Variability and Grammatical Architecture[1]

David Adger

December 2010

1. Introduction

What I want to do in this paper is to connect with the some of the work that Sjef Barbiers discussed in his paper --- the kind of work that sociolinguists do. I've been interacting with sociolinguists for a good 7 or 8 years now, and I've learned a lot from them. However, one of the things sociolinguistics largely lacks is an interest in syntactic models that they could work with. The syntactic models we generative linguists provide them with do not seem to be considered by sociolinguists to be germane to their interests. So I've been interested recently in trying to address some of the data that the sociolinguistics are interested in tackling with a view to exploring how our models of syntax, devised with quite different goals in mind, might be useful for the analytical purposes of sociolinguistics. What will be crucial here is not the sociolinguistic side of relevant data, but rather the question of how to theoretically model community-internal variability, and more specifically person-internal variability.

There are models out there of this kind of stuff: Tony Kroch has had for a long time the idea that there are competing, multiple grammars (Kroch 1994), and more recently Andrew Nevins and Jeff Parrott have been working out a proposal where they propose probabilistic versions of the Distributed Morphology operation of Impoverishment to deal with this kind of variation, resurrecting the basic proposals of Labov (1969), in a modern theoretical setting (Nevins and Parrott 2010).

---

[1] Many thanks to the audience at the Barcelona Workshop on Linguistic variation in the minimalist framework, January 2010, and to Terje Lohndal and Peter Svenonius for comments on an earlier draft. Thanks too to Jenny Cheshire, Sali Tagliamonte and Jennifer Smith for interim sociolinguistic advice.

I have a slightly different line of attack that I'm going to try and pursue here, and this line of attack connects again with issues that have been mentioned by a bunch of people here. Cedric Boeckx, in his talk, raised questions of 'memorization' or 'idiomaticity', however that might fit into what we do, and how that relates to non-compositionality when a piece of morphology is linked to a certain piece of syntax. Sjef Barbiers talked about the same kind of issue but called it something slightly different. He called it conventionalization and also raised questions of frequency, and that's something that I'll talk about here as well. I'm hoping that by the time I'm done with all this, we'll be able to connect this really lovely, beautiful, minimalist syntax to the kind of things that people outside of syntax have an interest in.

2. Internal Factors

The particular empirical domain I'll focus on is what sociolinguists call internal factors. These are individual-internal factors connected to grammatical features and structures (or possibly to the processing of grammatical features and structures) that impact upon the choice of a particular variant in a particular context. I'll focus on two examples, because I think I understand them. The first is from Buckie, a little town in the North East of Scotland. It's work that I did with Jennifer Smith, who is from Buckie and a sociolinguist who has spent some time working there. As an example of the phenomenon I am interested in, have a look at (1):

(1a) Buckie boats were a' bonny graint.

'Buckie boats were all nicely grained'

(b) The mothers was roaring at ye comin' in

'The mothers were shouting at you to come in'

All speakers in this particular speech community allow this kind of variation. What we see here is that in both cases we've got plural subjects, and in one case we have an apparently agreeing auxiliary *were* and in the other case we have *was*. We find this

pattern also with certain pronouns. I've given you a table with the frequencies that we find in the corpus in (2).

(2)

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

For 2nd person singular and 1st person plural, we see that two thirds of the time the form is *was*, whereas in Standard English we expect *were*. All speakers allow this variation, although we don't have sufficient numbers in this data to show that all speakers have this pattern of distribution, so this is only available at the community level. Similarly, for plural noun phrases, examples like 1a and 1b, what we find is 56 per cent of *was*. Again all speakers allow this. We can't analyze this by just proposing that, in Buckie, there is simply the possibility of substituting *was* in place of Standard English *were*. We cannot say that because, as can be seen in the table, the 3rd plural pronoun *they* has only the *were* form, apparently displaying obligatory agreement.

Variable agreement with plural NP subjects appears also with main verbs in the present tense, as well as with present tense *be*:

(3a) The money that that lads *is* making is nae canny compared to what I made.

"The money that those boys are making isn't good compared to what I made"

(b) Aye, and some of them *are* drawing dole on three, four different names.

"Yes, and some of them are receiving unemployment benefit using three or four different names"

(4a) What bairns walk any distance?

"What children walk any distance?"

(b) When they go back, the teachers asks them to write something and they send them till's

"When they go back, the teachers ask them to write something and they send them back to us)"

What we see here is a plural subject in both (a) and (b) examples, but variation in whether the verb is marked with the *–s*.

Furthermore, once again with the pronominal *they*, the present tense main verb, just like the past and present auxiliary, appears with plural agreement (i.e. no *–s* is possible). There is just 0.6% appearance of *–s* with *they* in resent tense (N=1271), and Adger and Smith (2010) show that these are all narrative presents with the verbs *say*/*tell*, suggesting that *–s* here is performing a distinct function. This conclusion is backed up by judgment data we carried out on the individuals who were recorded for the initial corpus gathering phase of the research. We took cases where *–s* appeared with a plural subject and swapped in *they* in the same context. This led to uniform rejection by our speakers:

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

(6) *When they gets home Buckie judgment

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

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

(9) *They is up in Elgin Buckie judgment

This phenomenon of variable agreement which becomes obligatory in the presence of a pronoun has been tackled in the literature by various people: Alison Henry (Henry 1995) had a view on the manifestation of this phenomenon in Belfast English which related it to differential possibilities in where pronominal vs non-pronominal subjects could be syntactically placed and how this syntactic placement impacts on agreement morphology. Tortora and Den Dikken (2010) extend this idea to Appalachian English, although more recently Bernstein and Zanuttini (2010) have proposed a featural perspective on this close to what Jennifer Smith and I argued for. And as I said, Nevins and Parrot also have a view on this that ties it to variable application of impoverishment rules tied to markedness effects.

Adger and Smith (2005) proposed a basic architecture for such variable phenomena that was further developed in Adger (2006, 2007) and Adger and Smith (2010). What we argued is that the explanation of the variability is to be tied only to the featural composition of the agreeing finite element, and it has nothing to do with the syntactic position of the relata. Adger and Smith (2010) provide a number of arguments that this is the right way to approach the Buckie data. An important claim that we made is that you don't need to do anything particularly special to the architecture of current theoretical models of syntax. We already have all the technology at our disposal to tackle this issue: the featural representation of functional categories and the matching of morphological feature matrices to syntactic ones (as in, for example, vocabulary items in Distributed Morphology). Schematically, take some functional category X, specified with features F, G and H, and take three vocabulary items each of which matches one of these features:

(10a) X[F, G, H]

(bi) [F] < -- > i

(bii) [G] < -- > j

(biii) [H] < -- > j

The syntax builds X, and then X needs to be `vocabularized'. Any of (bi-iii) will match, and so any can vocabularize X. On the assumption that the choice of vocabulary item is random, then we expect each of (bi-iii) to be inserted 33% of the time. However, if there is homonymy in the vocabulary list, as in (10b) above, then the surface forms will be unevenly distributed in terms of their frequency. In the schematic example given here, we will see 66% j to 33% i, approximately.

The idea, then, is that the variability arises because of an underspecification of the relation between the syntactic feature bundles and their morphological exponents; variability in the frequency of the forms emerges because of homonymy in the set of relevant exponents, and I proposed an algorithm (Adger 2006) that generates the set of exponents from the input, and termed this approach to variation *Combinatory Variability*.

The system is different in one important respect to standard models of the morphology syntax interface that involve underspecification, though: it does not incorporate a Subset Principle for resolving competition between lexical entries, at least not as a grammatical principle. The algorithm in Adger (2006) does reduce a certain amount of homophony, and therefore incorporates the intuition that a morphological learner generalizes beyond its input data, but the algorithm produces what Adger (2007) calls a Pool of Variants, any of which can in principle be inserted. I showed that this was the right result in a number of cases.

The categoricality of agreement with pronominal *they* arises because of parametric variation in the usual sense: Buckie, like Welsh, but unlike standard English, bundles a pronominal feature in with person and number features in its finite agreement head (see Adger and Smith 2010 for details). The checking of this pronominal feature is what leads to the impossibility of *they* with *–s*.

However, this basic story is not quite satisfying. It lacks a place for the internal factors that I mentioned above in (at least) two important ways. The first is frequency effects: it's well known in the variationist literature that the frequency of a particular lexical item can impact on how robust the variability is: more frequent verbs, such as auxiliaries, appear

more frequently with *–s* (see the discussion in the final section on the frequent verbs). In fact, it is generally true that variation is more robust for items that are more frequent (this is strikingly attested in historical change, see for example Ellegard 1953 on *do*). So if you look at, for example, frequent verbs like *be* or *do* (even just focusing on their main verb uses), vs. less frequent verbs like *kick* or *interrupt*, there is a higher proportion of *-s* with the more frequent verbs than there is with the less frequent verbs. There is no obvious way in the combinatorial variability system to capture that.

The second effect is what I raised at the end of Sjef Barbiers's talk. As first noted in the Labovian tradition (e.g. Labov 1969), it turns out that there are effects that the grammar has on how frequent variants are; that is the grammatical features don't merely determine whether a structure is grammatical or ungrammatical. For variable cases what these grammatical properties do is rather impact on the probability that one variant rather than another will appear. That is, beyond the social effects on variability which are well known, there are also grammatical effects on variability. In general these effects are more `global' than simply the features of a lexical item. So this is the problem of internal factors, that I mentioned above.

For example, take verbal *–s* (as the appearance of *–s* with plural subjects is often known): it turns out that in Buckie it's more likely that verbal -s will appear in relative clause constructions than it is that it will appear in non relatives (Selected as significant, with Varbrul factor weights of .64 and .48 respectively, with a range of 16).

(11)

Subject relatives vs. other			
Subject relatives		other	
%	N	%	N
42	53	**34**	327

So the syntax matters in this case. The interesting thing about this kind of effect is that the syntactic context is impacting not on the grammaticality or ungrammaticality of a particular form, but rather on the probability of that form. All the context does is push the observed frequencies in one direction or the other.

Again the combinatorial variability model as it stands apparently has no obvious way of doing this. What the model does is just this: the agreeing verb form is just V plus some set of agreement features of plus tense features. But we seem to need more information in there: specifically whether the clause that the verb appears in is a relative clause. The question is how we get that information into the system.

More generally, though, the issue of internal factors raises a theoretical question: are there formal limits on what internal factors can be and if so, what characterizes these limits?

These are the issues that in the theory of variation that I want to tackle: variability, frequency, internal factor effects.

3. Connecting Syntactic Structure to Variable Form

What I want to do now looks extremely disconnected from this. I want to look next at Merge. I'm going to develop a theory of how Merge works, which is going to give us a distinction essentially between extended projections and specifiers, where the information high up in an extended projection can have an effect on what is vocabularized low in that projection. The way that spellout, or morphological vocabularization to be more exact, is executed in this theory will give us a handle on frequency effects as well as on the limits of the impact of internal factors.

Merge is standardly taken to be binary, I have given below a (slightly simplified) definition from a recent paper by Chris Collins and Ed Stabler (Collins and Stabler 2009), who formalize certain aspects of minimalism, which clarifies the stipulations inherent in Merge:

(12) Let W be a workspace and let A, B be syntactic objects where A, B belong to W and A and B are distinct. Then, external merge of A, B is equal to the set {A,B}.

Note that this definition (and all others I'm aware of that are set theoretic), has a crucial stipulation in it, that A and B are distinct. One might ask why that stipulation is there. If we remove it then we have the Merge of A and A being simply the set {A}. In fact, the idea of removing this stipulation and allowing *Self Merge* has been proposed already in the literature. Guimaraes (2000) suggested it and then spent some time ruling certain applications of Self Merge out. Richie Kayne suggested it in his recent Antisymmetry and the Lexicon paper (Kayne 2008) as a way of distinguishing nominal from verbal extended projections. Both Guimaraes and Kayne however strictly limit the application of Self Merge. For example, Kayne says in a footnote in his paper, "alongside {x} there seems to be no need for {{x}}, which would be unavailable in principle if every merge operation must directly involve a head."

However, why do we need to say that? Let's jettison that assumption as well. If we do this, then what emerges is a unary branching structure of the kind that we all thought we'd ruled out when we started doing bare phrase structure. However, I think unary branching is just given to you by the system, if you remove these various stipulations. So what we get then is essentially something like (13). We have a root, and then we essentially have a sequence of elements which are the set containing that root, the set containing the set containing that root, and so on:

(13)
x

|

x

|

x

|

root

The immediate issue to address here is that of the label of these various constituents. That might be one way to rule out iterated self-merge: there is nothing to give the output of self-merge a label. However, I think that Collins in his 2002 paper actually was giving an argument not for the elimination of labels, but rather for the redundancy of triggered merge and labels. So I'm going to suggest we jettison triggered merge (following Chomsky 2004) and instead have a labeling system.

In such a system, labels are given exogenously, by something outside the core computational system: a language-external cognitive schema applies to the fractionated structure that is given to you by the syntax, thereby labeling that structure (cf Chomsky 2007, 14). For example, assume an antecedently given conceptual distinction between kind, sortal and substance concepts (see, for example, Xu 2010, Li, Dunham & Carey 2009). These cognitive distinctions are represented as linguistic distinctions via labels of pieces of structure (e.g the label Ki(nd) or Cl(assifier) from Svenonius 2009, following Zamparelli 2000 and Borer 2005) and the relationships between the cognitive distinctions are represented as relations between labeled structures (e.g dominance, identity). Syntax provides a structure, which is labeled by cognitive schemata: since these schemata are finite in size, the structures labeled by them are also finite, so syntax using just self-Merge provides a finite set of labeled structures (effectively, fragments of extended projections). Syntactic computation (feature matching, and binary Merge) operates over these syntactic structures, not cognitive ones, which is why languages have a limited and universal set of syntactic operations.

However, I take it to be false that the functional structure is identically represented in all languages for every root, or we would expect no variation at all. Rather language acquirers make parochial selections from the finite set of structures given by self-Merge and labeling during the process of language development. These parochial selections are what bear the 'parametric' properties discussed by Rizzi in his contribution: instructions for syntactic actions (what Adger and Svenonius 2010 call 'second order features'). In a sense, these parochial selections are a kind of 'bundling' of extended projections, giving a finite set of sequences of categories which may bear an instruction to the spellout

systems. For convenience, I'll call this set the s(equence)-lexicon, and it is what is responsible for category level generalizations about any particular language, effectively playing the same role as functional category feature bundles in standard minimalist approaches.

For example, a V2 language will have in its s-lexicon a sequence of categories that runs from V through Fin to C which bears an instruction that the phonology of the entire projection line is to be linearized at C (see below).

That, then, is one kind of Merge, which results in a set of sequences which are rooted (and thereby directed, since if there is an asymmetry between the two terminae of a sequence, you have a directed sequence) and labeled, some subset of which is extracted during acquisition to provide an s-lexicon.

The other case will just be our familiar set-merge in both its internal and external manifestations. I suggest that what's particularly linguistic in both self-merge and in set-merge is that there's a matching requirement for both of them (that is, Merge is licensed to apply, although not triggered, when two elements match, perhaps with the edge-feature of Chomsky (2007) being an existential precondition for such matching). In self-merge the two inputs to the operation obviously match since they are the same thing, giving token-identity. For set-merge, our familiar binary merge, I suggest that what we have is essentially type identity: that is, identity of (some subpart of/features of the two) labels. This results in a system with generalized spec-head agreement (Chomsky 1993, Brody 2000, Koopman 2006).

What we have now derived is a system which generates (i) a (possibly finite) set of unidirectional unary branching sequences of labeled nodes (traditionally, sequences of functional categories, or extended projections, Grimshaw 1991) some subset of which is 'extracted' during acquisition and to which is added spellout instructions and (ii) an infinite set (because of recursion) of tree structures whose nodes may branch (in which case we have a specifier structure) or not. Interestingly, this system, although derivational, is very similar in its yield to Brody's (2000) Mirror Theory system of phrase structure. It is a consequence of both systems that the syntactic complement relation is

reserved for extended projections, with binary specifierhood always involving spec-head agreement (that is, feature sharing).

Take, for example, a fragment of the Buckie sentence in (4b):

(14) the teachers asks them

The verb root *ask* here Self Merges giving a syntactic object that contains just the root. This object is then labeled as a V (presumably there is a small set of choices as to how roots are labeled, perhaps at most N, V and A (if Baker 2003 is correct)). If Borer (2005) is correct, then the choice here is available to a speaker throughout his/her life, so that roots can be ambiguously nominal or verbal.

The label V bears some feature that it can share with an object (or further Self Merge leads to an aspectual category, as in Ramchand (2008), that bears such a feature), allowing Merge of *them* followed by further Self Merge which is labeled v, which allows Merge of the subject, as is standard. Further Self Merge builds T, Internal Merge raises the subject, and further Self Merge builds matrix C. So much is fairly standard, with Self Merge and Labelling taking the place of external merge of heads. The final structure is a `telescoped' structure, similar to Brody's:

(15)

C

T

The teachers v

< The teachers > V

them ask

Following on from the discussion above, English effectively allows compilation of much of this process as its s-lexicon contains the sequence (V-v-T-C). This sequence captures

many properties of English verbs, such as the fact that aspectual properties of the verb are spelled out as part of the verb itself (that is there is no progressive inflection on an English verb); that the verb on modern English does not raise to declarative C (or the s-lexicon would contain (V-v-T-Fin-C, as it does in V2 languages), etc.

With this in place, let us turn in more detail to how these structures are spelled out. In Brody's system, a core principle is Mirror, which states that:

(16) The syntactic relation "X complement of Y" is identical to an inverse-order morphological relation "X specifier of Y" (Brody 2000, 42)

Brody's system gives us an agglutinative interpretation of the morphology, so that the morphological structure associated with (15) is:

(17)

0

s

0

0

ask

The second part of Brody's system is a diacritic mark that specifies at which point of the extended projection (what Brody calls the extended word) the morphological form is pronounced. If v bears this mark, then the word *asks* will be spelled out 'in' v: that is, it will be spelled out after the (trace of the) subject and before the object. Brody takes variation in the height of verbs to be given by this diacrtitic.

The system I have just proposed can make use of exactly the same intuition: the order of the morphological units is just given as a reverse order of the complement line which

derives from Self Merge, and the height of the verb can be given by a diacritic mark on the relevant label in the projection given by Self Merge of the verb: that is, a second order property related to spellout is attached to a sub-part of an element in the relevant s-lexical item.

However, as is well known, not all morphology is agglutinative; we also need to be able to handle fusional/suppletive morphology. I suggest that we do this by taking word-level vocabulary items to `span' the functional structures in the s-lexicon in a way akin to that proposed by Williams (2003). For example, the vocabularization of [be V v T[past] C] is just listed as *was*, capturing the suppletion. Note however, that in addition to this vocabularization, in the absence of any further stipulation, we also have to have [be V v T[past] C] vocabularized as be-0-0-ed-0, since each of these labels has a vocabularization independently of the others. Standardly, one would appeal to a blocking principle (Kiparsky 1973) to rule out *beed* as the past tense of *be*. However, given the framework outlined above for dealing with variation, blocking cannot be a principle of grammar, it must rather be a fact about the routinization of structures that is emergent from the learning process, so both *was* and *beed* have to coexist as possible vocabularizations of this syntactic structure. In terms of the Combinatorial Variability model, they are both in the Pool of Variants, however *beed* is so low in that Pool of Variants (that is, its input probablility is miniscule compared to *was*) that although it matches the category label sequence, it is never (or very rarely) going to be put to use.

During development the child is using all of the syntactic resources given to her by UG to build up syntactic structures and to linearize them into morphological chunks. However, as the child's development progresses, variation emerges as a result of the tension between the child's generated structures and the evidence from the input. This variation then reduces as the child begins to attempt to match the frequency of her caregivers' variants (Smith et al 2007). At some point, the accessibility of the non-routinized version is very low, and the variation vanishes from the child's production. Blocking is then simply a side effect of this process: the routinized version is associated with such a high probability that the regular version is judged as unacceptable.

Given this, a non suppletive structure like *asks* will also exist in (at least) two forms, a routinized version (or perhaps more than one of these, see below) and the version which is the algorithmic output of the linearization of the syntax (the equivalent of *beed*).

In situations where the child's input is variable, more than one vocabularization of a single s-lexical sequence is possible: one of these will be the output of the morphological linearization of the structure on a label by label basis, but the child's input may consist of only routinized suppletive forms (for example, *was* and *were*), and so the child will have these forms matched to sequences of syntactic categories and keep them both in the pool of variants.

With this framework in mind, consider again the variable agreement in Buckie. Maintaining the approach sketched above, we have two vocabularizations for the root *ask* in the context where it has a plural NP subject: *ask* and *asks*, dependent on the featural specification of the subject (and hence, given spec head agreement, of T). The latter of these itself has both a routinized and a generated source. The more frequent the verb *ask* is in the input, the more routinized its forms are. It is always possible, for infrequent verbs, to generate the *–s* form, (as indeed it is for the *–0* form) but there may be no routinized form. In the absence of such a routinized form, or if such a form is barely routinized, then there will only be one way to get to the *–s* form: linearize the syntax into an appropriate morphology. On the other hand, if the verb is frequent, then the routinized form will be accessible, and, moreover, it will always also be possible to linearize the syntax if the verb is regular. In this way, the system captures the generalization that the variability between the forms is more robust for frequent verbs than for infrequent ones.

The approach also gives us a handle on the problem of internal factors. Indeed it begins to give shape to a possible theory of internal factors. Note that in the structures generated by the theory, C is actually part of the possible vocabularization of the verb. Although the verb is pronounced 'in' v, C is part of the projection line, and hence a vocabularization that results in a verbal *–s* form may span C, and hence the relative/non-relative specification of C will be relevant to that vocabularization.

Thinking about the higher proportion of *–s* in relatives in Buckie, we can see this as combinatorial variability applied not to an X0 feature bundle, as in the theory presented in Adger (2006), but rather to a sequence of labels, where certain vocabulary items span relative C, while others don't. Putting aside the featural details covered in Adger (2006) and Adger and Smith (2010), we would have something like the following:

(18a) [V v T C[rel]] < -- > -s

(b)[V v T] < -- > -s

(c) [V v T] < -- > -0

Now in a non-relative, only (b) and (c) match the syntactic context, predicting a rough equality in the frequency of each variant. However, in a relative clause, all of (18a-c) are available for insertion, which predicts a higher frequency of the surface form *–s*. Of course, this is a toy example, as numerous other factors will be relevant (the featural content of the subject and hence T, possible aspectual information contributed by v (see, e.g. Poplack and Tagliamonte 1989) or an Asp projection, the frequency of the verb root itself, etc). However, even in this toy case we can see that not any global featural fact about the clause will be relevant, rather just those that are encoded in the projection line of the verb. We predict, then, that this kind of agreement variation cannot be affected by, for example, the number of the subject of a clause embedded by the agreeing verb, or the subject of a sentential subject of the verb.

4. Language Change and Variability

Let me now turn to some broader issues connected to questions of acquisition and change. In terms of acquisition, of course, the approach outlined here allows us to connect fairly well with some of the findings some of the work that's been done by the Construction Grammarians and their predecessors (e.g. Bybee 2006, and for work on variable agreement systems see, especially, Pietsch 2005), who have pointed out that there are frequency effects in the process of development that have an impact on frequencies in corpora and on ease of processing. However, we do this while keeping the notion of frequency or probability firmly outside the grammar: it is relevant only for

questions of lexical access during production. The grammar generates a restricted range of variants depending on the syntax (the Pool of Variants) and then some function will apply to the Pool of Variants to select a single one in any occasion of use. The structure of the Pool of Variants, how homonyms are distributed within it and whether there are both generated and routinized forms has a direct impact on the frequency distribution of the surface forms. In addition, the Pool of Variants may be externally structured by the processing systems that effect lexical access: some items are generally highly routinized and easily accessible, so that when they appear within the Pool of Variants this property itself has an effect on whether that item is chosen.

On the question of language change there are two interesting phenomena I want to mention here. The first is related to the loss of variability and the emergence of a possible parametric setting, while the second is related to the emergence of grammatical structuring in previously unstructured variation.

In Buckie there is a change happening in variable agreement. Older speakers use more *–s* forms with plural NP subjects than younger speakers:

(19)

Distribution verbal s- by age

old		middle		young	
%	N	%	N	%	N
73	59	**24**	136	**32**	185

There's an interesting dip in the middle aged speakers which appears in many of the variables which have been investigated for this corpus (see Smith 2000). Smith argues convincingly that this is due to a particular sociolinguistic factor connected to standardization for middle aged females which depresses the numbers. We will concentrate here just on the older vs the younger generation.

Furthermore, this change is affecting auxiliaries more than it is affecting main verbs:

(20)

Table X: Distribution verbal s- by age and verb type						
	old		middle		young	
	%	N	%	N	%	N
aux	**77**	13	**43**	37	**36**	67
be cop	**68**	19	**16**	55	**48**	63
other	**72**	25	**19**	42	**11**	54

Given the system outlined here, we can attribute this change to the frequency difference alluded to above (and in fact to the further effect of the suppletive forms of the auxiliaries, which implies that they must span the sequence of labels). Given that children track the frequencies of their caregivers, the existence of change in the frequencies of variables seems odd. However, at some point in development, children alter their frequency of use of particular variables. This is possibly connected to changes in schooling and hence to the alteration of how the developing child fixes its linguistic affiliation, changing from its family to its peer group, or at least adding to its linguistic repertoire. Assume that children alter this frequency uniformly, depressing their use of a particular variant by a stable amount. In such a case, the more frequent verbs will maintain a higher proportion of use of the disfavoured variant. This is especially true for the memorized or routinized forms. However, the grammatical system is still the same: the grammar still produces plural NP subjects with *–s* and we would predict that novel verbs, for example, would be grammatical in both forms. Low frequency verbs which are variable would arise just via application of the grammatical principles involving the linearization and vocabularization of category labels. However, given the frequency of auxiliaries, we then see a bifurcation between auxiliaries and main verbs that depends precisely on the fact that auxiliaries are in almost every sentence, whereas the different

main verbs are much rarer. Now, when these speakers bring up the next generation, after iteration of the process, the variable agreement, depending on its sociolinguistic status, may become yet more disfavoured until at some point the bifurcation between auxiliary and non-auxiliary verbs becomes so sharp that acquirers attribute what was a variable distinction to a categorical property of grammar (for example the auxiliary/main split). This would then correspond to the emergence of a parametric distinction between varieties that develops from the loss of variation interacting with frequency effects.

The other side of language change is explored in Cheshire, Adger and Fox (2010). In that paper we examine the appearance of a new condition on variation in the speech of a group of London teenagers. The study shows that, for both the teenagers and the older generation, both *who* and *that* can be used as subject relativizers with animate antecedents of the relative clause, just as in Standard English:

(21) apparently a chav is like someone that wears like big gold chains

(22) I'm the only one who's gone to college

As in standard English, *who* does not appear with inanimates. However, although the frequency of *who* as a subject relativizer with animates is the same in each generation, there are very different conditions on the use of *who*. The younger generation are statistically more likely to use *who* than *that* when the antecedent of the relative is an ongoing topic of the discourse. No such effect is found for the older generation, so what we see is the emergence of a structuring of the variation.

We capture this behaviour within the kind of system developed here by specifying the relativizers as follows:

(23) a. [C, relativiser:+] <--> *that*

b. [C, relativiser:+, animate:+] <--> *who*

c. [C, relativiser:+, animate:+, topic:+] <--> *who*

Any of these feature bundles can occur with an animate topical antecedent. However, when the antecedent is non-topical the vocabularizations in (23) are not all possible: non-

topical antecedents can only appear with (a) (pronounced as *that*) and (b) (pronounced as *who*). Schematically, we have the following situation:

(24) a. NP[animate:+, topic:+] (a) *that*, (b) *who*, (c) *who*

b. NP[animate:+, topic:-] (a) *that*, (b) *who*

This captures the higher proportion of relativizers pronounced as *who* that occur when the antecedent is topical: since both (b) and (c) can be used with a topic, but only (b) with a non-topic, topics are expected to occur with *who* more often than with *that*, while there is no such expectation for non-topics.

Cheshire, Adger and Fox argue that the availability of the topic feature in the specification of the relative complementizer (we take *who* just to be a realization of C) arises because of the multilingual environment the speakers develop their linguistic competence within, an environment which includes many languages which have grammatical marking for topicality and other information packaging features (see the paper for discussion). However, what is crucial for the discussion here is that the co-opting of this feature creates a differentiation in the Pool of Variants available to speakers in topical vs non-topical syntactic environments, and hence a difference in the frequencies with which the variants are produced in the different environments. That imbalance may then lead to a change in the grammatical system, in this case actuated by the multilingual environment rather than the attenuation of grammatical contexts due to sociolinguistc pressure to differentiate generational usage patterns.

5. Conclusion

To conclude, I think what is interesting about this system is that it allows us to maintain a strictly modular grammatical architecture, which keeps our theory of syntax fairly pristine, both in terms of the mechanisms and operations involved, and in terms of the theoretical vocabulary over which they operate. Moreover it allows us as theoretical linguists to begin to address questions of individual variability, frequency and internal factor structuring of variability that have not been addressed in generative grammar for many years. At the same time, this system connects well with current sociolinguistic

work on language variability and change, and provides a viable grammatical architecture for sociolinguists interested in linking their work with linguistic theory.

Thank you.

6. References

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.

Adger, David and Smith, Jennifer (2010) 'Variation in agreement: A lexical feature-based approach', *Lingua* 120. 1109–1134

Judy B. Bernstein and Raffaella Zanuttini (2010) 'What is verbal -s in English?' Ms.

Borer, Hagit (2005). *In Name Only: Structuring Sense*. Oxford: Oxford University Press.

Brody, Michael (2000) 'Mirror theory: syntactic representation in perfect syntax'. *Linguistic Inquiry* 31:29-56.

Bybee, Joan (2006) 'From usage to grammar: The mind's response to repetition' *Language*. 82.4, 711-733

Cheshire, Jenny, Adger, David and Fox, Sue (2010) 'Relative *who* and the actuation problem', Ms Queen Mary University of London

Chomsky, Noam (2004) 'Beyond Explanatory Adequacy' in Adriana Belletti, ed, *Structures and Beyond*, OUP, New York.

Chomsky, Noam (2007) 'Approaching UG from below'. In Hans-Martin Gaertner Uli Sauerland, eds., *Interfaces + Recursion = Language?* , Berlin and New York:

Mouton de Gruyter.

Collins, Chris (2002) 'Eliminating labels'. In *Derivation and Explanation in the Minimalist Program*, 42-64, Blackwell Publishing.

Collins, Chris and Stabler, Edward (2009) 'A formalization of minimalist syntax', ms

NYU and UCLA.

Ellegard, Alvar (1953) *The Auxiliary 'Do': The Establishment and Regulation of Its Use in English*. Almqvist & Wiksell, Stockholm.

Grimshaw, Jane (1991). 'Extended projections', ms, Brandeis University.

Guimarães, M. (2000) 'In Defense of Vacuous Projections in Bare Phrase Structure,' in

M. Guimarães, L. Meroni, C. Rodrigues & I. San Martin (eds.) *University of*

Maryland Working Papers in Linguistics, 9, 90-115.

Henry, Alison (1995) *Belfast English and standard English: Dialectal variation and parameter setting*. New York and Oxford: Oxford University Press.

Kayne, Richard (2008) 'Antisymmetry and the lexicon', ms NYU.

Kiparsky, Paul (1973) 'Elsewhere in Phonology', S. Anderson and P.

Kiparsky, eds, *A festschrift for Morris Halle*, 93-106. New York: Holt, Rinehard and Winston.

Kroch, A. (1994) 'Morphosyntactic variation'. In Beals, K., J. Denton, R. Knippen, L. Melnar, H. Suzuki and E. Zeinfeld, eds., *Papers from the 30th regional meeting of the Chicago Linguistic Society, volume 2: the parasession on variation in linguistic theory*. Chicago: Chicago Linguistics Society, 180-201.

Labov, W (1969), 'Contraction, deletion, and inherent variability of the English Copula', *Language* 45. 715-762.

Li, P., Dunham, Y. & Carey, S (in press). Of Substance: The Nature of Language Effects on Entity Construal. *Cognitive Psychology*.

Nevins, A., Parrott, J.K. (2010) 'Variable rules meet impoverishment theory'. *Lingua* 120, 1135–1159.

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

Poplack, Shana, & Tagliamonte, Sali A. (1989) 'There's no tense like the present: Verbal

-s inflection in early Black English'. *Language Variation and Change* 1:47–84.

Ramchand, Gillian (2008) *Verb Meaning and the Lexicon: A First Phase Syntax*.

Cambridge: Cambridge University Press.

Smith, Jennifer (2000) *Synchrony and diachrony in the evolution of English: evidence from Scotland*. Ph.D. Thesis, University of York.

Smith, Jennifer, Durham, Mercedes and Fortune Liane (2007) '"Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect'. *Language Variation and Change,* 19.63–99.

Svenonius, Peter (2008) 'The position of adjectives and other phrasal modifiers in the decomposition of DP' in Louise McNally and Chris Kennedy (eds) *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, Oxford University Press, 16-42.

Tortora, Christina, and Marcel den Dikken (2010) 'Subject agreement variation. Support for the configurational approach'. *Lingua* 1089–1108.

Williams, Edwin (2003) *Representation Theory*, MIT Press, Cambridge, Mass.

Xu, F. (2010) Count nouns, sortal concepts, and the nature of early words. J. Pelletier (ed.), *Kinds, things, and stuff: New Directions in Cognitive Science , Vol. 13 (pp. 191-206)*. Oxford University Press.

Zamparelli, Roberto (2000) *Layers in the Determiner Phrase* Garland, New York.