization: a transition in the epistemic states of the interlocutors from not-knowing the content of the mirative marked proposition to knowing it; for Mexas, the strength of the unexpectedness/surprise component of mirative meaning comes about via pragmatic inferencing. This characterization fits particularly well with the discovery inference that is attached to LDEs, since ultimately the primary function of these expressions is to obligatorily mark a sharp change in information about the speech situation. The strength of the unexpectedness component of LDEs arises from the pragmatics.

We propose that the mirative inference of LDEs is encoded in a specialised mirative complementizer, C_{MIR}, which is typically null. We analyse the movement of the locative as cliticization to the C_{MIR} head, rather than movement to a specifier position, and so the movement rule involved is more like the cliticization of locative proforms in Iberian Romance than focus movement in Hungarian and Italian (cf. Roberts 2010). A cliticization analysis allows us to account for the fact that LDEs only allow fronting of the proforms *here* and *there*, since only minimal categories like proforms may undergo cliticization.[12] We propose that C_{MIR}, in addition to its semantically interpreted mirative feature, also bears a syntactic feature which requires it to establish a syntactic dependency with a locative; call this a [Loc] feature. Cliticization of a locative to C_{MIR} satisfies this feature's requirements, but it might also be satisfied by other mechanisms. The structure of a simple example with the order given in 12, repeated here as 40a, is in 40b.[13]

(40)　　a.　　*t/here* pronoun$_{nom}$ be$_{fin}$

b.
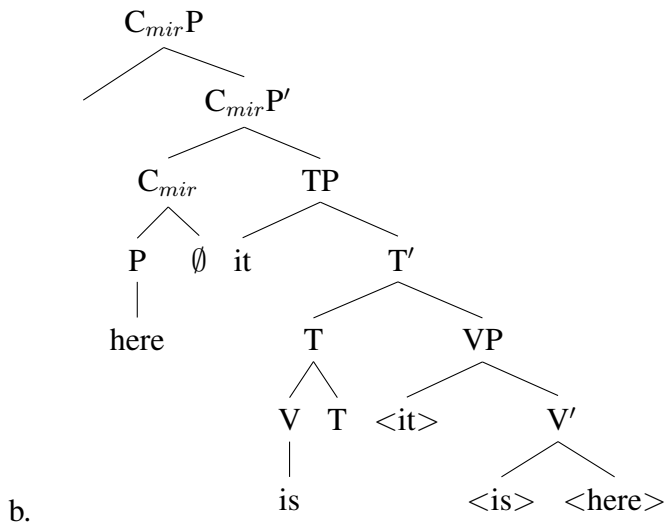
Since the locative undergoes movement and leaves finite *be* at the edge of a movement gap, we expect that structures of the format in 12 will be subject to the gap restriction with respect to auxiliary contraction, and indeed we have already seen that this is the case in non-Scots varieties. In section 4.3 we will explain why it does not hold in the nonstandard LDEs in Scots. But first we need to expand on some relevant aspects of the analysis and how they account for the properties of LDEs identified in Section 2.

---

[12]The fact that the fronted category is a proform that undergoes cliticization does not necessary mean that it is a phonologically weak clitic which may not bear focal stress, as there are clitics in other languages which may bear focal stress. For instance, Irish incorporated subject pronouns are stressed (Bennett et al. 2019), and Slovenian pronominal clitics can be stressed in certain circumstances (Dvořák 2007).

[13]Angled brackets indicate lower copies of movement, what in earlier stages of the theory would have been analysed in terms of traces. The adoption of the Copy Theory of Movement becomes important in what follows; see section 4.3 in particular.

The analysis above accounts for the restricted nature of the initial element: the cliticization analysis accounts for the fact that the initial locative must be a proform, and positing a [Loc] feature on C_{MIR} ensures that we do not predict a wider range of mirative fronting phenomena than we find. We propose that the restrictions on the clause below the locative follow from the mirative component of the construction's meaning and general pragmatic considerations. The fact that LDEs are typically made up of a minimal clause follows from the fact that the extra content contributed by addition of a modifier or an embedding predicate would conflict with a report of an immediate realization based on directly perceptible evidence concerning the location of the subject. This is clear with the addition of verbs like *seem*: a report of sudden realization based on directly perceptible evidence is usually incompatible with *seem*'s lexical meaning that the attitude holder's evidence is indirect and based on inferencing; The pragmatic oddness of *he suddenly seemed to be here* follows from this.[14] The same logic extends to the cases with adverbs like *definitely* or *predictably*: these indicate there is some epistemic uncertainty, which is inconsistent with the inference that a sudden discovery has been made on the basis of directly perceptible evidence. This pragmatic explanation allows for exceptions where the adverb contributes little in the way of additional information about the event, as is the case with the use of *suddenly* in past tense LDEs, since one can retrospectively describe a discovery as sudden.

The restriction against focussed subjects follows from the fact that focus involves generating focus alternatives and negating them (Rooth 1985): that is, *here HE is* involves computing the alternative *it's not the case that here some other person is*. These examples are degraded for the same reason that *here he isn't* is degraded: the mirative reports on directly perceptible evidence (see also Rett 2011 on the incompatibility of negation with exclamatives). The same logic extends to other non-discourse-old subject types like indefinites, which give rise to scalar inferences which also involve computing negated alternatives.[15] The fact that LDEs are only possible with the copula and low-information locational verbs such as *put*, *come* and *go* can also be explained in this way. The use of more informationally rich verbs like *run* contributes information about manner of motion, much like an adverb, and hence these are excluded for the same reason that manner adverbs are excluded, while low-information lexical verbs lack this component and so are less likely to clash with the discovery inference.[16]

The fact that an additional PP may occur in the postverbal position (in what we have called elaborated LDEs) is to be expected since the only contribution of these PPs is to elaborate on the precise location of the discovery. These examples can be derived from the simple cases in 41, where there seem to be two locative predicates appearing together in the postverbal position.

(32) a. Here it is on the shelf.
      b. There it is in the corner.

(41) a. It's here on the shelf.
      b. It's there in the corner.

In the spirit of Kayne (2005a), Collins (2007) and Svenonius (2010) we assume that *here/there* is a PP which

---

[14]An anonymous reviewer suggests that the verb *seem* is in fact compatible with reports of sudden realization, such as in *It suddenly seemed that we were going to win*. Our pragmatic story is consistent with this observation, as the example is in the past tense and reports on a complex event; sudden realization is pragmatically plausible in such a circumstance.

[15]Note that this leaves open the possibility that discourse-new subjects are semantically compatible with LDEs; that is, they would be possible if they did not give rise to focus alternatives. Focus which does not involve alternative computation is known as 'information focus' (Zubizarreta 2001), and quite typically placing information focus on the subject in English involves putting it in a postverbal position (Birner 1992). This may be why LDEs with discourse-new subjects are of the locative inversion-type order, i.e. *here's John*, and the general ban on putting unfocused pronouns in that position would follow from the fact that pronouns cannot be in information focus. Proper consideration of this matter would take us too far afield here.

[16]The mode of explanation here is broadly in line with Levin and Rappaport Hovav's (1995, ch.5) account of restrictions on the verbal predicate in English locative inversion clauses, such as the fact that it is most commonly used with unaccusative verbs, and impossible with transitive verbs.

occurs as the specifier of a larger PP which take a complement denoting a specific location (see also Cresti & Tortora 2000), and we suggest that the contribution of *here/there* is to signal whether that location is proximal or distal. 42 illustrates the structure for 41a.
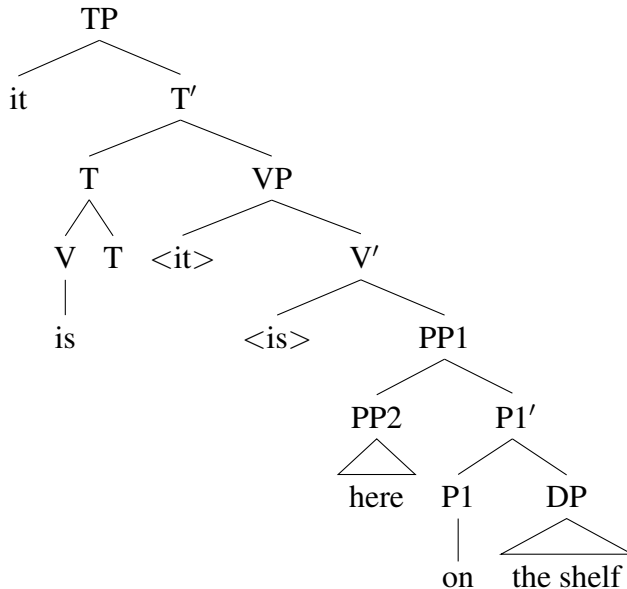
(42)

The fact that most speakers of English disallow contraction in Elaborated LDEs, as shown in 33, repeated here, is thus of a piece with the fact that predicate extraction can still lead to a gap restriction problem even if there is an adverbial to the right of the gap, as in 3b.

(33) a. *Here it's on the shelf.
      b. *There it's in the corner.   StE

(3b) *I don't know where the bathroom's ___ in this building.

To summarize, the various restrictions on standard LDEs follow from the combination of two factors: the cliticization that is part of the fronting rule involved, and the semantic/pragmatic properties of the mirative inference which defines the construction. Having set out the analysis of standard LDEs, we will now turn to nonstandard LDEs, beginning with the most productive system, that of Glasgow.

4.2. Nonstandard LDEs in Glasgow Scots.

REPEATING LDEs. We begin with repeating LDEs, of which there are two kinds, doubled LDEs and elaborated LDEs. Recall that doubled LDEs involve a second locative occurring in the postverbal position which is identical to the initial one, as in 43, and elaborated LDEs involve a second locative which is non-identical, for instance a PP or a modified version of the initial locative, as in 44.

(43) a. Here it's here!
      b. There you're there!
      c. Here I'm here!

(44) a. Here it's over here!
      b. There it's in the corner!

The fact that these are unacceptable with contraction for speakers of non-Scots varieties (and indeed some va-

rieties of Scots, e.g. Shetland Scots) indicates that such speakers can only derive such structures by movement, as with standard LDEs like *here it is*. It must therefore be the case that some alternative analysis is available for Glasgow Scots speakers. We propose that the key factor is that in Glasgow Scots the initial *t/here* in 43 is a pronunciation of the mirative complementizer C_{MIR}. That is, in structures like 43 the two overt locatives are base-generated separately, and they are not the same category: the postverbal locative is a regular PP predicate, and so the core of the clause is largely the same as with regular locative clauses, but in the case of Glasgow Scots the initial *here/there* is an overt realization of C_{MIR}. 45 provides a structure for a simple example like *here it's here*.
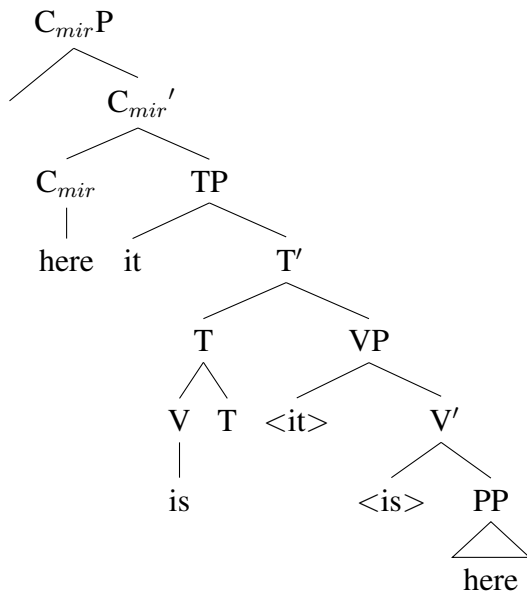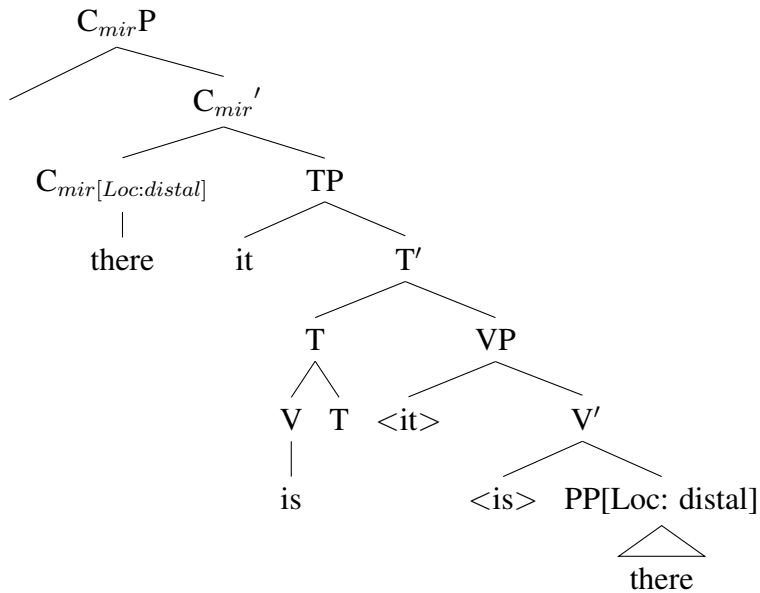

(45)

The overt C_{MIR} in this structure has a similar feature specification to its covert counterpart in non-Scots varieties (i.e. 40b), as it bears a [Loc] feature which must be valued by entering into a dependency with another locative, but this need not be done by movement: some kind of agreement dependency would suffice, for instance using the probe-goal technology of Chomsky (2000) (see Déchaine & Wiltschko 2014, and Zu 2018 for comparable cases where complementizers agree with clause-internal elements). We take the two overt forms of the overt C_{MIR} to correspond to different values of the [Loc] feature, specified for whether it is distal or proximal.[17] The relevant Scots varieties thus have two exponents for C_{MIR}, given in 46a-46b. A tree for a *there*-based version is given in 47.

(46) a. $C_{MIR} \rightarrow$ [hir] / [proximal] Glasgow Scots
      b. $C_{MIR} \rightarrow$ [ðɛr] / [distal] Glasgow Scots

---

[17]Tortora (1997, 2001, 2014) proposes that the correct feature specification on null locative expressions in Romance varieties involves [±speaker] rather than proximal/distal. We acknowledge this is a possibility here too, but keep to a more familiar feature specification to simplify presentation.

(47)

```
                    C_mir P
                   /      \
                    C_mir ′
                   /       \
        C_mir[Loc:distal]   TP
              |            /   \
            there        it     T′
                               /   \
                              T      VP
                             / \    /   \
                            V   T <it>   V′
                            |          /    \
                           is       <is>  PP[Loc: distal]
                                            /\
                                          there
```

Elaborated LDEs are derived from very similar structures, where there is a base-generated overt C_MIR in the left periphery and a syntactic dependency connects the C_MIR to the lower locative, determining its form: *there* if [distal], *here* if [proximal]. The only difference is that the higher locative (*here/there*) does not have the same phonological form as the lower locative it establishes a dependency with (e.g. *in the corner*).
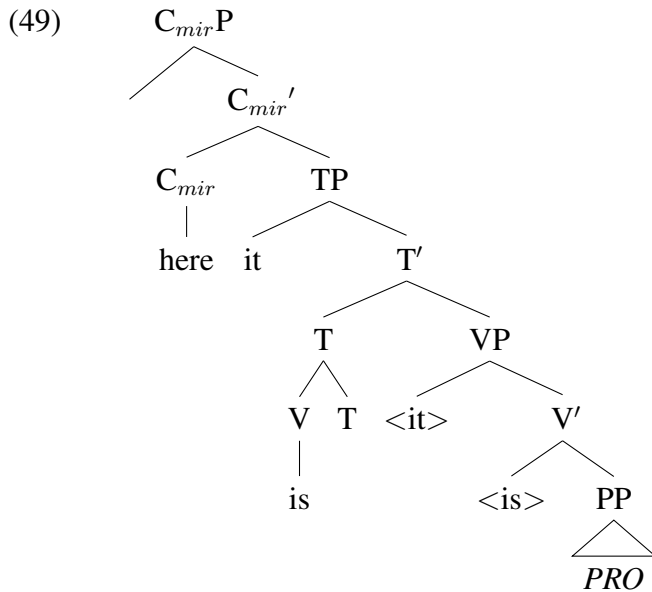
(48)

```
                    C_mir P
                   /      \
                    C_mir ′
                   /       \
        C_mir[Loc:distal]   TP
              |            /   \
            there        it     T′
                               /   \
                              T      VP
                             / \    /   \
                            V   T <it>   V′
                            |          /    \
                           is       <is>  PP_[Loc : distal]
                                            /\
                                       in the corner
```
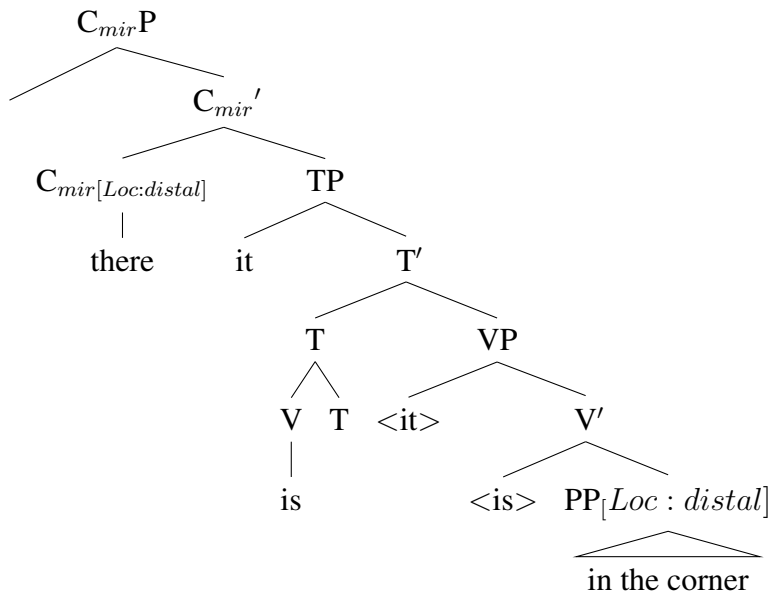
The fact that auxiliary contraction is possible in doubled and elaborated LDEs with this analysis is unsurprising, since there is no movement from the post-auxiliary position. In contrast, most other varieties will always derive such structures with movement of *here/there* to the initial position, even when there is a locative PP predicate, and hence will disallow contraction.

MINIMAL LDEs. Turning now to minimal LDEs – those with no overt material following *be* – our starting point is the observation that the possibility of minimal LDEs entails the option of doubled LDEs, one of the most striking empirical findings in section 3. We claim that this is so because minimal LDEs are a subspecies

of doubled LDEs, with the same gross structure and an identical structure for the left periphery. Specifically, minimal LDEs are repeating LDEs but with a **null locative proform** in the predicate position (cf. Tortora 1997, 2001, 2014). This null locative proform, PRO$_{loc}$, which bears a specification for [proximal/distal], is an additional innovation which is specific to a subset of varieties of Scots, including Glasgow Scots. It is the availability of this proform plus the overt C$_{MIR}$ which makes minimal LDEs an option. Any variety which lacks either component will not be able to derive minimal LDEs.

(49)

```
                C_mir P
               /      \
                    C_mir'
                   /      \
             C_mir          TP
              |            /   \
            here        it       T'
                               /    \
                             T        VP
                            / \      /   \
                          V   T  <it>     V'
                          |             /    \
                         is          <is>     PP
                                              /\
                                            PRO
```

A key claim here is that PRO$_{loc}$ does not count as a gap in the same way that a movement trace or an ellipsis site does, and so it does not obey the gap restriction.[18] We discuss exactly why traces and PRO would differ in this respect in the next subsection, but before we do that, a few comments are in order regarding the derivation of minimal LDEs.

As the C$_{MIR}$ is the same as the C involved in repeating LDEs, the relevant feature on the complementizer will create a syntactic dependency with PRO$_{loc}$. In the case of minimal LDEs, this dependency is required to identify the locative proform. We use 'identify' in the sense of Rizzi (1986) and much subsequent work, where null proforms are only licensed when the features that they encode are recoverable from some other head in the structure; in the domain of pro-drop, this is the requirement that verbal morphology be sufficiently rich to allow speakers to infer the phi-specification of the null pronoun. Following Sigurðsson (2011), we take identification to be a syntactic requirement that the proform and its identifying category enter into a syntactic dependency and that the linked elements must share the relevant distinctive features required for identifying the category. In the case of LDEs, this is done by the Loc feature of C$_{MIR}$, which sets up a dependency with the locative proform. The *here/there* form of C$_{MIR}$ overtly signals whether this feature has [proximal] or [distal] as its value. If C$_{MIR}$ is not present, the null locative proform will not be licensed, much like a null subject would not be licensed in the absence of agreement features. Anticipating the discussion to come somewhat, this allows us to explain the fact that Glasgow Scots speakers do not allow the null locative proform to be used in various other constructions where we could in principle imagine inserting the null locative in place of an elided VP, such as 2b above, repeated here:

(2b)   *Mary's usually there when Kim's.

---

[18]The distinction between traces and PRO here is similar to that which is drawn in Chomsky and Lasnik's (1977) classic account of *wanna* contraction (but see Postal & Pullum 1978, 1982, Pullum 1997).

Thus we claim that it is not the case that null locative can be used in any context, but rather its presence is tied to the presence of a specific, independently motivated, functional category in the structure,[19] and so the option of minimal LDEs is restricted to those grammars where both C_{MIR} and PRO_{loc} are available.

With these basic analyses established, We now return to the question raised earlier of why a trace of movement and a null locative PRO may differ with respect to auxiliary contraction.

4.3. An analysis of the gap restriction and its exceptions in Glasgow Scots.
In this section we will first set out our assumptions about how the gap restriction on auxiliary contraction is derived. With this as background, we will then turn to an analysis of the apparent exceptions to this restriction in the dialect where these exceptions are the most extensive: Glasgow Scots. First, we show how the assumption of a null PRO_{loc} allows minimal LDEs to be derived; we then discuss the availability of clitic subject pronouns, and show how this plays into the analysis, and finally we show how the analysis handles repeating LDEs.

DERIVING THE GAP RESTRICTION. The core of our analysis of the gap restriction on auxiliary contraction is that contracted auxiliaries are clitics which must be prosodically incorporated into a prosodic host in their immediate context. If the clitic is incorporated rightward, and its host is subsequently deleted, the clitic will be stranded, leading to ungrammaticality.

Clearly, such an explanation requires specific assumptions concerning prosodic incorporation. We follow various authors in taking this to be a postsyntactic rebracketing rule, which manipulates the structure of the sentence in the prosodic component of the grammar to ensure that clitics are grouped with an appropriate host (Zwicky 1982, Nespor & Vogel 1986, Hayes 1989, Selkirk 1996, Erteschik-Shir 2005, Clemens 2014, Bennett et al. 2019). Prosodic incorporation can be thought of as a rule which sorts out the placement of clitics following an initial pass of the algorithm for converting syntactic structure to prosodic structure, which first parcels the structure up into phonological words. Clitics are by definition elements which do not form phonological words on their own, and so in order to ensure that the structure is exhaustively parcelled up prior to the next stage of the derivation, clitics are incorporated into an appropriate host. Hayes (1989) calls this stage CLITIC GROUP FORMATION, and prosodic incorporation is the rule which does the work of putting clitics into groups. 50 presents how this would be done for a simple clause with a contracted auxiliary: the clitic *'s* is left stranded once p-words are formed, and so it is incorporated rightward onto the following predicate *here* to form a Clitic Group with the host word to its right.

(50) He 's here   → parse into p-words
      ($_{PWD}$ He) *'s* ($_{PWD}$ here)   → prosodic incorporation of *'s*
      ($_{PWD}$ He) ($_{CLG}$ *'s* ($_{PWD}$ here)

We adopt Hayes' (1989) approach for concreteness here, but other theories which treat the prosodic hierarchy differently would suffice.

In order to explain how this prosodic incorporation leads to the gap restriction, it is also necessary to consider exactly when it takes place. We take the non-pronunciation of lower elements of a movement chain (the creation of a 'gap' resulting from movement) to involve deletion, henceforth 'copy deletion', in line with Chomsky (1995) and much work since. It is essential for our account of the gap restriction – with standard LDEs, *wh*-questions, VP-ellipsis and so on – that prosodic incorporation must precede this deletion. That is, the

---

[19]We should note that some of the examples which motivate the gap restriction might be ruled out by additional factors which do not relate to our account. For instance, Pullum & Zwicky (1997) argue that some examples of auxiliary contraction adjacent to ellipsis sites are bad because independent properties of those configurations require placing light accent on the auxiliary. While their account has some promise, it is hard to see how it could be extended to cover our data, without buying into core details of our analysis relating to differences between movement and non-movement-derived dependencies.

gap restriction arises when prosodic incorporation groups a clitic auxiliary with a host which is subsequently deleted, stranding the auxiliary.[20]

(51)   *I don't know where he's
        ... where he 's where $\rightarrow$ parse into p-words
        ($_{PWd}$ where) ($_{PWd}$ he) 's ($_{PWD}$ where) $\rightarrow$ prosodic incorporation of 's
        ($_{PWd}$ where) ($_{PWd}$ he) ($_{ClG}$'s ($_{PWd}$ where) $\rightarrow$ copy deletion
        ($_{PWd}$ where) ($_{PWd}$ he) ($_{ClG}$'s (~~$_{PWd}$ where~~)) $\rightarrow$ clitic stranded!

There is independent evidence to believe that deletion occurs late in the postsyntactic component. Indirect support comes from recent proposals in which subsequently deleted copies of movement are shown to interact with suprasegmental phonological processes such as stress (Ahn 2015) and tone assignment (Korsah & Murphy 2016). These proposals motivate a theory where copy deletion is late in the postsyntactic component, at least following linearization and vocabulary insertion, and given that both processes have been shown to interact with articulated prosodic structures (see e.g. Chen 1979 on the sensitivity of tone sandhi rules to prosodic structure) we may conclude that prosodic structures must be parcelled up before both stress/tone assignment *and* copy deletion. More direct support comes from Thoms & Sailor (2018), who argue that the restrictions on extraction from 'British *do*' follow from the fact that encliticization of clitic auxiliaries is interrupted by copies of movement.

Note that we assume that clitic auxiliaries incorporate rightwards by default. This assumption is adopted from Wilder (1997), and it has precedent in Bresnan (1978), who took the attachment rule in question to be syntactic (working with an earlier framework). At first blush, this is an unintuitive analysis, as the auxiliary seems to form a word with the preceding subject.[21] Wilder (1997) addresses this by analysing auxiliary contraction as a two-stage process: first, the auxiliary is procliticized rightwards, and then it is encliticized onto the preceding subject. We essentially follow Wilder in taking auxiliary contraction to be a two-stage process, although we differ in the details. On our analysis, the first stage is incorporation of the clitic auxiliary, and the second stage is rebracketing of the subject with the following constituent. If the subject is a clitic pronoun, it may also be incorporated rightward into the complex formed by the auxiliary and the predicate.[22]

---

[20]The same result is achieved by Wilder (1997), who also takes rightward prosodic incorporation of clitic auxiliaries to be responsible for the gap restriction. However his account does not explicitly adopt the assumption that prosodic incorporation precedes copy deletion, presumably because he does not entertain the possibility that clitic auxiliaries can cliticize leftward if they do not find a host rightward, as we do below; rather, he simply says that they cliticize rightward, and if there is nothing in the following context then the clitic is stranded.

[21]For an alternative approach see Anderson (2005, 2008), who proposes that contracted auxiliaries are clitics which attach leftward, with no rightward cliticization involved at all. Anderson proposes that the gap restriction comes about because leftward cliticization leaves the phonological phrase to the right of the subject without any phonetic content, in violation of some general constraint against vacuous phonological phrase formation. Interesting though it is, Anderson's account cannot be straightforwardly applied to account our of the Scots data, since it has little scope to allow the abstract structure of the post-auxiliary context to play a role in conditioning the availability of auxiliary contraction.

[22]Strong pronouns and phrasal subjects will behave differently: both project their own p-phrase and so they cannot be rebracketed by prosodic incorporation. We propose that sentences with such subjects are parsed by recursive p-phrasing of the predicate, with the first pass assigning the subject and predicate distinct p-phrases and the second pass grouping them in one larger p-phrase. A simplified example is given below.

(52)   Peter Pan's here $\rightarrow$ parse into p-words
        ($_{PWd}$ Peter) ($_{PWd}$ Pan) 's ($_{PWd}$ here) $\rightarrow$ prosodic incorporation of aux
        ($_{PPH}$ ($_{PWd}$ Peter) ($_{PWd}$ Pan)) ($_{PPH}$ ($_{ClG}$'s ($_{PWd}$ here)) $\rightarrow$ p-phrase formation
        ($_{PPH}$ ($_{PPH}$ ($_{PWd}$ Peter) ($_{PWd}$ Pan)) ($_{PPH}$ ($_{ClG}$'s ($_{PWd}$ here)))

This derives different representations for different subject types, and it leads us to expect differences between contracted auxiliaries attached to pronouns, which are in a very local configuration with the auxiliary in cases such as 53, and contracted auxiliaries

(53)   It 's here.   → parse into p-words
       It 's ($_{\text{PWd}}$ here)   → prosodic incorporation of pronoun and aux
       ($_{\text{CIG}}$ It 's ($_{\text{PWd}}$ here))

Having set out these basic assumptions about the derivation of the gap restriction, we are now in a position to consider the apparent exceptions, considering first of all Glasgow Scots, since this is the dialect where there is the widest range of such cases.

MINIMAL LDES IN GLASGOW SCOTS.   As discussed above in section 4.2, one crucial component of our analysis of the apparent 'exception' to the gap restriction found in Glasgow Scots minimal LDEs (*Here it's!*) is that here the context following the clitic is a null proform—what we have called PRO$_{loc}$. In contrast to both ellipsis and traces of movement, PRO$_{loc}$ is not null because it undergoes deletion; rather it never has phonological features. This entails that there is no prosodic structure to the right of the contracted auxiliary when PRO$_{loc}$ is generated in this position, and so there is no host for rightward incorporation of the auxiliary.

As mentioned earlier, we assume that contracted auxiliaries can cliticize leftward when they do not find a host to their right, as in this case.[23] Here there is no host for prosodic incorporation and so the clitic is incorporated leftward instead. More specifically, we claim that the auxiliary is incorporated leftward *with the subject pronoun* into the initial locative, which serves as the host.

(54)   Here it 's PRO   → parse into p-words
       ($_{\text{PWd}}$ Here) it 's   → no host rightward, so incorporate leftward
       ($_{\text{CIG}}$ ($_{\text{PWD}}$ Here) it 's)   no clitics stranded!

One reason that we take the locative to be the host for incorporation here is that the subject in this case is *it*, which, as a clitic itself, is not eligible to serve as a host for other weak elements.

There is evidence which supports the idea that the auxiliary prosodically incorporates leftward. This evidence comes from the fact that minimal LDEs are sensitive to the structure of the subject preceding the contracted auxiliary. In particular, when the subject is an epithet, rather than a clitic pronoun, it turns out that minimal LDEs with auxiliary contraction are not acceptable in Glasgow Scots, as shown in 56. The contrast with the acceptable LDES in 55, which are identical except for the lack of contraction, shows that indeed it is the impossibility of the full NP allowing the auxiliary to cliticize leftwards that causes the unacceptability.

(55)   a.   There the wee bugger is!
       b.   Here the bastard is!

(56)   a.   *There the wee bugger's!
       b.   *Here the bastard's!   Glasgow Scots

We take this to indicate that the leftwards prosodic incorporation rule is restricted to incorporating weak elements only into heads in the spine of the clause. Incorporation into phrasal elements is disallowed (see Hayes (1989) for a precise formulation of the conditions on prosodic incorporation which has this effect). The sensitivity of minimal LDEs to the phrase structure of the preceding subject provides support for the assumption that the auxiliary is incorporated leftwards. Moreover, the fact that phrasal subjects do not generally preclude

---

attached to phrases, which are separated from the auxiliary by a prosodic boundary, although they are both still contained within a single prosodic domain. As it happens, we do indeed find such differences, as is noted by Zwicky (1970), Kaisse (1985) and Close (2004): many contracted auxiliaries (such as *'ll* and *'d*) only form a word with pronoun subjects.

[23]The idea that a clitic element may cliticize leftward after failing to find a host to its left is also entertained by Zwicky (1982) in a discussion of the behaviour of nonfinite *to* in ellipsis contexts. We leave it to future work to spell out exactly how our own analysis of the gap restriction can be made compatible with Zwicky's account of the *to*-facts.

auxiliary contraction (in particular with *'s – The car's here*) indicates that leftward incorporation is not the default: if it were, then we would expect that the gap restriction would be lifted whenever the subject is a phrase or a strong pronoun, contrary to fact (cf. 3b, **I don't know where the bathroom's in this building*).

THE CLITIC STATUS OF SUBJECT PRONOUNS IN GLASGOW SCOTS. One important aspect of the analysis in 54 is that the subject pronoun is itself analysed as a clitic and is incorporated into the preceding locative along with the auxiliary to form a single clitic cluster. Given that without this incorporation the subject is itself not an appropriate host for a clitic auxiliary, it is a necessary condition for a minimal LDE that the subject pronoun be analysable as a clitic. However, not all nominative subject pronouns in all varieties of English are clitics. It is not difficult to motivate a clitic analysis for *it*, since it cannot be stressed or coordinated, but as noted by Cardinaletti & Starke (1996, :24), the other pronouns in Germanic behave much like strong pronouns (in the sense familiar from studies of Romance varieties, e.g. Cardinaletti 1994), in that they can be coordinated and bear stress.

Cardinaletti and Starke's discussion of English pronouns focuses on standard varieties of English. When we turn to subject pronouns in Scots, we see that there is good evidence that some Scots varieties have a more Romance-like pronoun inventory, in that they have a set of distinct strong and weak subject pronouns. In Glasgow Scots, there are regular pronominal forms, which are much like in all other varieties of English, and there is also a set of reduced versions of *I, you, he, they* and *we* which differ either in vowel quality (more back) or in deletion of a consonant (as in the case of *he*). *She* has no distinct weak form, and *it* only has a weak form (as in more standard varieties). These facts are summarized in Table 5. In a neutral context, either the clitic or the strong pronoun can be used, although the preference is for use of the weak form.

Pronoun	strong	weak
I	[æ]	[ʌ]
you	[yu]	[yʌ]
he	[hi]	[i]
she	[ʃi]	[ʃi]
it		[ɪt]
they	[ðe]	[ðə]
we	[wi]	[wə]

Table 5: Glasgow Scots distinctive weak and strong pronouns

(57)　　a.　{[æm] / [ʌm] } leaving.
　　　　b.　{[yur] / [yʌr]} leaving.
　　　　c.　{ [hiz] / [iz] } leaving.
　　　　d.　{ [ðer] / [ðər] } leaving.
　　　　e.　{ [wir] / [wər] } leaving.

If the subject is focussed, only the strong form can be used, typically with some lengthening. The weak form is impossible.

(58)　　I don't know about John but ...

　　　　a.　{[æ:m] / *[ʌm] } leaving!
　　　　b.　{[yu:r] / *[yʌr]} leaving!
　　　　c.　{ [hi:z] / *[iz] } leaving!
　　　　d.　{ [ðe:r] / *[ðər] } leaving!

e.   { [wi:r] / *[wər] } leaving!

In minimal LDEs, only the weak form of the pronoun is possible. Full forms are not possible in minimal LDEs for Glasgow Scots speakers, whether lengthened or not.[24]

(59)   a.   Here { [ʌm] / *[æm] }
           b.   There {[yʌr] / *[yur]}
           c.   Here {[iz] / *[hiz] }
           d.   There {[ðər] / *[ðer]}
           e.   Here {[wər] / *[wir]}

Thus Glasgow Scots has both weak and strong forms of the nominative personal pronouns, and the versions which occur in 59 (in minimal LDEs) are the weak pronouns. These weak pronouns do not project p-words of their own, and so they are subject to prosodic incorporation; in other words, they are clitics. The unacceptability of the strong forms of the pronouns in 59 therefore supports our claim that the subject needs to be a clitic in order to survive in a minimal LDE. The auxiliary needs to prosodically incorporate into a head in the clausal structure (in this case the mirative complementizer *(t)here*). If another clitic intervenes, then both clitics can incorporate, leading to a well formed structure. If the subject is not a clitic, the auxiliary cannot prosodically incorporate, leading to the unacceptability of the examples with strong pronouns in 59. This account predicts that there may be variation with respect to whether a given pronoun may occur in minimal LDEs; we return to this shortly below.

REPEATING LDEs IN GLASGOW SCOTS. To complete our account of the Glasgow Scots system, let us turn back to repeating LDEs, in which the immediately postverbal position is filled by a locative expression, and show how the account of the gap restriction outlined above predicts that these would be acceptable for Glasgow Scots speakers but unacceptable for speakers of most other varieties of English.

(60)   a.   Here he's here.   doubled LDE
           b.   There it's in the corner.   elaborated LDE

Recall that for speakers of non-Scots varieties, structures such as these must be derived by movement of a locative from the postverbal position to the initial position. If there is movement from the immediately postverbal position, then there will be the same scenario as represented in 51 above: the contracted auxiliary will incorporate into the copy left behind by movement of the locative, and subsequent deletion of the lower copy of the locative will leave the auxiliary without a host. The presence of an additional PP in the following context makes no difference because it is not a viable host for prosodic incorporation at the point where prosodic incorporation takes place. 61 provides a schematic for how this would proceed in standard English (and many other non-Scots varieties).

(61)   *There it's in the corner.   e.g. standard English
            there it 's there in the corner   → parse into p-words
            ($_{PWd}$ there) it *'s* ($_{PWD}$ there) in the ($_{PWD}$ corner)   → prosodic incorporation
            ($_{PWd}$ there) ($_{ClG}$ it *'s* ($_{PWd}$ there) ($_{ClG}$ in the ($_{PWD}$ corner))   → copy deletion
            ($_{PWd}$ there) ($_{ClG}$ it *'s* ~~($_{PWd}$ there)~~) ($_{ClG}$ in the ($_{PWD}$ corner))   → clitic stranded!

There is no such problem in Glasgow Scots, because there is no movement from the postverbal position: the initial locative *there* is base-generated in C_{MIR}, and the locative that appears in the immediately postverbal

---

[24]We should emphasize that these two pronoun inventories are not distinguished in sociolinguistic terms, for instance along a vernacular vs standard dimension, as both are used in variation in vernacular speech.

position is a viable host, just as it would be in *it's in the corner*.

(62)   There it's in the corner.   Glasgow Scots
        there it 's in the corner   → parse into p-words
        ($_{PWd}$ there) it 's in the ($_{PWD}$ corner)   → prosodic incorporation
        ($_{PWd}$ there) ($_{ClG}$ it 's in the ($_{PWD}$ corner))   → clitic not stranded!

Thus, the crucial difference between, for example, Standard English, schematized in 61, and Glasgow Scots, as shown in 62, is the absence of a lower copy of *there* in 62.[25]

SUMMARY.   We account for the gap restriction by analysing contracted auxiliaries as clitics which are prosodically incorporated rightwards into the following context by a rule which takes place prior to copy deletion. The gap restriction arises from the opaque interaction of copy deletion and prosodic incorporation, and this does not happen when the following context is a null locative, since null elements are 'born silent' and thus cannot interact with any postsyntactic rules. We therefore expect the gap restriction to be lifted just when the following context is a null element, as we have claimed to be the case in Glasgow Scots LDEs.

4.4. Dialectal variation within LDEs.   So far we have provided an account of nonstandard LDEs in Glasgow Scots, which is the most productive of the systems observed. Recall from section 3 that there are at least three other systems in varieties of Scots.

1. Systems which allow minimal LDEs only with *it's*, but allow doubled and elaborated LDE's with all other pronouns. This is attested in Kilmacolm, and the atlas data suggests that similarly restricted systems are found in the Borders.[26] We call this the KILMACOLM SYSTEM.

2. Systems that do not allow minimal LDEs, but allow double and elaborated LDEs with all pronominal subjects. This is attested in Dundee, and the atlas data suggests that it may be attested elsewhere in the surrounding area. We call this the DUNDEE SYSTEM.

3. Systems that do not allow minimal LDE's, and allow doubled LDE's only with *it's*. This is attested in locations which are peripheral or remote to the dialect region where the productive systems are best represented, such as Buckie, Glenrothes and Thurso, and the atlas data suggests it may be found elsewhere in the east. We call this the PERIPHERAL SYSTEM.

The first two of these systems are both productive, as they allow (in the form of elaborated LDEs) for a wide range of novel nonstandard LDEs, while the third system can be boiled down to the availability of *here it's here* and *there it's there*. We analyse the near-productive systems in Kilmacolm and Dundee as variations on the productive system of Glasgow, while the unproductive peripheral system is treated in a distinct way.

THE KILMACOLM SYSTEM .   The restriction of minimal LDEs to *it's* clusters in the Kilmacolm system can be accounted for if we consider the role of leftward cliticization. In the previous subsection we proposed that the pronoun and contracted auxiliary are both required to cliticize leftward onto the initial *t/here* in minimal LDEs, and that this requires the pronoun to itself be a clitic pronoun and subject to prosodic incorporation. As we noted, Glasgow Scots has a distinct set of weak subject pronouns and only these forms occur in minimal LDEs.

---

[25]In both of these schematics we represent the preposition *in* as a clitic which also incorporates into the following nominal, much like in Hayes (1989), but this is not important.

[26]We have also confirmed the availability of this system in informal interviews with speakers from Castle Douglas (Dumfries and Galloway) and Kelso (the Borders).

In this respect, the Glasgow Scots pronominal inventory is like Italian, which has a full range of weak and strong pronouns. Another kind of system is the German and Standard English system, where there is only one weak pronoun, *es/it*, with all the other pronouns only having strong forms. Our hypothesis is that the Kilmacolm system is distinguished from the Glasgow system precisely in behaving like standard English varieties. It lacks special clitic forms, except for *it's*. Other subject pronouns in Kilmacolm may be phonologically reduced, but they are not clitics. This means that they project a phonological word (before reduction). The clitic auxiliary cannot then prosodically incorporate into a head in the clausal spine, as a phonological word intervenes.

(63) Here I 'm PRO   → parse into p-words
      ($_{PWd}$ Here) ($_{PWd}$ I) 'm   → no host to right or left for aux!

We should stress that this difference is currently speculative, as supporting it would require a full phonological analysis of the varieties, taking us beyond the scope of this paper. However, the variation we see here is just what is expected given the independently established nature of *it* as a clitic in English (Cardinaletti & Starke 1996).

Looking beyond minimal LDEs, the Kilmacolm and Glasgow Scots systems are predicted to be identical with respect to all other nonstandard LDEs. Doubled and elaborated LDEs do not involve leftwards incorporation in their analysis, so the form of the subject pronoun is predicted to be irrelevant in these cases. As can be seen in Table 3, this is the correct prediction: there is no subject pronoun effect with doubled or repeating LDEs in the Kilmacolm variety.

THE DUNDEE SYSTEM. The Dundee system lacks minimal LDEs of any kind but is otherwise productive, allowing all kinds of elaborated and doubled LDEs. We propose this is a difference in syntactic inventory: whereas Glasgow Scots has the null locative proform, PRO$_{loc}$, this is absent from Dundee Scots.[27] Without PRO$_{loc}$, Dundee Scots cannot generate minimals LDEs. Otherwise it has the same overt C$_{MIR}$ as Glasgow Scots, so all doubled and elaborated LDEs may be derived in the same way as they are in the Glasgow system.

As with the Kilmacolm system, the subject pronoun inventory facts are irrelevant for the derivation of doubled and elaborated LDEs, since the auxiliary leans rightward onto the following locative predicate and this option does not rely upon the subject being a clitic;[28] rather, when the auxiliary leans rightward, the subject can be a full DP. That this is the case with doubled LDEs in these dialects is confirmed by examples such as 35f, repeated here, which are acceptable for Dundee speakers. (The same judgments hold for speakers with the more productive systems as well.)

(64) There the bastard's there.   Dundee Scots

If the auxiliary was leaning leftwards, as it is in our analysis of minimal LDEs, then the acceptability of such examples would be unexpected, since the non-pronominal subject could not serve as a host for the clitic nor could it be cliticized leftwards along with it.

THE PERIPHERAL SYSTEM. Finally there is the peripheral system, which allows doubled LDEs with *it's* but no other nonstandard LDEs. This was attested in a few different locations which were peripheral to the main dialect region, with Glenrothes being closest. LDEs in this dialect are therefore unproductive, suggesting an

---

[27]An alternative analysis is that Dundee Scots has PRO$_{loc}$ as part of some generally available inventory of null elements, but it is deficient, in that it does not bear a feature specification for [proximal/distal], and so it therefore cannot be licensed. Such an analysis is more in the spirit of contemporary work on null categories (e.g. the references cited in the discussion of licensing above), where linguistic variation with respect to null categories is explained in terms of features on functional heads rather than variation in in the inventory of null roots (but see Kayne 2005b).

[28]Thus it is possible that speakers of the Dundee dialect do not have a full range of weak pronouns.

alternative analysis is required.

(65)   a.   Here it's here.
       b.   There it's there.   doubled LDEs with *it's*
       c.   *Here I'm here.   no doubled LDEs with *I'm*
       d.   *There you're there.   no doubled LDEs with *you're*
       e.   *There it's in the corner.   no elaborated LDEs with *it's*

Given the peripherality of these locations with respect to the main dialect region, we might expect that the nonstandard LDEs are generally less common in speech in these areas, with only the most frequent versions being attested in the everyday speech that forms the input for learners, which is the *it's*-based doubled LDEs. This opens up the possibility that the peripheral system is an 'imperfectly' learned version of the core system acquired by speakers in the main dialect region, one which lacks the key ingredient that is common to the productive systems.[29] In terms of the analysis developed above, this would be the overt C_{MIR}.

This is the approach we take here. We claim that learners in these areas analyse instances of *here it's here* and *there it's there* as true doubling constructions, where the two instances of the locative are related by movement. On our analysis, this involves MULTIPLE COPY SPELLOUT, where there is fronting of the locative to the null C_{MIR} (as in most other varieties of English) but both the top and bottom copy of the fronted locative are spelled out. Thus the tree structure is minimally different from 40b:
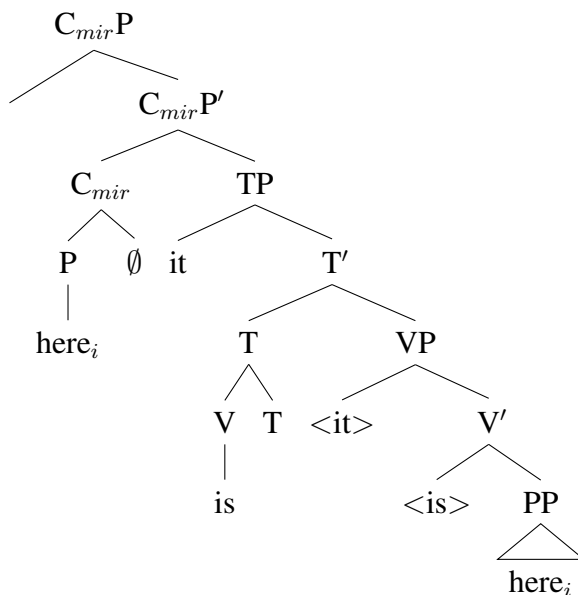
(66)

Multiple copy spellout in movement chains is attested in various languages and constructions, and it seems to be most common in cases where what is moved is minimal unit, typically a head or a proform (see especially Nunes 2004 but also Landau 2006, Kandybowicz 2007, Barbiers et al. 2008, Barbiers et al. 2010, Barbiers 2014[30]). Nunes (2004) proposes that doubling comes about when lower copies of a moved proform are morphologically reanalysed so that they form a morphological unit with some other element in the structure; this is only open to proforms or other minimal units. Nunes claims that this follows if reanalysis leads to a situation where the two occurrences of X are sufficiently distinct so that copy deletion skips the lower one, which would otherwise be

---

[29]We thank an anonymous reviewer for suggestions regarding this alternative analysis.

[30]For doubling phenomena in English specifically, see Close (2004, :3), which follows up on brief remarks on auxiliary doubling in Nunes (2004, :170) and Brenier & Michaelis (2005, , ch.3) on copula doubling, Radford & Felser (2011) on preposition doubling, and Wood (2012) on comparative doubling.

deleted later in the postsyntactic derivation.[31]

(67)   X ... X Y ...   → X ... [X Y] ...

We claim that this is a possible analysis for *here it's here* because the two parts of the *it's* cluster are both clitics which obligatorily incorporate rightwards into the following context. In the case of LDEs, this involves incorporation into the lower copy of the moved locative. When this happens, the lower copy of *here* is not deleted, and we derive doubling.

(68)   here ... it 's here ...   → here ... [it's here] ...

Note that this requires learners to posit a reanalysis rule which rebrackets the *it's* cluster with the following locative. This is the additional ingredient which distinguishes Scots varieties with the peripheral system from other non-Scots varieties which disallow contraction in all of these cases, and we claim that learners will only posit this rule on the basis of positive evidence, for example in instances of *here it's here*.

An important aspect of this analysis is that the reanalysis rule responsible for doubling needs to be distinct from the prosodic incorporation rule which we invoked for regular cases where an auxiliary leans to its right (*it's here*). This is so because if they were the same rule, then we would predict that doubling would occur in gap restriction contexts much more freely, leading to the generation **I don't know where it's where* and other such cases. We can capture this by analysing the rebracketing involved in 68 as a morphological rule which takes place at an earlier stage in the postsyntactic derivation than prosodic incorporation. Such morphological rules are parochial and morpheme-specific, while the prosodic rebracketing rules are general rules which apply in ignorance of which morphemes are involved. Independent evidence for distinguishing two different kinds of rebracketing is hard to come by in this particular case, but there is precedent for distinguishing the domains formed by cliticization (clitic groups) from those formed by morphological rules (phonological words): see Hayes (1989) and Nespor & Vogel (1986).

We conclude this discussion with a prediction of the analysis: if doubling in the peripheral system is triggered by morphological reanalysis of the subject with the following predicate, and this is only possible with minimal proforms, then doubled LDEs should not be possible with phrasal subjects in these varieties. This prediction is borne out, as speakers of these varieties report that examples such as 64, repeated here, are degraded with contraction.

(69)   ??There the bastard's there.   Glenrothes Scots

4.5. Summary. In this section we have proposed analyses of the syntax of various LDE types and shown how their corresponding prosodic structures predict the availability of auxiliary contraction. Table 6 provides a summary of the different syntactic analyses developed, focusing on cases with *it*. These interact with the different pronoun inventories in the way described above.

5. Conclusion. In this article we have described a surprising set of exceptions to the general ban on auxiliary contraction before gaps. The exceptional cases are found only in a subset of Scots varieties, and there only in what we have termed locative discovery expressions. We have shown that the patterns of variation are very orderly, and the systems are those that we'd expect given the possible parameters of variation (C_{MIR}-locative dependency requiring / not requiring movement; pronoun inventories; inventory of null elements). We have provided a uniform explanation of the gap restriction and its exceptions, according to which auxiliary

---

[31]Nunes' own theory requires a different order of operations, since it is couched within an analysis where copy deletion occurs prior to linearization, working with the theory of linearization in Kayne (1994). It is possible to recast his analysis in the terms we have adopted here, but it requires a different theory of linearization.

Type	Syntax	Prosody
Standard LDE	$[_{CP}$ $[_{C'}$ here$_i$+C $[_{TP}$ it $[_{T'}$ is $[_{VP}$ <here$_i$ >]]]]]	(here) (it (is))
Minimal LDE	$[_{CP}$ $[_{C'}$ here $[_{TP}$ it $[_{T'}$'s $[_{VP}$ PRO$_{loc}$]]]]]	(here it's)
Elaborated LDE	$[_{CP}$ $[_{C'}$ here $[_{TP}$ it $[_{T'}$'s $[_{VP}$ $[_{PP}$ in the corner]]]]]]	(here) (it's (in the corner))
Doubled LDE (productive)	$[_{CP}$ $[_{C'}$ here $[_{TP}$ it $[_{T'}$'s $[_{VP}$ $[_{PP}$ here]]]]]]	(here) (it's (here))
Doubled LDE (peripheral)	$[_{CP}$ $[_{C'}$ here$_i$+C $[_{TP}$ it $[_{T'}$'s $[_{VP}$ here$_i$]]]]]	(here) (it's here)

Table 6: Summary of LDE types and analyses

contraction involves prosodic incorporation of clitic auxiliary forms onto adjacent material, with the kind of abstract structure in the 'gap' position determining whether contraction is possible, and shown that the dialectal variation attested falls out of our account. The resulting picture is one where apparent unexpected dialectal variation at the PF interface turns out to follow from minor variations in the syntactic inventories of different dialects, with the points of variation being parametric (different inventories of features). Crucial to revealing this was the microcomparative data, which revealed connections between different constructions and different dimensions of variation that would not have become apparent with a narrower dataset.

References.

AHN, BYRON. 2015. *Giving reflexives a voice: twin reflexives in English*. Los Angeles, CA: UCLA dissertation.

AIKHENVALD, ALEXANDRA. 2004. *Evidentiality*. Oxford: Oxford University Press.

AIKHENVALD, ALEXANDRA. 2012. The essence of mirativity. *Linguistic Typology* 16.435–485.

ANDERSON, STEPHEN R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

ANDERSON, STEPHEN R. 2008. English contracted auxiliaries really are simple clitics. *Lingue e linguaggio* 7.1–18.

BARBIERS, SJEF. 2014. Syntactic doubling and deletion as a source of variation. *Linguistic variation in the minimalist framework*, ed. by M. Carme Picallo, 197–223. Oxford University Press.

BARBIERS, SJEF, and HANS BENNIS. 2007. The Syntactic Atlas of the Dutch Dialects. A discussion of choices in the SAND project. *Nordlyd* 34.53–72.

BARBIERS, SJEF; OLAF KOENEMAN; and MARIK LEKAKOU. 2010. Syntactic doubling and the structure of *wh*-chains. *Journal of Linguistics* 46.1–46.

BARBIERS, SJEF; OLAF KOENEMAN; MARIKA LEKAKOU; and MARGREET VAN DER HAAM (eds.) 2008. *Microvariation in syntactic doubling*. Bingley: Emerald.

BENNETT, RYAN; EMILY ELFNER; and JAMES MCCLOSKEY. 2019. Prosody, focus and ellipsis in Irish. *Language* 95.

BIRNER, BETTY. 1992. *Inversion in English*. Evanston, IL: Northwestern University dissertation.

BOUDAHMANE, KARIM; MATHIEU MANTA; FABIEN ANTOINE; SYLVAIN GALLIANO; and WILL BARRAS. 2008. *Transcriber: a tool for segmenting, labeling and transcribing speech*. Copyright © 1998–2008, DGA, http://trans.sourceforge.net/.

BRENIER, JASON M., and LAURA A. MICHAELIS. 2005. Optimization via syntactic amalgam: syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1.45–88.

BRESNAN, JOAN. 1978. *Contraction and the transformational cycle in English*. Bloomington, IA: Indiana University Linguistics Club.

BRESNAN, JOAN. 1994. Locative inversion and the architecture of universal grammar. *Language* 70.72–131.

CARDINALETTI, ANNA. 1994. On the internal structure of pronominal DPs. *The Linguistic Review* 11.195–220.

CARDINALETTI, ANNA, and MICHAL STARKE. 1996. Deficient pronouns: A view from germanic. a study in the unified description of germanic and romance. *Studies in comparative Germanic syntax*, ed. by Hoskuldur Thráinsson, Samuel Epstein, and S. Peter, vol. 2, 21–65. Dordrecht: Kluwer.

CHAMBERS, J.K., and PETER TRUDGILL. 1998. *Dialectology*. Cambridge: Cambridge University Press.

CHEN, MATTHEW Y. 1979. Metrical structure: evidence from Chinese poetry. *Linguistic Inquiry* 10.371–420.

CHOMSKY, NOAM. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

CHOMSKY, NOAM. 2000. Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.

CLEMENS, LAUREN EBY. 2014. *Prosodic noun incorporation and verb-initial syntax*. Cambridge, MA: Harvard University dissertation.

CLOSE, JOANNE. 2004. *English auxiliaries: a syntactic study of contraction and variation*. York: University of York dissertation.

COLLINS, CHRIS. 2007. Home sweet home. *NYU Working Papers In Linguistics* 1.

COOPMANS, PETER. 1989. Where stylistic and syntactic processes meet: English locative inversion. *Language* 65.728–851.

CRESTI, DIANA, and CHRISTINA M. TORTORA. 2000. Aspects of locative doubling and resultative predication. *Proceedings of the 25th annual meeting of the Berkeley Linguistics Society*, ed. by Steve S. Chang, Lily Liaw, and Josef Ruppenhofer, Berkeley, California, 62–73.

CULICOVER, PETER W., and ROBERT D. LEVINE. 2001. Stylistic inversion in English: a reconsideration. *Natural Language and Linguistic Theory* 19.283–310.

DÉCHAINE, ROSE-MARIE, and MARTINA WILTSCHKO. 2014. Micro-variation in agreement, clause-typing and finiteness:comparative evidence from blackfoot and plains cree. *Proceedings of the 42nd algonquian conference*, ed. by J. Randolph Valentine and Monica Macaulay, 69–101. Albany, NY: SUNY Press.

DELANCEY, SCOTT. 1997. Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1.33–52.

DVOŘÁK, BOŠTJAN. 2007. Slovenian clitic pronouns and what is so special about them. *Slovenski jezik – Slovene Linguistic Studies* 6.209–233.

ERTESCHIK-SHIR, NOMI. 2005. Sound patterns of syntax: object shift. *Theoretical Linguistics* 31.47–93.

HAYES, BRUCE. 1989. The prosodic hierarchy in meter. *Rhythm and meter*, ed. by Paul Kiparsky and Gilbert Youmans, 201–260. Orlando, FL: Academic Press.

HUDDLESTON, RODNEY, and GEOFFREY K. PULLUM. 2002. *The cambridge grammar of the english language*. Cambridge: Cambridge University Press.

INKELAS, SHARON, and DRAGA ZEC. 1993. Auxiliary reduction without empty categories: a prosodic account. *Working Papers of the Cornell Phonetics Lab* 8.205–253.

KAISSE, ELLEN M. 1985. *Connected speech: the interaction of syntax and phonology*. London: Academic Press.

KANDYBOWICZ, JASON. 2007. On fusion and multiple copy spell-out. *The copy theory of movement*, ed. by Norbert Corver and Jairo Nunes, 119–150. Amsterdam: John Benjamins Publishing Company.

KAYNE, RICHARD S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

KAYNE, RICHARD S. 2005a. Here and there. *Movement and silence*, 65–84. Oxford: Oxford University Press.

KAYNE, RICHARD S. 2005b. *Movement and silence*. Oxford: Oxford University Press.

KING, H. V. 1970. On blocking rules for contraction in English. *Linguistic Inquiry* 1.134–136.

KORSAH, SAMPSON, and ANDREW MURPHY. 2016. What can tone tell us about successive-cyclic movement? Evidence from Asante Twi. *Proceedings of NELS 46*, ed. by C. Hammerley and B. Prickett, 227–240. Amherst, MA: GLSA Publications.

LABOV, WILLIAM. 1972. *Sociolinguistic patterns*. Oxford: Blackwell.

LABOV, WILLIAM. 1984. Field methods of the project on linguistic change and variation. *Language in use: readings in sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28–66. Englewood Cliffs, NJ: Prentice Hall.

LANDAU, IDAN. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9.32–66.

LEVIN, BETH, and MALKA RAPPAPORT HOVAV. 1995. *Unaccusativity*. Cambridge, Massachusetts: MIT Press.

MACKENZIE, LAUREL. 2012. *Locating variation above the phonology*. Philadelphia, PA: University of Pennsylvania dissertation.

MEXAS, HARIS. 2016. *Mirativity as realization marking: a cross-linguistic study*. Master's thesis, Universiteit Leiden, Leiden.

NESPOR, MARINA, and IRENE VOGEL. 1986. *Prosodic phonology*. Dordrecht, Holland ; Riverton, N.J.: Foris.

NUNES, JAIRO. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

POSTAL, PAUL M., and GEOFFREY K. PULLUM. 1978. Traces and the description of English complementizer contraction. *Linguistic Inquiry* 9.1–29.

POSTAL, PAUL M., and GEOFFREY K. PULLUM. 1982. The contraction debate. *Linguistic Inquiry* 13.122–138.

PULLUM, GEOFFREY K. 1997. The morpholexical nature of English *to*-contraction. *Language* 73.79–102.

PULLUM, GEOFFREY K., and ARNOLD M. ZWICKY. 1997. Licensing prosodic features by syntactic rules: the key to auxiliary reduction. Paper presented at the LSA annual meeting; accessible at https://web.stanford.edu/ zwicky/PZ1997.pdf.

RADFORD, ANDREW, and CLAUDIA FELSER. 2011. On preposition copying and preposition pruning in wh-clauses in English. *Essex Research Reports in Linguistics* 60.

RETT, JESSICA. 2011. Exclamatives, degrees and speech acts. *Linguistics and Philosophy* 34.411–442.

RIZZI, LUIGI. 1986. On chain formation. *The grammar of pronominal clitics(19)*, ed. by Hagit Borer. New York, New York: Academic Press.

ROBERTS, IAN. 2010. *Agreement and head movement: clitics, incorporation and defective goals*. Cambridge, Massachusetts: MIT Press.

ROOTH, MATS. 1985. *Association with focus*. Amherst, MA: University of Massachusetts, Amherst dissertation.

SCHÜTZE, CARSON T., and JON SPROUSE. 2013. Judgment data. *Research methods in linguistics*, ed. by Robert Podesva and Devyani Sharma, 27–50. Cambridge: Cambridge University Press.

SELKIRK, ELISABETH. 1996. The prosodic structure of function words. *Signal to syntax: bootstrapping from speech to grammar in early acquisition*, ed. by James Morgan and Katherine Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.

SELKIRK, ELISABETH O. 1984. *Phonology and syntax : the relation between sound and structure*. Cambridge, Mass.: MIT Press.

SIGURÐSSON, HALLDÓR ÁRMANN. 2011. Conditions on argument drop. *Linguistic Inquiry* 42.267–304.

SMITH, JENNIFER. 2015-2019. The Scots Syntax Atlas. Research grant, Arts and Humanities Research Council, AH/M005550/1.

SMITH, JENNIFER, and MERCEDES DURHAM. 2012. Bidialectalism or dialect death? Explaining generational change in the Shetland Islands, Scotland. *American Speech* 87.57–88.

SVENONIUS, PETER. 2010. Spatial P in English. *Mapping spatial PPs: the cartography of syntactic structures volume 6*, ed. by Guglielmo Cinque and Luigi Rizzi, 127–160. Oxford: Oxford University Press.

THOMS, GARY, and CRAIG SAILOR. 2018. When silence gets in the way: extraction from *do*-ellipsis in British dialects. *Proceedings of NELS 48*. Amherst, MA: GLSA Publications.

TORTORA, CHRISTINA M. 1997. *The syntax and semantics of the weak locative*. Newark, DE: University of Delaware dissertation.

TORTORA, CHRISTINA M. 2001. Evidence for a null locative in Italian. *Current studies in Italian syntax: studies offered to Lorenzo Renni*, ed. by Guglielmo Cinque and G. Salvi, 313–326. London: Elsevier.

TORTORA, CHRISTINA M. 2014. *A comparative grammar of Borgomanerese*. New York, NY: Oxford University Press.

WILDER, CHRIS. 1997. English finite auxiliaries in syntax and phonology. *Clitics, pronouns and movement*, ed. by James. R Black and Virginia Motapanyane, 321–362. Amsterdam: John Benjamins.

WOOD, JIM. 2012. Double comparatives. Wed page, accessible at https://ygdp.yale.edu/phenomena/double-comparatives. Online: https://ygdp.yale.edu/phenomena/double-comparatives.

WOOD, JIM, and RAFAELLA ZANUTTINI. 2018. Datives, data and dialect syntax in American English. *Glossa* 3.1–22.

ZANUTTINI, RAFAELLA. 2017. Presentatives and the syntactic encoding of contextual information. *Elements of comparative syntax: theory and description*, ed. by Enoch Oladé Aboh, Eric Haeberli, Genoveva Puskás, and Manuela Schönberger, 211–255. Berlin: Mouton de Gruyter.

ZU, VERA. 2018. *Discourse participants and the structural representation of the context*. New York, NY: New York University dissertation.

ZUBIZARRETA, MARIA LUISA. 2001. The constraint on preverbal subjects in Romance interrogatives: A minimality effect. *Subject inversion in Romance and the theory of universal grammar*, ed. by Aafke C. Hulk and Jean-Yves Pollock, 183–204. Oxford: Oxford University Press.

ZWICKY, ARNOLD. 1982. Stranded *to* and phonological phrasing. *Linguistics* 20.3–58.

ZWICKY, ARNOLD M. 1970. Auxiliary reduction in English. *Linguistic Inquiry* 1.323–336.