Variation in Agreement: a lexical feature-based approach*

David Adger and Jennifer Smith

November 2007

1   Introduction

This paper has two main theoretical aims: (i) to argue for the utility of a lexical, feature-based, approach to intra-dialectal and inter-dialectal language variation, confirming and extending Borer's (1984, 251–254) suggestion that language variation is ultimately a matter of the properties of the lexicon of functional categories; (ii) to argue that the variability found in an individual speaker is two-dimensional: it may involve varying featural specification of functional categories and/or underspecification in the mapping between these categories and morphological forms, the former modeling the kind of

variation usually thought of as 'parametric' and the latter modeling the kind of variation usually captured by the notion of linguistic variable (Labov 1994, Labov 2000). We thus offer a unified model of the grammatical representations that underly language variation of both types.

We defend this theoretical position through an analysis of a pattern of variability in subject-verb agreement in a variety of English spoken in the town of Buckie in North-East Scotland (see Smith 2000). The variability arises with plural Noun Phrase subjects which, from the perspective of Standard English, display apparently singular agreement. We find the variability in the past tense of the verb *be*, and in the present tense of *be*, main verbs and other auxiliaries. We call variability in the past tense of *be was/were* variation and will call the appearance of the morpheme *-s* on a present tense verb 'verbal *-s*' (see, e.g. Poplack and Tagliamonte 1991). (1) gives examples of *was/were* variation, while (2) shows variable verbal *-s* with present tense *be*:[1]

(1)   a.   Buckie boats **were** a' bonny graint. (older female)

              'Buckie boats were all nicely grained.'

    b.   The mothers **was** roaring at ye comin' in. (older male)

              'The mothers were shouting at you to come in.'

(2)   a.   Nearly all the houses **is** bought with the English (older female)

              Nearly all the houses are bought by the English

    b.   All my old pals **are** all away but me (older female)

              All my old friends are away except me.

---

[1] The majority of examples are taken from a corpus of about 40 hours of spoken data (Smith 2000). In the remaining cases, grammaticality judgments are used. These are indicated in the text and were checked by the second author and by participants in the original study.

Verbal -*s* also appears variably with main verbs and auxiliaries such as *do* and *have*:

(3)   a.   What bairns **walk** any distance? (middle aged female)
            'Which children walk any distance?'

      b.   When they go back, the teachers **asks** them to write something
            and they send them till's. (older male)
            When they go back, the teachers ask them to write something
            and they send them back to us.'

(4)   a.   Maybe some churches **does** na do it yet, (older female)
            'Maybe some churches don't do it yet.'

      b.   Most adults' clais you get have a spare button, but bairns' ones
            do na (younger female)
            'Most adults' clothes you get have a spare button but children's
            ones don't.'

All thirty eight speakers represented in the corpus exhibit variability in form. Overall percentages for *was/were* variation with pronominal subjects can be seen in (5) and for verbal -*s* in (6).

(5)

pronoun	percentage of *was*	N
1st singular	100	691
2nd singular	69	161
3rd singular	100	2290
1st plural	67	368
2nd plural	10	10
3rd plural	0	435
Singular NPs	100	762
Plural NPs	56	187

Variable verbal *-s* in Buckie is only found with plural non-pronominal NP subjects (see section 2 for a more detailed characterization), and the overall percentages are given in (6):

(6)

verbal *-s* with third person plural non-pronominal NPs	
%	N
58	588

We will argue that the agreement patterns here are an integral part of the grammar of this variety of English, developing our previous analysis on the past tense agreement system of the verb *be* in Buckie to account for this pattern (Adger and Smith 2005, Adger 2006). We will argue that, while *was/were* variation involves just underspecification of features, verbal *-s* variation in addition requires a difference in the inventory of grammatical features at play in the grammar. Thus we will see the interaction of the two modes for capturing language variation: underspecification, leading to 'Labovian' variation, and variation in feature composition, leading to 'parametric' variation.

4

1.1 Combinatorial Variability

Adger (2006) develops an approach to the variable use of *was/were* forms of preterite *be*. The fundamental idea is that variation in surface form can be traced back to the lexical inventory of a language and more specifically to an underspecification of the grammatical features of the preterite forms of *be*. The Buckie paradigm is especially interesting because it seems to form a system which cannot be derived from the standard system by simply saying that one particular standard form can also have a local variant. The paradigm looks as follows (see (5) for the percentage of *was* in each paradigm cell):

(7)

I was	We **was/were**
You **was/were**	You (ones) **was/were**
He/she/it was	They were

We would like to immediately forestall a possible analysis that takes the Buckie paradigm to be identical to that of standard English but where *were* always has a *was* variant. This will not work for the Buckie paradigm. *Was* is simply impossible with a third person plural pronominal subject. There are 435 possible contexts for *they was*, but not a single occurrence.

The theory of this system developed by Adger (2006) adopted a Minimalist framework of assumptions based on Chomsky (1995). In this framework, the input to the syntactic system is a set of lexical items, where each lex-

ical item is itself a set of features. The language learner has to posit the correct features and also correctly bundle them up into lexical items. The grammatical structures of the language are then simply derivative of universal invariant syntactic principles coupled with the particular lexical items. Crucially, within this framework, lexical items include functional categories (in fact, this is where all the featural action is). Adger (2006) proposes an algorithm which generates a set of lexical items for agreeing verbs directly on the basis of the morphological input (see also Marr 1982 for a similar idea in the field of vision, and Pertsova 2007 for recent discussion of such learning algorithms for grammatical phenomena). The algorithm proposed in Adger (2006) looks as follows:

(8) Seek Maximal Generalization by

    a. Generating all n-feature LIs, where n=1.

    b. Mapping them to forms, so that a successful mapping is made if there is a form which the LI always matches—essentially reduces multiple exponence (Reject Optionality)

    c. Delete any spurious LI's (i.e. delete each n-feature LI and check if coverage is reduced; if it is, reinstate, if not continue) - essentially reduces synonymy (Reject Synonymy)

    d. Recursing over n=n+1, with the proviso that if a form has been successfully analysed in the n–1th step, LIs capturing it in the nth step will be rejected (Minimize Lexicon).

It is important to realize that the algorithm given here is simply a model of the relationship between the primary linguistic data and the lexical specifications of the functional elements. It may well be used as part of the acquisition process, but it is primarily intended as a model of why the struc-

ture of the lexicon is what it is (analogously to the notion that derivations model syntactic competence even though they are not used in online processing of sentences). It is the nature of the representations that is at issue here, rather than the process of their acquisition or use.

The features we will take as relevant for the specification of English agreement here are those of Halle (1997) (see also Harley and Ritter 2002 and Nevins and Parrott (this volume)): [singular:±], [participant:±] and [author:±]. These three features allow us to distinguish the various paradigm slots of the most richly agreeing verb in English, (*be*), and we will take it that they are the features that are always relevant to agreement phenomena in the language.

On a first pass, the algorithm matches single feature lexical items with forms as follows:

(9) (a) [singular:+] was
     (b) [singular:–] were
     (c) [participant:+] was
     (d) *[participant:–] was, were
     (e) [author:–] were
     (f) [author:+] was

Since it is not the case that there is one form which [participant:–] can always be mapped to, this lexical item is rejected, leaving the following fragment of the lexicon:

(10) (a) [singular:+] was
      (b) [singular:–] were
      (c) [participant:+] was
      (d) [author:–] were

(e) [author:+] was

The items in (10) should be conceived of as mappings between phonological (or morphological) forms and morphosyntactic features, that is, they serve an interface function. In a system of grammar like that outlined in Chomsky 1995, they would be the input to the syntactic derivation. However, under a slightly different architecture which allows 'Late Insertion' of phonological forms (e.g. the Distributed Morphology model of Halle and Marantz 1993), they can be taken as instructions to spell out certain features as certain forms (this was the architecture assumed in Adger and Smith 2005; see also Parrott 2006 and Nevins and Parrott 2008). We think that the basic ideas we defend here are implementable in either framework of assumptions and we will remain largely neutral in the remainder of the discussion.

Now, imagine that the syntactic derivation contains a second person singular pronoun in the subject position. There are three items from (10) that will match with such a pronoun, giving a **Pool of Variants** for the agreeing verb (see Adger 2007). In any particular utterance situation, various factors will impact on the selection of one of members of the pool of variants, so that the phonology of the variants in the surrounding phonological context, the sociolinguistic status of the variant, its recency etc, might all in principle have an effect on which variant is chosen. We can conceive of this as a choice function $\mathbf{U}$, selecting an element from the pool of variants. However, it could be that the phonological forms themselves are not evenly distributed across the pool: as in pools which emerge from the derived lexicon in (10), where we have grammatical homonyms. In such cases, the final *surface* form chosen will of course also depend on this. In our example with a second singular subject, the pool of variants looks as follows:

(11)   [singular:+, participant:+, author:–]   (a) was; (c) was; (d) were

Now the various factors that go into **U** will have an impact in any particular situation. However, the impact of the uneven distribution of phonological forms in the pool of variants itself is a constant factor, and can be seen as responsible for uneven distribution of forms in the corpus. If U applies to the pool of variants x times, then the surface form of the variant will be *was* $0.66x$ times and *were* $0.33x$ times, on the assumption that nothing else is affecting the behaviour of U (see §4.3 for some discussion of potentially important factors that may impact on the behaviour of U). It turns out that this system actually makes a rather good prediction about the frequency of forms in the corpus, as can be seen from (12):

(12)

pronoun	percentage of *was*	N
2nd singular	69	161
1st plural	67	368
2nd plural	10	10

The corresponding pool of variants for *we* and *you*[plural] are as follows:

(13)   [singular:–, participant:+, author:+]   (b) were; (c) was; (e) was

(14)   [singular:–, participant:+, author:–]   (b) were; (c) was; (d) were

It's clear that the prediction is straightforwardly confirmed for *you*[singular] and *we*, both of which are predicted to occur with *was* approximately two thirds of the time. The case of *you*[plural] is less clear. The system predicts 7 *were* and 3 *was* and we find 9 *were* and 1 *was*. We take this to be an artifact of the paucity of the data. Informant judgment tasks show that both are possible, if we force a plural interpretation of *you* by using a quantificational modifier like *all*:

9

(15)   You was all there.   Buckie judgment

(16)   You were all there.   Buckie judgment

See section 4.1 for further discussion of this kind of data.

The preceding discussion takes the systems of use to apply to the pool of variants, and assumes that the properties of elements of the pool, and properties of the utterance (both what has been produced and what is planned) are relevant to the selection. We think that this is consistent with both the 'lexicalist' and 'Late Insertion' interpretations of (10), although both might need different additional assumptions about the connection between syntax and processing.

1.2   Standard English Present Tense Agreement

Let's see how this system works for present tense agreement, which is where verbal -*s* appears, focussing first on a non-variable system like that of standard English. The present tense agreement system of Standard English is fairly impoverished, with inflection appearing overtly only in the third person singular:

(17)   singular   plural

    1   0   0

    2   0   0

    3   -s   0

Following the algorithm, we first generate all one feature lexical items:

(18)   a.   *[singular:+] 0/-s

    b.   [singular:–] 0

    c.   [participant:+] 0

d. *[participant:–] 0/-s

e. [author:+] 0

f. *[author:–] -s/0

We discard (a), (d) and (f), since they violate step (b) of the algorithm. However, at this point we still have no successful analysis of 3rd person singular, so we move to the second pass of the algorithm, generating all two feature lexical items:

(19) a. [singular:+, participant:+] → 0

b. [singular:+, participant:–] → -s

c. [singular:–, participant:+] → 0

d. [singular:–, participant:–] → 0

e. *[singular:+, author:–] → 0/-s

f. [singular:+, author:+] → 0

g. [singular:–, author:+] → 0

h. [singular:–, author:–] → 0

i. [author:+, participant:–] → inconsistent, as authors have to be participants

j. [author:+, participant:+] → 0

k. [author:–, participant:+] → 0

l. *[author:–, participant:–] → 0/-s

The only results that are relevant are those that have -s as an output. The others are rejected as they analyse parts of the paradigm which already have an analysis by virtue of the first step of the algorithm and are thereby ruled out by clause (d) of (8), Minimize Lexicon. Of the remaining three cases, two are rejected because there is no single form to which the relevant feature bundle can always be mapped (this is clause (b) of (8), Reject Optionality,

11

at play), leaving the only successful lexical item as (19-b). We end up with the final lexicon:

(20)   a.   [singular:–] 0

          b.   [participant:+] 0

          c.   [author:+] 0

          d.   [singular:+, participant:+] -s

A similar process, applied to the present tense forms of the verb *be*, results in the following lexicon:

(21)   a.   [singular:–] are

          b.   [author:–] are

          c.   [singular:+, author:+] am

          d.   [singular:+, participant:–] is

These mini-lexica will not generate any variability, as the reader can readily ascertain, unlike the case of the mini-lexicon for Buckie preterite *be*.

2   Variable Verbal -s in Buckie

The grammar of Buckie disallows the apparently third person singular form of the present tense verb with third plural pronouns, as noted above, analogously to the *was/were* pattern. Judgment tests, where we take examples from the corpus and substitute in pronominals in place of DP subjects with lexical context, are uniformly rejected by speakers:

(22)   *When they **comes** up to see us there   Buckie judgment

(23)   *When they **gets** home   Buckie judgment

(24)   *They **does** na do it yet   Buckie judgment

(25)   *They **does** na get what that cats get   Buckie judgment

(26)   *They **is** up in Elgin   Buckie judgment

This judgement task is borne out in the corpus, where only 0.6% of the non-agreeing -s forms occur with *they*:

(27)

plural 3rd pronoun subject		plural NP subject	
%	N	%	N
0.6	1271	58	588

The 0.6% in the left hand column consists of eight exceptions. These eight exceptions are of a particular type: they are 'verbs of communication' which Leech (1987, 11) points out have their own specialized system connected to the semantics of these verbs and which are well documented in the variationist literature (e.g. Johnstone 1987).

(28)   And they says till him 'Well, Mr. Smith, come in on Monday. (older female)
       'And they said to him "Well Mr Smith, come in on Monday".'

(29)   But they says 'That's it there, like.' (middle aged male)
       'But they said "That's it there, like".'

(30)   They says actually they're never all in thegither bar maybe the week-end. (middle aged female)
       'But they said actually that they are never in all together except maybe at the weekend.'

(31)   He started being sick and they says 'Positing.' (younger female)

13

'He started being sick and they said "Positing".'

(32)   And they just says it was magic, ken. (younger male)

'And they just said that it was magic, you know.'

(33)   Aye, aye, but they never tells you that in the paper, like. (older male)

'Yes yes, but they don't ever tell you that in the newspaper.'

(34)   Oh aye, a lot of things that they never speaks about, you ken. (middle aged male)

'Oh yes, there are a lot of things that they don't ever speak about, you know.'

(35)   But they tells you. (older male)

'But they tell you.'

We assume then that these examples involve some other system, where *-s* marking is used to mark a particular narrative/aspectual meaning associated with verbs of communication.

Putting these eight examples aside, we are left with the patterns in (36), which show variability in the present tense when the subject is a plural noun phrase, and variation more widely in the paradigm of the preterite tense of *be*, also including variability when the subject is a plural noun phrase. The mini-lexicon in (10) will capture the categorical variable split in preterite *be* with pronominal subjects, so the remaining task is to capture the rest of the system in (36).

(36) a.

present main verb	singular	plural
1	-0	-0
2	-0	-0
3 pronoun	-s	-0
3 NP	-s	-s/-0

b.

present *be*	singular	plural
1	am	are
2	are	are
3 pronoun	is	are
3 NP	is	is/are

c.

preterite *be*	singular	plural
1	was	was/were
2	was/were	was/were
3 pronoun	was	were
3 NP	was	was/were

3   Previous Accounts of Agreement Variability

3.1   Syntactic Explanations

Henry (1995) examines a similar construction in Belfast English, which she terms SINGULAR CONCORD. She proposes an analysis based on the syntactic position of the subject (see also Tortora and den Dikken 2008). Obviously, it is of interest to see whether this hypothesis will capture the Buckie patterns.

The verbal *-s* phenomenon in Belfast English is similar to the Buckie pattern in that we find variability in the present tense between *-s* marking and

absence of -*s* marking. Henry (1995, 19) observes that in these constructions, the presence of an adverbial intervening between the subject and the finite verb impacts on the possibility of -*s* marking. Henry first notes that, in general, verbal -*s* marking in this dialect is possible with auxiliary verbs as well as with main verbs, just as we have seen for Buckie:

(37)   These cars **go/goes** very fast.

(38)   The eggs **are/is** cracked.

She then shows that when an adverbial intervenes between a subject and an auxiliary, verbal -*s* becomes impossible:

(39)   The children really **are** late

(40)   *The children really **is** late.

(41)   The children really **likes** pizza

In the Buckie corpus, we have few examples where an adverb intervenes between the subject and the finite verb. Two cases clearly involve a main verb, and are analogous to (41):

(42)   No, my folk originally **comes** fae Portsoy. (older female)
        'No, my family originally come from Portsoy.'

(43)   I says "Your other two sisters never **comes** over".(older female)
        'I said "Your other two sisters don't ever come over."

One case plausibly involves an auxiliary that has raised to T, but unfortunately it also has an analysis as a main verb that stays in the VP, and this analysis cannot be ruled out:

(44)   Helen and me often **has** a laugh thegither, you ken, (older female)

          'Helen and I often have a laugh together, you know.'

One might ask, then, whether Buckie displays the same pattern as is found in Belfast English. Our suggestion is that it does not.

Although we cannot tell from the corpus data available to us, we can use informant judgment tests. These show that verbal -*s* with an auxilliary is in fact well formed:[2]

(45)   The bairns probably/maybe **doesna** go there.   Buckie judgment

          'The children probably/maybe don't go there.'

(46)   Her sisters maybe **hasna** been there afore.   Buckie judgment

          'Her sisters maybe haven't been there before.'

(47)   The bairns really **isna** well behaved.   Buckie judgment

          'The children really aren't well behaved.'

A second property of the Belfast construction that Henry points out has to do with inversion structures. In Belfast English verbal -*s* is apparently impossible in inversion constructions Henry (1995, 19):

(48)   ***Is** the students here?

(49)   ***Has** the children arrived yet?

This finding is not replicated in Buckie, where we find examples like the

---

[2]We use negated auxiliaries here to control for any preference to contract *be/have* auxiliaries to the subject. These sentences are also good with non-negated auxiliaries, but then there is an alternative where the auxiliary contracts with the subject and the adverb follows. Since we are interested in syntactic factors here, we have attempted to control for the impact of any preference for contracted forms in general.

following:

(50)   **Is** your lights gan on and off? (middle aged female)

'Are your lights going on and off?'

(51)   **Has** your granny and them aie been in Rathven, aye? (middle aged female)

'Have your grandmother and her family always been in Rathven?'

However, although Buckie and Belfast behave differently with respect to adverbial intervention effects and subject auxiliary inversion, they do show the same patterning of verbal -s with pronominal vs NP subjects: in both varieties singular concord is impossible with a pronoun (a constraint known as the 'Northern Subject Rule' Murray 1873):

(52)   *They **goes** very fast.   Belfast judgment

(53)   *They **is** cracked   Befast judgment

We find the same categorical impossibility of singular agreement with a pronominal in Buckie:

(54)   Her two daughters **is** up in Elgin (older female)

'Her two daughters are up in Elgin.'

(55)   *They **is** up in Elgin

In her analysis of singular concord in Belfast English, Henry firstly argues that what we have here is a syntactic, rather than a semantic phenomenon, and that it cannot be reduced to cases of collective singular agreement found in British English examples like the following:

(56)   The government **have/has** resigned.

Henry establishes this point by showing that the inversion and adverb intervention effects do not apply to this kind of agreement. It is certainly also the case that in Buckie, the semantic effect of collectivity is not found with singular concord cases. We can see this fairly clearly by comparing Standard British English judgments on collective nouns with singular agreement. If we force the plural reading of the collective noun, then the floated quantifier *all* forces plural agreement:

(57) a. *The government **has** all resigned.   Standard British English judgment

  b. The government **have** all resigned.   Standard British English judgment

However, analogous examples from the Buckie corpus show that there is still variability:

(58) a. The state rooms **are** all used (younger male)

  b. The rebates at the end of the year**'s** all based on fuel (younger male) 'The rebates at the end of the year are all based on fuel.'

These attested examples are backed up by grammaticality judgments on the following cases:

(59) a. The quines **has** all gone.   Buckie judgment
        'The girls have all gone'

  b. The quines **have** all gone.   Buckie judgment
        'The girls have all gone'

Henry argues that verbal -*s* is a lack of agreement marking, and is correlated with nominative case marking. She proposes a clause structure with both Agr and T heads, and assumes that both the specifier of AgrP and TP are

potential subject positions but that only Agr checks nominative case in its specifier:

(60)
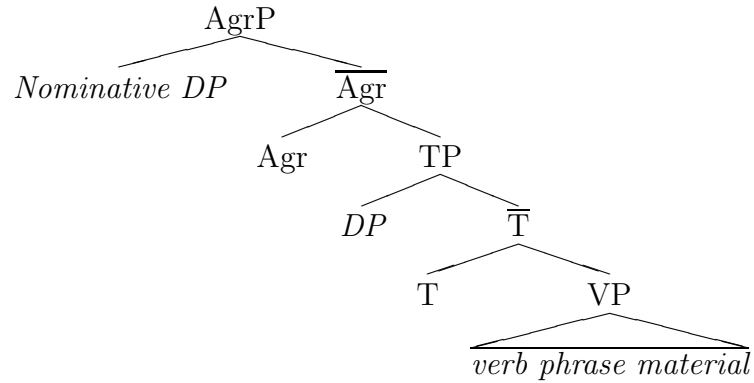
The next step in the proposal is that pronouns which are morphologically marked as nominative must check their nominative case. This will force a nominative marked subject pronoun to raise to the specifier of AgrP, as the pronoun's nominative case cannot be checked in the specifier of TP:

(61) a. $[_{AgrP}$ pronoun[nom] [Agr[nom] $[_{TP}$ DP T ...]]
      b. *$[_{TP}$ pronoun[nom]T ...]]

However, DPs in general do not bear overt nominative case. Henry suggests that they can be marked with a default case, which is checked in the specifier of TP. However, such non-pronominal DPs can also check nominative case, so they too can appear in the specifier of AgrP, leading to an optionality of positioning for DPs but not for pronouns:

(62) a. $[_{AgrP}$ DP[nom] [Agr[nom] $[_{TP}$ ⟨ DP⟩ T ...]]
      b. $[_{TP}$ DP[default] T[default] ...]]

The system now has a difference in the position of pronouns and DPs which can be used to explain the categorical/variable agreement split. As-

sume that whenever the specifier of Agr is filled, the finite verb is marked for agreement, but when the specifier of T is filled, the verb marks with-*s*. It follows that pronouns always cooccur with agreement, while DPs may, or may not.

As evidence that there are two positions involved in this construction, Henry draws on the adverbial adjacency effects discussed above. Recall that verbal -*s* in Belfast English is impossible between a subject and an auxiliary if an adverb intervenes:

(63)   The children really **are** late.

(64)   *The children really **is** late

Henry proposes to capture this by assuming that there is an adverbial position between the agreeing subject and the finite verb, but this position is absent between the non-agreeing subject and the finite verb. To implement this, Henry proposes that the adverb attaches to TP, and the finite auxiliary is always in T (with agreement lowering at PF):
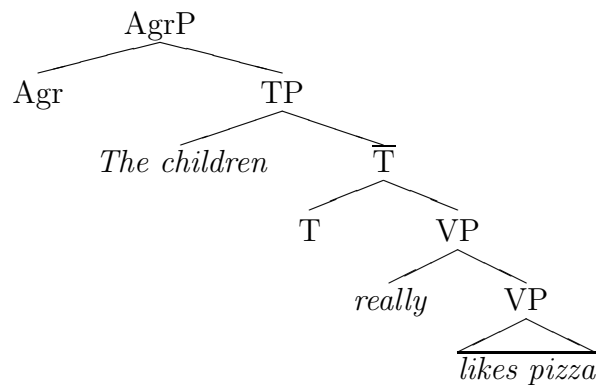
(65)
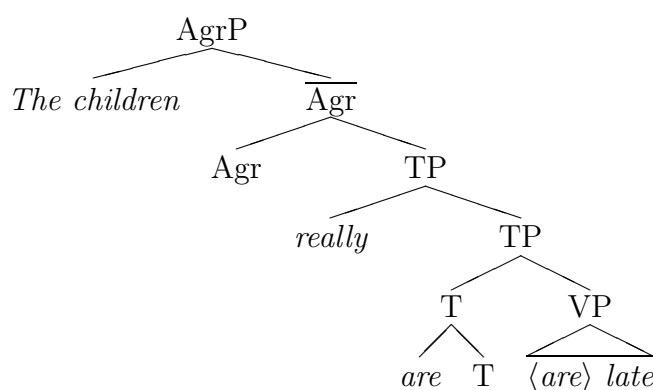
If the subject is in the specifier of TP, then there is no adverbial position available, since whereas TP is a possible adjunction site, $\overline{\text{T}}$ is not:

(66)

```
                    AgrP
                 ╱        ╲
              Agr          TP
                        ╱      ╲
                 The children    T̄
                              ╱     ╲
                            T         VP
                          ╱  ╲      ╱    ╲
                        is    T  ⟨is⟩  late
```
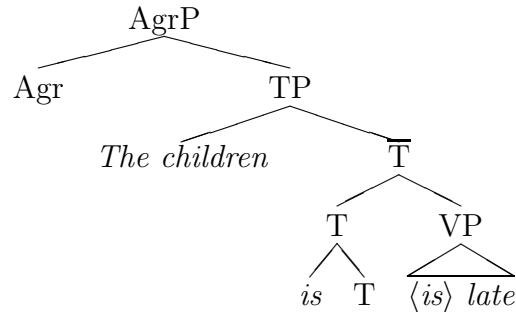
To capture the fact that main verbs allow verbal -s even when the adverb intervenes, Henry capitalizes on their in situ position inside VP. In this case the adverb would attach lower than TP:

(67)

```
                    AgrP
                 ╱        ╲
              Agr          TP
                        ╱      ╲
                 The children    T̄
                              ╱     ╲
                            T         VP
                                    ╱    ╲
                                really    VP
                                        ╱    ╲
                                    likes pizza
```

Furthermore, Henry proposes that this structure also allows us to capture the absence of verbal -s in inversion constructions in Belfast English. The idea is that T to C movement must pass through Agr, because of the Head Movement Constraint. This "activates" Agr, and forces the subject to raise to the specifier of AgrP, and agreement thereby appears on the finite auxiliary.

Henry's system captures the Belfast English patterns fairly effectively, although it does so at the expense of a number of stipulations. The system requires us to allow both default and nominative case in Belfast English,

and to allow both to be optional. For example, when the subject is a full DP, the nominative case on Agr must not be available, so either it must be optional, or projection of Agr must be optional. Similarly, when the subject is a pronoun, the default case on T must be optional. A second assumption is that DPs, but not pronouns, have to be able to check both nominative and default case. Furthermore, in order to capture the adverbial effects, the adverbs must be stipulated to adjoin to VP or TP, but not to AgrP.

Although the analysis is capable of capturing the patterns reported, it is difficult to extend it to closely related dialects. Recall that a pronoun marked with nominative case must be in the specifier of AgrP, and is hence impossible with singular concord.

'*we, they,* and *youse* . . . all require agreement'

(Henry 1995, p 18).

Henry doesn't report data where the subject is first person plural and the verb is the past tense of *be*. However, her system would predict only *we were* and not *we was*. As we have already seen, *we was* is well formed in Buckie (along with *you was*) and in addition, we have cases of spoken narrative from conversations between members of the Portavogie speech community in Northern Ireland (County Down, about 26 miles from Belfast) with *we was* (Tagliamonte 2000–2003):

(68)   And we **was** comin' by with the Queen Mother (older male)

(69)   When we **was** wee fellers (older male)

It may be that there is some dialectal difference between Henry's informants and the Buckie and Portavogie speakers, but clearly the analysis developed above which ties nominative case on a pronoun to full agreement will not

immediately extend. Tortora and Den Dikken (this volume) also report a nominative marked complex pronoun in Appalachian English, *we'uns* which allows verbal *-s*, even though the Appalachian system is otherwise like Belfast and Buckie in disallowing verbal *-s* with simplex pronouns.

A second issue arises with the interpretation of the adverbials. If the adverb *really* can adoin to TP, and a full DP subject can be in the specifier of TP, then we predict the following sentences should have the same reading for the adverb (which is in the same position in both cases, adjoined to TP):

(70)   The children $[_{TP}$ really $[_{TP}$ are late $]$ $]$

(71) $[_{TP}$ Really $[_{TP}$ the children is late $]$ $]$

However, two Belfast informants have confirmed that, for them at least, the modificational reading is lost in (71) and only a speaker oriented reading is possible.

Note that in Buckie, the system is different. There is no constraint excluding verbal *-s* in inversion constructions, and, moreover, there does not seem to be an adverbial intervention effect. However, the pronoun constraint is still in effect.

There are a number of ways to capture this, while maintaining the spirit of Henry's system. One possibility would be to say that the relevant adverbials have a slightly looser distribution in Buckie than in Belfast, and may adjoin to $\overline{\text{T}}$. That is, we'd have a structure like the following:

(72)
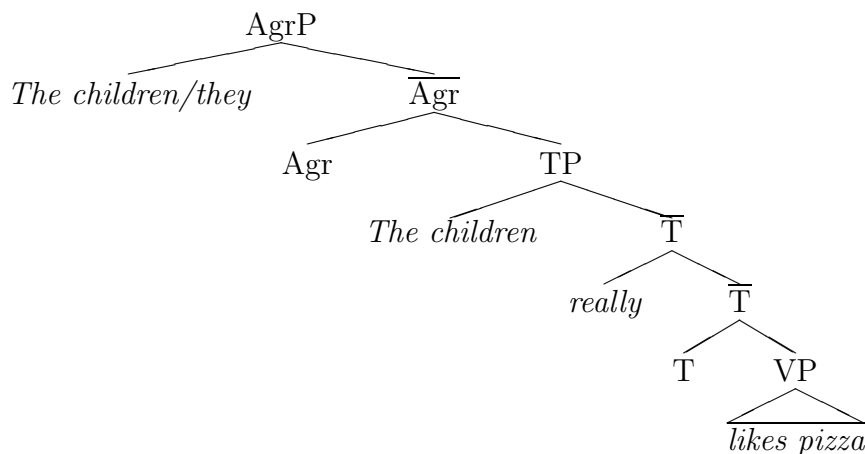
This structure would allow adverbs to intervene with both DP and pronominal subjects. However, it would not explain why inversion is also possible with a singular verb. In Henry's system, movement of T to C "activates" Agr and an activated Agr ends up being spelled out as agreement. We'd have to say that movement through Agr in Buckie doesn't activate it. Overall, Henry's approach doesn't seem to extend naturally to the Buckie system.[3]

A rather different syntactic analysis of this kind of data is given by Börjars and Chapman 1998. Börjars and Chapman propose that pronouns are essentially ambiguous between arguments and inflectional markers, and that an inflectional marker pronoun will be, by its very nature, in complementary distribution with inflection. If we take the pronoun *they* as an example, and assume that it may only be an inflection marker (we notate it as *they*$_{\text{Infl}}$), then we predict:

(73)   *pro* they$_{\text{Infl}}$ go

---

[3]A reviewer suggests that Henry's analysis, which relies on syntactic position, is possibly correct for Belfast English, while a lexical analysis along the lines of that developed below is correct for a system like that of Buckie, see also Tortora and Den Dikken (this volume). This may well be correct, although we think that the theoretical problems pointed out here for Henry's system, as well as the issue of adverbial readings discussed immediately above suggest a more thoroughgoing reappraisal is necessary (see also Henry 2005).

(74)   *pro they$_{\text{Infl}}$ go-es

(73) is essentially an inflected form of the verb, with an empty subject. (74) is ungrammatical as *they* and verbal inflection are in complementary distribution. This approach also predicts:

(75)   *They$_{\text{Infl}}$ often go

(76)   *We$_{\text{Infl}}$ and they$_{\text{Infl}}$ go.

This will follow, since *they*$_{\text{Infl}}$ is an inflection, and cannot be separated from its verb.

Börjars and Chapman suggest pronouns in dialectal English are in general ambiguous, so as well as *they*$_{\text{Infl}}$, there is also a simple pronoun *they*. While *they*$_{\text{Infl}}$ cannot be separated from the verb, and cannot be coordinated (as it is essentially a clitic), *they* can be:

(77)   We and they **goes**.

(78)   They often **goes**.

This then predicts a pattern of data where verbal -*s* appears when the plural pronoun is not directly adjacent to its verb. Of course, the system doesn't immediately capture the ungrammaticality of the string *they goes*, since this has a well formed parse where *they* is a pronoun rather than a clitic, so the system needs to be augmented with an economy principle that avoids pronouns in favour of inflection (cf. Chomsky 1981, 65). The principle would roughly say the following:

(79)   When there is a choice between *they*$_{\text{Infl}}$ and *they* in a syntactic context, choose the former.

This will now rule out *they goes*, since (79) tells us to choose the inflectional variant of *they* and that variant is incompatible with *-s* appearing on the verb.

As we will see in the next subsection, the ban on *they* with verbal *-s* in Buckie is not sensitive to adjacency: (78) is ungrammatical in Buckie. To capture this within the system under consideration, one would have to say that only the inflectional variant of the third person plural pronoun is available in Buckie. This would then correctly rule out the third person plural pronoun with verbal *-s*. However, the problem would then be cases where the third person plural pronoun occurs in a sentence where it is not adjacent to its verb:

(80)   They often **go** (grammatical in Buckie)

Such sentences are perfect in Buckie, but if the pronoun is only inflectional, they are predicted to be ungrammatical by Borjars and Chapman's system. For the Buckie agreement paradigm, something else is required.

3.2   A Usage Based Analysis

Pietsch (2005) suggests a usage based account. This takes a rather different view of what a grammar is. Rather than assuming that a grammar is a set of principles or rules that generate structures on the basis of information stored in lexical items, usage-based approaches assume that grammars are collections of constructions, where a construction is a kind of template or schema (see, e.g. Goldberg 2003). The language acquirer abstracts these constructions from linguistic events, and since different linguistic structures are encountered in discourse at different frequencies, these lead to qualitatively different linguistic representations. More frequent constructions be-

come more deeply entrenched, and such deeply entrenched constructions are easier to use. From this perspective, there is no separation between grammar and lexicon; rather we have a network of constructions, from very schematic to very specific.

Pietsch uses this basic model to suggest an approach to the kinds of pattern we are interested in here. He proposes that *they+verb* is a specific construction, instantiating the more general schema subject+verb. Given that it is more specific, it may override some aspects of the latter (e.g agreement). High frequency specific schemas may become entrenched, and coexist in the adult grammar with the more general schemas. These then compete in any production of a sentence.

Under Pietsch's view, a Standard English system will consist of a network of constructions, of the following general sort:

(81)   a.   Subj-Verb
       b.   Subj[+3sg] V-s
       c.   Subj[–3sg] V-0
       d.   he V-s
       e.   Subj[+3sg] is
       f.   I V-0
       g.   Subj [–3sg] are
       h.   . . .

There is massive redundancy in such a system, but the idea is that this need not concern the analyst. Memory is cheap, it's processing that's costly; so it is an efficient use of resources to maintain a redundant system like this.

A system like Buckie can then be straightforwardly represented by a set of slightly different schemata, crucially including (e) in (82), which will allow

any NP to cooccur with -s inflection (irrespective of that NP's plurality). Since such a system will also have a low level schema for *they+verb+0*, plus other low level schemata for NPs specified for plurality, we have something like the following:

(82) a. they V-0

b. she V-s

c. NP[sg] V-s

d. NP[pl] V-0

e. NP V-s . .̣ .

We can then add on extra specifications to this, enriching the representations so that, for example, particular verbs might have more entrenched representations than others (e.g [NP goes] is more entrenched than [NP-runs]).

This model can explain more or less any set of constraints on local co-occurrence patterns in strings but crucially there is no *grammatical* explanation for why a particular paradigm is the way it is. This means that the model makes predictions about what one might expect when the subject is separated from its verb in various ways. Since it is 'they V-0' that is entrenched, we expect verbs which are separated from 'they' to allow verbal *-s*. At least in Buckie, this doesn't seem to be possible:[4]

(83) The quines aie **sings** too loud   Buckie judgment
'The girls always sing too loudly.'

(84) *They aie **sings** too loud   Buckie judgment
'They always sing too loudly.'

An even stronger argument can be made on the basis of coordinate construc-

---

[4]Although this is possible in some varieties. See Pietsch 2005 for a description of these.

tions:

(85)   The loonies **kens** fit's good for them and **ken/kens** fit's bad for
        them an a'.
        'The boys know what is good for them and know/knows what is bad
        for them as well.'

(86)   The loonies **ken** fit's good for them and **ken/kens** fit's bad for them
        an a'.
        'The boys know what is good for them and know/knows what is bad
        for them as well.'

These judgments show that variability in verbal -*s* is preserved even when
the verb is not adjacent to its subject.[5] With this in hand, we can now look
at verbs in coordinations with pronominal subjects:

(87)   *They **kens** fit's good for them and **ken/kens** fit's bad for them an
        a'.
        'They know what is good for them and know what is bad for them
        as well.'

(88)   They **ken** fit's good for them and **ken/*kens** fit's bad for them an
        a'. 'They know what is good for them and know what is bad for
        them as well.'

Here we see a pattern predicted by a grammatical agreement requirement
on pronominal subjects, but a mystery on Pietsch's account. There is a

---

[5]There is a very weak parallelism effect which creates a preference for the verb in the
second member of a coordinate construction to have the same expression of agreement as
the verb in the first member. However, there is no doubt in speakers minds of the contrast
between this and the ungrammatical cases reported in (88).

constraint against using the verbal -*s* marked form even though it is separated from its subject by an entire verb phrase. The same facts can be seen with preterite *be*:

(89)   The smokies **were** magic and **were/was** na dear. Buckie judgment
        'The smoked fish were really good and weren't expensive.'

(90)   The smokies **was** magic and **were/was** na dear.   Buckie judgment
        'The smoked fish were really good and weren't expensive.'

(91)   They **were** magic and **were** na dear.   Buckie judgment
        'They were really good and weren't expensive.'

(92)   *They **were** magic and **was** na dear.   Buckie judgment
        'They were really good and weren't expensive.'

This data show that the constraint against *they was* is one that is deeply embedded in the grammar of Buckie, and that it is a categorical constraint preserved in various syntactic contexts, rather than being a surface effect of entrenchment of the *they were* pattern.

4   Combinatorial Variability in verbal -*s*

The split between the capabilities of full NPs and pronouns to trigger agreement is one which is well-established for other languages. For example, in Welsh, subject pronouns trigger agreement on their verbs:

(93)   Gwelodd ef y   car.
        Saw   he the car
        He saw the car.

(94)   Gwelsant hwy y   car.
           Saw-3pl   they the car
           They saw the car.

However, this contrasts with full NPs, which do not trigger subject agreement:

(95)   Gwelodd y dyn y car.

           Saw the man the car

           The man saw the car.

(96)   Gwelodd y   dynion y   car.
           Saw   the men   the car
           The men saw the car.

(97)   *Gwelsant y   dynion y   car.
           Saw-3pl   the men   the car
           The men saw the car.

We assume then that UG makes available a pronoun/NP split, as well as a singular/plural split. We can implement this in our feature system by first assuming that full NPs consist of a determiner layer selecting an NP (Abney 1987):

(98)   [$_{DP}$ D NP]

Second, we take [pronominal:±] to be a feature which can be specified on D. D's which have a positive specification for this feature occur without an overt complement (cf. the analysis of pronouns in Postal 1966; see also Bianchi 2006 for a recent argument for a basic distinction between pronominal and non-pronominal DPs). We can think of [pronominal:+] as licensing NP ellipsis, leaving behind just a bare D which is spelled out as a pronoun (see Elbourne 2005 for a recent defense of this idea). For example, the structure

of the first person singular pronoun *I*, ignoring case, would be:

(99)   [$_{DP}$ D[pronominal:+, singular:+, participant:+, author:+]] ∅]

The Welsh data above suggest that agreement on T can be sensitive to this feature, much like it is sensitive to the ϕ-features of D. The feature content of T will then be:

(100)   [T, pronominal, singular, participant, author, past]

The algorithm we developed in the previous chapter will apply to verbal *-s* inflection as follows. First we generate all single feature mappings:

(101)   (a) *[singular:+] 0, -s

(b) [singular:–] 0

(c) [participant:+] 0

(d) *[participant:–] 0, -s

(e) *[author:–] 0, -s

(f) [author:+] 0

(g) [pronominal:–] -s

(h) *[pronominal:+] 0, -s

We discard (a), (d), (e) and (h) as violations of the Reject Optionality clause of the algorithm in (8). However, the whole paradigm is not yet analysed so we then generate all two feature mappings:

(102)   a.   [singular:+, participant:+] → 0

b.   [singular:+, participant:–] → -s

c.   [singular:–, participant:+] → 0

d.   [singular:–, participant:–] → 0

e.   [singular:+, pronominal:–] → -s

33

f. *[singular:+, pronominal:+] → 0/-s

g. [singular:–, pronominal:+] → 0

h. [singular:–, pronominal:–] → 0

i. [singular:–, pronominal:–] → -s

j. [pronominal:+, participant:+] → 0

k. *[pronominal:+, participant:–] → 0/-s

l. [pronominal:–, participant:+] → inconsistent, as participants have to be pronouns

m. [pronominal:–, participant:–] → -s

n. [author:+, pronominal:–] → inconsistent, as authors have to be pronouns

o. [author:+, pronominal:+] → 0

p. *[author:–, pronominal:+] → 0/-s

q. *[author:–, pronominal:–] → 0/-s

r. *[singular:+, author:–] → 0/-s

s. [singular:+, author:+] → 0

t. [singular:–, author:+] → 0

u. [singular:–, author:–] → 0

v. [author:+, participant:–] → inconsistent, as authors have to be participants

w. [author:+, participant:+] → 0

x. [author:–, participant:+] → 0

y. *[author:–, participant:–] → 0/-s

Discarding the feature mappings that violate Reject Optionality, and those which are ruled out by Minimize Lexicon, the only remaining feature mapping is (b).

The result of the application of the combinatorial variation algorithm for verbal *-s* is then:

(103) a. [singular:+, participant:–] → -s

b. [singular:–] → 0

c. [participant:+] → 0

d. [pronominal:–] → -s

These analyse all the possible forms of both pronominal and full DP subjects. Let's see how this pans out.

(104) a. T[participant:+, singular:+, pronominal:+] (e.g. *I, you*) (c), so 100% zero inflection

b. T[participant:+, singular:–, pronominal:+] (e.g. *we, you*) (b) or (c), so 100% zero inflection

c. T[participant:–, singular:+, pronominal:+] (e.g. *he, she, it*) (a) so 100% verbal *-s*

d. T[participant:–, singular:–, pronominal:+] (e.g. *they*) (b) so 100% zero inflection

e. T[participant:–, singular:+, pronominal:–] (e.g. *the man*) (a) or (d) so 100% verbal *-s*

f. T[participant:–, singular:–, pronominal:–] (e.g. *the men*) (b) or (d) so 50/50 split.

Discarding all cases with a pronominal third person plural subject, we find that the distribution of verbal *-s* in the corpus is actually 58%, occurrence, rather than the predicted 50%. We take this as support for the combinatorial system in general, since the predicted and actual percentages are very close. Moreover, our model also allows us enough leeway to provide some

understanding of why the numbers are not absolutely identical.

Recall that the general model takes an instance of a surface variant to be the outcome of applying the function U to the pool of variants (in this case (b) and (d)). However, the output of U is also influenced by various other factors: "internal" factors such as the structural and phonological properties of the variants, the recency of a variant, what similar structures have been recently parsed, etc, and potentially social factors, such as the age of the speaker. These are the kinds of factors that enter into a classical variationist analysis and they tell us a great deal about the way that U functions. We take it that the 50/50 split predicted in (104-f) is not accurate precisely because of the impact of these other factors. See section 4.3 for further discussion of the interaction between the pool of variants and U.

The same procedure with past and present tense *be* gives the following mini-lexicon(we use [be] here as a feature which indexes this particular form; it should be seen as a shorthand for a more detailed syntactic analysis—*be* in English certainly involves a number of syntactic features):

(105)   a.   [singular:–, be, present:+] are

            b.   [author:–, be, present:+] are

            c.   [singular:+, author:+, be, present:+] am

            d.   [singular:+, participant:–, be, present:+] is

            e.   [pronominal:–, be, present:+] is

(106)   a.   [singular:+, be, present:–] was

            b.   [singular:–, be, present:–] were

            c.   [participant:+, be, present:–] was

            d.   [author:–, be, present:–] were

            e.   [author:+, be, present:–] was

f. [pronominal:–, be, present:–] was

An interesting property of this approach is that we have two sources for the variation. The lexical assembly of features includes [pronominal:–] in Buckie and Welsh grammar but not in many other varieties of English, and this claim about the inventory of features and how features are specified on lexical items is an example of the approach to parametric variation that stems from Borer (1984), while the way that the algorithm gives multiple forms for a single meaning via underspecification in the mapping between feature structures and morphological forms is a formal method of capturing what we might call Labovian variation. The claim we make in this paper is that a speaker's lexical inventory can vary on either dimension.

4.1 Further Evidence: syntactically complex pronouns

Independent evidence for this system comes from the behaviour of pronouns that have nominal complements. Verbal -s is impossible with bare *you*:

(107) *You **is** here   Buckie judgment

(108) *You **kens** him   Buckie judgment

Recall from section 1.1 that preterite *be*, however, is variable in form with both *you* and *we*. We give examples from second person here:

(109) He says 'I thocht you were a diver or somethin.' (middle aged male)
       He said 'I thought you were a diver or something'.

(110) 'Aye, I thocht you was a scuba diver.' (middle aged male)
       'Yes, I thought you were a scuba diver.'

The *was* variant is also possible with plural *you*, where the plurality can be forced by a floating quantifier:

(111)   You **was** all here.   Buckie judgment

(112)   You **all** was here   Buckie judgment

However, even if we force a plural version of *you* via a floating quantifier, verbal *-s* is still impossible:

(113)   *You **is** all here   Buckie judgment

(114)   *You all **is** here   Buckie judgment

(115)   *You all **kens** him.   Buckie judgment

The generalization is then that pronominal *you* does not occur with verbal *-s*, and this follows from the lexicon given above: The only lexical items that match with the featural specification of *you* are (103) (b) and (c), both of which have a null realization.

We can now test the impact of pronominality, by taking the plural version of *you* with an overt plural nominal complement. This will have the same feature specification, with the exception that it will be [pronominal:–], rather than [pronominal:+], since its NP complement is not elided. The prediction is that verbal *-s* should then become possible, as (103) (d) will then be a possible match, and this has verbal *-s* as its realization. We then predict variability for such cases. This prediction is confirmed:

(116)   He would be twenty-nine and you ones **are** thirty aye. (older male)

(117)   I used to think 'Christ, you ones **is** old' ken (younger female)

(118)   You ones **kens** him   Buckie judgment

This is also true with other nominals besides the pro-form 'ones'. For example, we can use the nominal *bairns*, 'children', with the same effect:

(119)   You bairns **is** for it/You bairns **are** for it.   Buckie judgment

This is straightforwardly predicted by our analysis, as in these cases T will have an [pronominal] feature which is valued as minus by the [pronominal:–] feature on the D that takes *ones/bairns* as a complement.

4.2   Demonstratives

A second case which leads to the same conclusion is the syntax of demonstratives in Buckie. The use of demonstratives to express plurality is highly variable, with the singular form of the demonstrative being variably used with plural nouns. This is true for both proximal and distal demonstratives:
Plural proximate

(120)   a.   I'd a' **these** cuttings. (older female)
               I had all these cuttings.

          b.   My mam had all **this** stories o' ootside folk. (middle aged female)
               My mother had all these stories about foreign people.

Plural distal

(121)   a.   She's one of **those** bonny big houses. (older female)
               She has one of those lovely big houses

          b.   But I 'd piles of **that** photos of the dancing. (older female)

But I had piles of those photographs of the dancing

    c.   That was ain o' **them** grogain suits. (older female)

          That was one of those grogain suits

    d.   It was a' bonny, able drifters, **thon** steel drifters. (older male)

          They were all good able drifters, those steel drifters

Here we have a number of variants used, including the relic form *thon/yon* which is still used in some Scottish dialects and the pandialectal use of *them*. (122) shows how these various forms appear in different contexts:

(122)   Distribution of plural proximate demonstratives

	N	%
these	33	34
this	65	66

(123)   Distribution of plural distal demonstratives

	N	%
those	7	4
that	132	69
thon	40	21
them	13	7

    Semantically plural demonstratives have high rates of use of apparently non-plural forms in both proximate (66% *this*) and distal contexts (69% *that*). In addition, there are relatively high rates of the older distal form *thon* in this context.

    In Adger and Smith (2005) we provided an explanation for verbal *-s* vari-

ability with plural DPs which was based on the idea that these demonstratives show us that D's can be formally singular in this dialect. We assumed in that paper that T is specified as follows (where the features [num] and [pers] are roughly equivalent to the features [singular] and [participant] used in the current paper):

(124)   T[$u$num:, $u$pers:]

If the subject is a DP, it will variably have the specification [$u$num:sing]/[$u$num:pl] on D, perhaps depending on whether the feature is lexically specified as valued or not:

(125)   T[$u$num:, $u$pers:] ... DP[$u$num:sing, pers:–] $\rightarrow$
         T[tense:past,~~$u$num~~:sing,~~$u$pers~~:–] ... DP[~~$u$num~~:sing, pers:–]
         The mothers dances

(126)   T[$u$num:, $u$pers:]   DP[$u$num:pl, pers:–] $\rightarrow$
         T[tense:past, ~~$u$num~~:pl,~~$u$pers~~:–] ... DP[~~$u$num~~:pl, pers:–]
         The mothers dance

This analysis relies on the idea that there is a meaningless feature which forces singularity in agreement irrespective of whether the reference of the DP is singular or plural. Obviously, this is not the ideal analysis from a theoretical perspective. Moreover, the analysis makes a prediction: when the demonstrative is singular, the verb will also be singular. This prediction turns out not to be true, a fact that we did not notice in our previous paper. We have both singular verbs and plural ones with the *that* form:

(127)   When **that wards is** up, they 've got to bide (middle aged female)
         'When those wards are up, they have to stay.'

(128)   A lot of families does na get what **that cats get**. (older male)

         'Many families don't get what those cats get.'

It appears we must look elsewhere for our explanation.

The alternative we pursue here relies on the idea that the demonstratives can be underspecified for number features, an idea which unifies the treatment of demonstrative agreement with the underspecification approach we take to variable agreement in the verbal system. Under such a hypothesis, we can distinguish the demonstratives via the specification [distal:$\pm$], allowing the form *that* to bear only this feature:

(129)   a.   [distal:+] that, thon

         b.   [distal:−] this

         c.   [distal:+, singular:−] them (those)

         d.   [distal:−, singular:−] these

This will allow the form *that* to appear with plural nouns, since the demonstrative is unspecified for agreement. It will also predict that verbal *-s* with such an DP should be possible, since we can pick T specified with [pronominal:−]. This is correct, as can be seen from (127) and (130):

(130)   I would rather go to a decent hotel, cos some of **that B and Bs is**
         just rubbish(younger female)

         'I'd rather go to a decent hotel, because some of those B(ed)-and-
         B(reakfast)s are just rubbish.'

Of course, since the noun itself can be plural, an alternative choice of T [singular:−], will give rise to plural agreement on the verb, even though the demonstrative has the form *that*:

(131)   A lot of families does na get what **that cats get**. (older male)

42

This analytical system captures the patterns we have seen so far, however, there is an interesting twist for any approach to this data that relies on underspecification: when there is no N complement to the demonstrative, *that* can have plural reference but categorically appears with verbal *-s* when it is in subject position.

We can see that *that* can have plural reference in object position by looking at contexts of use of this work in the corpus. For example, during one of the interviews for the corpus, the interviewee was looking at people in a photograph. Pointing out a subgroup of people, she said:

(132)   I na ken nane o' **that** nor nane o' **that**

        I don't know any of those or any of those.

In cases where there is a bare demonstrative with plural reference in subject position, we can see that 100% of them have singular agreement:

(133)

Plural this/that		other plural NP	
N	% verbal *-s*	N	% verbal *-s*
20	100	374	36

This is odd on an underspecification account, since the demonstrative does not bear [singular:+] and, given it has no NP complement, it is [pronominal:–]: apparently there is no source for verbal *-s* marking.

One potentially interfering factor here is the possibility of having singular agreement with *that* in subject position when the demonstrative is not fully referential. This is also possible in standard English:

(134)   **That's** the girls I meant.

However, this is only possible in Standard English when the *be* auxiliary is contracted. Compare question and negation versions of (134) in standard

English:

(135)   ***Is** that the girls you meant

       for Are those the girls you meant?

(136)   *That **isn't** the girls I meant.

       for Those aren't the girls I meant.

In Buckie, even in these non-contractable environments, it is still impossible to have plural agreement with a plural referring demonstrative:

(137)   ***Are** that the quines you meant

       Are those the girls you meant.

(138)   *That **are** na the quines I meant.

       Those aren't the girls I meant.

The generalization is, then, that in the absence of N, non plural demonstratives have verbal -*s* categorically.

How can we capture this behaviour? We can't just keep the simple checking system we have developed up to this point, since then either a singular or a plural T should be able to appear with [distal:+], but we have seen that only the T that is realized as -*s* is possible.

There are two analytical possibilities we would like to sketch here, but we will not choose between them. One approach compatible with the theory we have developed so far, is to take the algorithm to give Distributed Morphology style Vocabulary Items (as in Adger and Smith 2005. See Nevins and Parrott (this volume) for further discussion), and to allow any features which are not valued in the syntax to be valued at the interface. More specifically, we stipulate that the syntax/morphology interface supplies unmarked feature

44

values for all unvalued features.[6] For the case at hand, we have the following representation in the syntax (note that we need to take the plus value of [singular] to be unmarked and we also need to take the plus value of [pronominal] to be unmarked—the marked version of D takes an NP complement):

(139)   T[singular: , participant: , author: , pronominal:] ... D[distal:+]

Nothing is valued, all features are specified with unmarked values at the interface:

(140)   T[singular:+, participant:–, author:–, pronominal:+] ... D[distal:+]

In this situation, T spells out as *-s* as this is the only phonological form that matches the [singular:+] feature or the [pronominal:–] feature, and D is, of course, spelled out as *that*. It follows that there is no variation.

In contrast to this, when the demonstrative has a plural nominal complement, then this entails a valued [pronominal] feature. Assuming that the [singular:–] feature of N values the [singular] feature of T, we have:

(141)   T[singular: , participant: , author: , pronominal:–] ... D[distal:+]
          N[singular:–] →
          T[singular:–, participant: , author: , pronominal:–]

Filling in unmarked values at the interface gives:

(142)   T[singular:–, participant:–, author:–, pronominal:–]

Now T spells out as either -0 or *-s*, since either the [singular:–] feature or the [pronominal:–] feature matches, correctly capturing the variation.

---

[6]There is an issue here with case features, which must be syntactically satisfied. Obviously, we do not want to fill in default values for these at the interface, since we would then be unable to rule out examples where a case feature is not appropriately licensed.

(143)   That cats get/That cats gets

Of course, if the complement N is singular, the result of valuation will be as follows:

(144)   T[singular: , participant: , author: , pronominal:–] . . . D[distal:+]
         N[singular:+] →
         T[singular:+, participant: , author: , pronominal:–]

Filling in the unmarked values gives:

(145)   T[singular:+, participant:–, author:–, pronominal:–] — only -s

which correctly predicts no variation.

(146)   That **gets**/***get**.

The main weaknesses of this proposal are the stipulations about markedness values for features and the system predicts that default forms will appear whenever a feature is unvalued, which is potentially problematic for case features.

A straightforward alternative would be to take demonstratives as non-pronominal, even when they lack a nominal complement. This idea would treat the demonstratives as syntactically complex, as follows:

(147)   [$_D$ th- [$_{NP}$ -is]]

(148)   [$_D$ th- [$_{NP}$ -at]]

Here the -is and -at parts of the demonstratives are actually morphemic, and perhaps specify the [±distal] distinction. Such analyses for demonstratives have been proposed by Delsing (1993) and more recently worked out by, for example, Leu 2006. If this is the right syntactic decomposition of demon-

46

stratives, then the overt NP would lead to [pronominal:–] marking on the D, and to categorical verbal *-s*. We leave the choice of these two different approaches open here.

4.3   Existentials

The final kind of agreement phenomenon we address here is the existential expletive *there*. In Buckie, present tense existential sentences with *there* in the structural subject position almost always have verbal *-s* on *be*.

(149)   There**'s** three possibilites.' (older male)
          'There are three possibilities.'

(150)   There**'s** the microwaves and there**'s** thon other things. (older male)
          'There are the microwaves and there are those other things.'

The overall distribution of the pattern for present tense existentials with plural associates, compared to non-existentials with plural subjects is given in (151):

(151)

existential		plural NP	
% verbal *-s*	N	% verbal *-s*	N
99.5%	210	36	380

However, unlike the case of verbal *-s* with pronominal subjects, there is still variation in existentials, with one example of agreement in the corpus:

(152)   There **are** very few conversations between the two (middle aged male)

Moreover, if we test an example like (153), we find that *are* is acceptable. This contrasts dramatically with the unacceptability of (154):

(153)   There **is/are** folk that ken him   Buckie judgment

         'There are people that know him.'

(154)   a.   They ***is/are** affa fine.   Buckie judgment

               'They are really fine.'

         b.   That ***is/are** both affa fine. Buckie judgment 'Those are both

               really fine.'

Furthermore, if we look at preterite *be*, we find that both *was* and *were* occur with a plural subject (90% of 162 examples of expletives with plural associates and preterite *be* appear with *was*):

(155)   There **was** no ministers in the Church of Christ.   (middle aged

         male)

(156)   There **were** people coming and going. (middle aged female)

This strongly suggests that T is able to agree with the plural subject, even in the presence of an expletive.

Unlike in the previous cases, we don't have a categorical/variable split with existentials which have plural associates, rather we have variability Although an overwhelming majority of expletive *there* sentences which have plural thematic subjects appear with a non-agreeing version of the copula, the grammatical system itself does not rule out agreement.

The system we have developed straightforwardly captures this. Recall the mini-lexicon for preterite *be*:

(157)   a.   [singular:+, be, present:–] was

         b.   [singular:–, be, present:–] were

         c.   [participant:+, be, present:–] was

         d.   [author:–, be, present:–] were

e. [author:+, be, present:–] was

f. [pronominal:–, be, present:–] was

We discount the feature [author], which is not specified on full DPs. Furthermore, a full DP must be [participant:–], so we discount (c). We are left with (a), (b) and (f), and a prediction that there should be 66% occurrence of *was* in this context.

As just noted, what we actually find is 90% *was*. In the system developed in Adger (2006), Adger (2007) and in this paper, there are two possibilities for tackling this disparity. One would be to say that it is a function of the various other factors that enter into the specification of U. For example, the fact that the semantic subject has not yet been processed at the point where choice of the form of *be* is made, and so no number feature is yet available, may lead to a preference for using the version of preterite *be* that lacks a number feature (that is, (f)). Alternatively, or perhaps additionally, it may be that *there* is featurally [pronominal:–], perhaps because it has a locative complement to its D or because it originates as part of the associate (as in Kayne 2008). If this is correct, then producing *there* might prime choice of (f), since both will have [pronominal:–] specification, even though there is no formal syntactic agreement relation between the expletive and the finite verb.

The alternative possibility is that the verb *be* takes into account the featural composition of the expletive in its mini-lexicon. That is, as well as (157), we also have something like (158):

(158) [expletive:+, be, present:–] was

However, languages do not seem to be sensitive in general to a specialized 'expletive' agreement form. When verbs display a special agreement form in

the presence of an expletive, that form is simply the default. It is probably the case that there is no [expletive] feature, and that UG thereby rules out the analysis in (158).

For present tense *be* in expletive constructions, one important factor not considered yet is the effect of contraction. The contracted present tense auxiliary appears as *-s* 100% of the time in existentials. If we confine our attention to non-contracted examples, there are only five of these, but interestingly, even in this extremely small number, we still have the one case of agreement noted above (that is, four *there is DP*[plural] and one *there are DP*[plural]. On the assumption that the expletive has no relevant features, this variation in agreement is what we would expect given the system outlined so far. The relevant mini-lexicon is repeated here:

(159)   a.   [singular:–, be, present:+] are

   b.   [author:–, be, present:+] are

   c.   [singular:+, author:+, be, present:+] am

   d.   [singular:+, participant:–, be, present:+] is

   e.   [pronominal:–, be, present:+] is

Once again, the processing of the sentence may increase the probability of choosing (e) over (a), but of course the numbers are too small to tell.

Finally, we have the question of why 100% of contracted present tense *be* with expletive subjects are marked with verbal *-s*. We take this to be an independent effect, possibly because the expletive and auxiliary are stored as a single form which is 'simultaneously' inserted into both the subject and auxiliary nodes (or alternatively, which selects a zero copula). The analytical options are too various to choose in this case, and the data we have do not push us in any particular theoretical direction.

5 Conclusion

Overall, this paper has argued for an analysis of agreement variability which extends our previous lexicalist, combinatorial, approach, rather than appealing to different structural licensing positions (see, Henry 1995 and Tortora and den Dikken 2008) or to surface, construction based, constraints (Pietsch 2005). There are no variable rules of the grammar, simply invariant Agree, Merge and Move (Chomsky 2000) working on a single lexicon. The eventual source of the variability is the structure of this lexicon, which is derived from a very general algorithm that links syntactic features with morphological forms. This set of feature-form mappings, together with the invariant syntactic operations, leads to a potential pool of variants from which the systems of use select a particular item on any occasion of utterance. The structure of this pool of variants is itself an important factor that impacts on the final choice of a form, inasmuch as a particular surface form may occur with more than one feature bundle in the pool, raising the probability that that form will be found more frequently. The choice of variant is also subject to pressures of processing and sociolinguistic status, neither of which are part of the grammatical system proper (see Adger 2007).

We have also shown that the two kinds of variation discussed by linguists under the rubrics of 'parametric variation' and 'the linguistic variable' can be brought together in a single theoretical model, that outlined immediately above. We hope that this model might be of some use in deepening the connections between generative and sociolinguistic approaches to the systems that underlie language variation.

References

Abney, S. P. 1987. The English noun phrase in its sentential aspect. Ph.D. thesis, MIT, Cambridge MA.

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42:503–530.

Adger, David. 2007. Variability and modularity: A response to Hudson. *Journal of Linguistics* 43:695–700.

Adger, David and Smith, Jennifer. 2005. Variation and the Minimalist Program. In Leonie Cornips and Karen Corrigan, eds., *Syntax and Variation: Reconciling the Biological and the Social*, 149–178, Amsterdam: John Benjamins.

Bianchi, Valentina. 2006. An asymmetry between personal pronouns and other DPs. *Snippets* 13:http://www.ledonline.it/snippets/.

Borer, Hagit. 1984. Restrictive relatives in Modern Hebrew. *Natural Language and Linguistic Theory* 2:219–260.

Börjars, Kersti and Chapman, Carol. 1998. Agreement and pro-drop in some dialects of English. *Linguistics* 36:71–98.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels, and Juan Uriagereka, eds., *Step by Step: Essays on Minimalist syntax in Honour of Howard Lasnik*, 89–115, Cambridge, MA: MIT Press.

Delsing, Lars-Olaf. 1993. The interal structure of noun phrases in Scandinavian languages. Ph.D. thesis, Lund, Sweden.

Elbourne, Paul D. 2005. *Situations and Individuals*. MIT Press.

Goldberg, Adele E. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences* 7:219–224.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In *MIT Working Papers in linguistics 30*, 425–449, MIT.

Halle, Morris and Marantz, Alec. 1993. Distributed Morphology and the pieces of inflection. In Ken Hale and Samuel Jay Keyser, eds., *The View from Building 20*, 111–176, Cambridge, Massachusetts: MIT Press.

Harley, Heidi and Ritter, Elizabeth. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78:482–526.

Henry, Alison. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.

Henry, Alison. 2005. Idiolectal variation and syntactic theory. In Leonie Cornips and Karen Corrigan, eds., *Syntax and Variation: Reconciling the Biological and the Social*, 109–122, Amsterdam: John Benjamins.

Johnstone, Barabara. 1987. "He says ... so I said": Verb tense alternation and narrative depictions of authority in American English. *Linguistics* 25:33–52.

Kayne, Richard. 2008. Expletives, datives and the tension between morphology and syntax. In Theresa Biberauer, ed., *The Limits of Syntactic Variation*, 176–217, John Benjamins.

Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Basil Blackwell.

Labov, William. 2000. *Principles of Linguistic Change. Volume II: Social Factors*. Oxford: Basil Blackwell.

Leech, Geoffrey. 1987. *Meaning and the English Verb*. London: Longman, second edition.

Leu, Thomas. 2006. These HERE demonstratives. *U. Penn Working Papers in Linguistics* 13:141–154.

Marr, David. 1982. *Vision*. San Francisco: Freeman and Co.

Murray, James A.H. 1873. *The Dialect of the Southern Counties of Scotland: Its Pronunciation, Grammar and Historical Relations*. London: Philological Society.

Nevins, Andrew and Parrott, Jeffrey K. 2008. Variable rules meet Impoverishment theory. *Lingua* this volume:xx–xx.

Parrott, Jeffrey. 2006. Distributed Morphological mechanisms of Smith Island *weren't* levelling, ms. Georgetown University.

Pertsova, Katya. 2007. Learning form-meaning mappings in the presence of homonymy. Ph.D. thesis, UCLA.

Pietsch, Lukas. 2005. *Variable Grammars: Verbal Agreement in Northern Dialects of English*. Tübingen: Max Niemeyer Verlag.

Poplack, Shana and Tagliamonte, Sali. 1991. Theres no tense like the present: Verbal -*s* inflection in early Black English. In Natalie Maynor Guy Bailey and Patricia Cukor-Avila, eds., *The emergence of Black English: Text and Commentary*, 275–324, Amsterdam: John Benjamins.

Postal, Paul. 1966. On so-called 'pronouns' in English. In F. Dinneen, ed., *Monograph Series on Language and Linguistics*, volume 19, 177–206, Georgetown University Press.

Smith, Jennifer. 2000. Synchrony and diachrony in the evolution of English: Evidence from Scotland. Ph.D. thesis, University of York.

Tagliamonte, Sali. 2000–2003. Back to the roots: The legacy of English dialects, ESRC Grant Number R000239097.

Tortora, Christina and den Dikken, Marcel. 2008. Subject agreement variation: Support for the configurational approach. *Lingua* this volume:xx–xx.