

Syntactic Theory and Linguistic Variation

David Adger

Chapter 1

Models of language variability

1.1 Intro

- outline 2 distinct approaches to studying language and the implications that these different concepts have had on the investigative methodology
- outline how variability in syntax fits into each of these different conceptions, and what problems it brings up.
- look in some more detail at how people have tried to incorporate syntactic variability into theoretical systems.

1.2 Variability in Language

1.2.1 I-language and E-language

Chomsky's distinction between the study of an external set of linguistic forms and the study of the internal system that is (in part) responsible for these.

a language as a collection of actions, or utterances, or linguistic forms (words, sentences) paired with meanings, or as a system of linguistic forms or events. Let us refer to such technical concepts as instances of externalized language (E-language), in the sense that the construct is understood independently of the properties of the mind/brain. From a point of view such as this, a grammar is a collection of descriptive statements concerning the E-language. The E-language is now understood to be the real object of study. Grammar is a derivative notion: the linguist is free to select the grammar one way or another as long as it correctly identifies the E-language. ?) pp19-20

A rather different approach was taken, for example, by Otto Jespersen, who held that there is some notion of structure in the mind of the speaker which is definite enough to guide him (sic) in framing sentences of his own, . Let us refer to this notion of structure as an internalized language (I-language). The I-language, then, is some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker-hearer. The

statements of a grammar are statements of the theory of mind about the I-language, hence statements about structures of the brain formulated at a certain level of abstraction from mechanisms. ?) pp21-23

- 'a mental organ'
- part of individual psychology,
- apparently species specific in large part;
- intensional

An individual grammar S_T is seen as the outcome of a function operating on the initial state (S_0) and the primary linguistic data (E)

$$(1) \quad L(S_0, E) \rightarrow S_T$$

Because I-language is part of the natural world, in the same way as livers or daffodils are, we can discover things about its properties experimentally, to a certain extent. Because I-language is something which is not directly observable itself (like quarks or neutrinos), a methodology of direct observation is not possible.

intuition based methodology: intuitions about acceptability of sentences provide indirect evidence about the structure of I-language. Note:

- intuitions are not a priori about grammaticality, that is something to be established by theory construction
- an intuition, in a way, is a matter of performance, and so may be affected by performance related factors (judgment fatigue etc)
- are you really getting at your I-language when you make an intuition judgment, or is there an overlay of prescriptive norms imposed by a standard? Especially problematic when working on certain phenomena (such as dialectal phenomena) with certain informants (such as educated ones).
- if the last point is true, is the imposed standard language responsible for reducing variability in judgments, hence leading theorists to build theories which have little room for intrapersonal variation.
- if the last point is true, is there more variability than there appears to be?

However, in the same way you can find things out about deep properties of matter by observing neutron stars and black holes, it should also be possible to find out some things about I-language by observation of its effects (e.g. corpora, observation of speech patterns, etc.). In short, you take evidence from wherever you can in order to develop understanding.

1.2.2 S-language

What Labov calls the 'materialistic conception' of language:

The materialistic conception is that language is a property of the speech community, an instrument of social communication that evolves gradually and continuously throughout human history, in response to a variety of human needs and activities. ?).

The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms. These norms may be observed in overt types of evaluative behavior, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage. Labov (1972:120-1)

So a speech community is really about attitudes and behaviour.

The materialist view is that 'competence' can only be understood through the study of 'performance', and that this dichotomy involves an infinite regress: if there are separate rules of performance to be analyzed, then they must also comprise a 'competence', and then new rules of 'performance' to use them, and so on.

Similar to some interpretations of the Saussurean concept of *Langue*: an abstraction over the linguistic behaviour of conspecifics. An individual who is a member of the community is attuned to this abstraction and can 'use' it.

- Social object
- massively variable
- Observational method has primacy: what do speakers do? Gather a corpus and study it
- Observer's paradox vs ethical methodologies
- corpus size - needs to be very large for (morpho)syntactic variation, unlike phonological.
- assuming that speakers actually do make mistakes, we are going to find performance errors. How do we work out whether these are 'errors' or not?
- further, there will just be unpluggable gaps - things that are rare or non-existent in the corpus.

The S-language camp generally criticize the generative camp for weak methodology in using judgments - see especially ?. However, everyone working on syntax or semantics uses judgments. It's just not possible to do away with them when faced with how meaning impacts on the analysis of data.

For generativists, the judgments themselves are the data, not the sentences (just try removing the starts from a syntax paper and see how far you can get with the analysis! Bob Borsley makes the analogy with going to the opticians, and being asked to give judgments about how well your eyes are focussing when covered by various lenses. For the optician, the data used to eventually create your prescription are the judgments. ?).

It's probably true that generativists should be clearer about our methods and materials in order to ensure better reliability of our data. It's also, of course, true that there is a huge mass of unproblematic, replicable and reliable data that we don't have good explanations for. However, it is also the case that theoretical choices might depend on the answers to fairly subtle judgments, and these are usually not checked for reliability, replicability etc.

The following are rarely dealt with in syntax papers. Perhaps syntacticians do this as a matter of course (and certainly it's close to necessary when working on a language which you are not a native speaker of), but it would be informative to say so.

- re-testing,
- controlling carefully for context,
- judgment fatigue
- distractors
- control for impact of prescription

(but see ?; ?; ?)

1.3 What varies?

1. In phonology and morpho-phonology, the traditional approach is to identify variants of a single overarching category, and to establish complementary distribution. e.g. The different allomorphs of the English plural morpheme /s/. These variants are determined categorically by context (the preceding segment, in this case)
2. However, take the -ing morpheme that forms the English progressive participle. It has two variants: /in/ and /iŋ/. Both are possible in the speech of single speakers, and across speech communities. A typical structuralist position would just be to say that these are in free variation.
3. This now gives us an example of what Labov calls the sociolinguistic variable (see, e.g. ?)? It's a case where a speaker has a choice of forms which convey the same meaning. The variationist idea is that, when a speaker has such a choice, it may be the case that speakers attach a social value to the variants (there may be no such social value either).
4. Such a definition of a variable is unproblematic in phonology, where the 'meaning' of the item concerned is fairly straightforwardly identifiable (not though, by just looking at a corpus!). The idea then is that there is an underlying form associated with the meaning which has a number of surface realisations (note the similarity to the architecture of the standard theory of TG).

I'll call this the Allomorphic Theory of Variability (ATV). It essentially says that variability works like this:

- (2) $\alpha \rightarrow a$ in context C1, A in context C2, \aleph in context C3

Of course, the effect of contexts is not deterministic here, but rather probabilistic, and the contexts can range from phonology to sociology. See section 1.3.2.

However, this approach is potentially problematic in syntax. Do topicalization structures and non-topicalized structures form a variable in this sense? The underlying form in the 1960s standard theory is identical, but in later analyses (e.g. ?) there are quite different underlying forms for such examples. Their truth conditional meaning will be the same, although their discourse embedding conditions will be different.

- (3) a. I like beans
b. Beans I like

Within the variationist literature, people have taken different approaches to this. Some have argued that this kind of variation is not a variable, in the standard Labovian sense, since notions of social class etc are not relevant, rather pragmatic factors are. Others have claimed that these pragmatic factors should always be incorporated into the notion of variable. (see, e.g. ?, ?)

What about particle and non particle shifted structures:

- (4) a. He took his coat off
b. He took off his coat.

Note that there are some categorical effects (examples from ?)

- (5) a. Lock it up.
b. *Lock up it.
- (6) a. *Lock all the doors on the second and third floors that lead into rooms with expensive equipment in them up.
b. Lock up all the doors on the second and third floors that lead into rooms with expensive equipment in them.

Svenonius gives the following context (page 50):

somebody walks into the room where you are sitting, engaged in the activity of balloon inflation, and asks you what you're doing.

- (12) a. -? I'm blowing balloons up. b. I'm blowing up balloons.

Here, Svenonius reports a preference for the (b) example in 8/12 speakers (the dis-preference for the (a) example is marked by a dash before it). Now contrast with the following (page 51):

the person who walks into the room knows that there is going to be a party, that you are making preparations, and furthermore that there are balloons. He asks the same question, ?What are you doing??

- (13) a. I'm blowing the balloons up. b. I'm blowing up the balloons.

Here, Svenonius reports a no preference either way. He proposes that this is because the syntax generates both structures, but that there is an interaction between the position of sentence stress and the fact that novel indefinites are best with stress, rather than without. Since sentence stress in English is at the right edge of the sentence, (12b) is an outcome which is more in keeping with the extra grammatical constraints imposed by the interfaces with syntax. In the situation with the definites, the whole VP is in focus, and this can be marked by stressing the right edge. Svenonius backs this idea up by contrasting this case with epithetic definites, which prefer to be destressed:

For example, imagine a situation in which there are two girls, Turid and Ingrid, and you want to know about their transportation. You ask, 'How are Turid and Ingrid going to get here??' and I respond:

(16) a. I'll pick the girls up. b. ? – I'll pick up the girls.

- (7) a. The cake which I made
 b. The cake that I made
 c. The cake I made

Some speakers make a difference between (a) and (b), so that (a) is only non-restrictive (e.g. the Word grammar checker!). This is (emphatically) not true for at least one speaker of English (me).

Both of these are part of Standard English. Compare with the following, where there doesn't appear to be an information structural effect:

- (8) a. There are three problems
 b. There's three problems
 (9) a. Where are my specs?
 b. Where's my specs?
 (10) a. I'd better go
 b. I better go

Again, these are (sometimes stigmatized) cases, but widely used.

Dialectal forms like the following also show variation:

- (11) a. I never went to school today
 b. I didnae go to school today
 (12) a. Did you never go to school today?
 b. Did you no go to school today?

No truth conditional or information structural difference.

Singular Concord ?

- (13) a. The eggs are cracked
 b. The eggs is cracked

Here we see variability in a rule of agreement.

Negative Concord vs negative polarity structures in Buckie (Smith 2001)

- (14)
- a. I na ken nane of that, nor I na ken nane of that. (1520.9:a)
I don't know any of those (people), nor do I know any of those.
 - b. Now, ee dinna ken ony o' that. (2131.17:a)
Now, you don't know any of those (people).
 - c. It wisna really gan naewye. (175.21:t)
It wasn't really going anywhere.
 - d. We didna get onywye. (408.4:t)
We didn't get anywhere.
 - e. I wisna sick or nothin, ye ken. (376.50:1)
I wasn't sick or anything, you know.
 - f. Never heard of a woman director or onythyn. (128.18:1)
Id never heard of a woman director or anything.

Auxiliary *have*:

- (15)
- a. I haven't Peter's brains
 - b. I don't have Peter's brains
 - c. I haven't got Peter's brains

In general, these syntactic variants appear to be reducible to

- whether Agree applies in the same way
- whether Move applies or not
- variation in lexical items chosen

1.3.1 Modelling variaation from an S-language perspective

Because of the nature of the conception of the language as involving a community grammar, within a speech community, variation is ubiquitous. Moreover, not only are individuals are 'attuned' to the grammar of their communities in terms of simple grammaticality, they are also attuned to it in terms of the patterns in the variation that is found. The patterns are argued to involve external and internal factors.

- external: gender; class; age; identity ...
- internal: featural constitution; structural position

A famous example is copula absence in AAVE:

- (16)
- a. She is my piano teacher.
 - b. Shes my piano teacher.
 - c. She my piano teacher.

?) originally argued that this was a phonological rule of consonant deletion, linking the possibility of contraction in Standard English with that of deletion in AAVE. But ?) points out cases like

- (17)
- a. That's the man they say in love.

- b. *That's the man they say's in love.
- (18) a. How old do you think his baby.
b. %How old do you think's his baby.

When copular absence is investigated wrt syntactic factors, like the category of the following complement, these have a strong effect. The numbers here are factor weights, which essentially say how important a factor this is in triggering copula absence. A weight of .5 means that there is no effect (from ?):

Table 1.1: Copular absence across varieties

	Form	Place	Source	NP	Loc	Adj	V+ing	gon
1	is	NYC Thunderbirds	Labov 1969	.2	.36	.48	.66	.88
2	is	NYC Jets	Labov 1969	.32	.52	.36	.74	.93
3	is	NYC Cobras	Baugh 1979	.14	.31	.72	.59	.78
4	is+are	Detroit WC	Wolfram 1969	37%	44%	47%	50%	79%
5	is	LA	Baugh 1979	.32	.29	.56	.66	.69
6	are	LA	Baugh 1979	.25	.69	.35	.62	.64
7	is+are	Texas kids	Bailey and Maynor 1987	.12	.19	.25	.41	.89
8	is+are	Texas adults	Bailey and Maynor 1987	.09	.15	.14	.73	.68
9	is+are	East Palo Alto	Rickford et al. 1991	.29	.42	.47	.66	.77

Note the effect of the internal factors here is fairly constant across varieties, with copular absence being disfavoured with NP complements and favoured with verbal complements. One S-language approach is to tie this down to history.

An alternative might be to tie this down to the effect of discourse function on frequency.

? take the case of Lummi (Straits Salish), where the person of the subject argument cannot be lower than the person of the non-subject. If this would happen in the active, the verb is passivized; if it would happen in the passive, the verb is obligatorily active:

- (19) *₋ The man knows me/you
- (20) xči- t- ŋ =sən/ =sx^w ə cə swəyʔqəy
 know- TR- PASS =1.SG.NOM/ =2.SG.NOM by the man
 I am/you are known by the man
- (21) xči- t- s cə swəyʔqəy cə swiʔqoʔəl
 know- TR- 3.TR.SUBJ the man the boy
 The man knows the boy
- (22) xči- t- ŋ cə swiʔqoʔələ cə swəyʔqəy
 know- TR- PASS the man the boy
 The boy is known by the man
- (23) xči- t =sən/ =sx^w cə swəyʔqəy
 know- TR =1.SG.NOM/ =2.SG.NOM the man
 I/you know the man

(24) *₋ The man is known by me/you

Bresnan et al. examine a parsed corpus of English (the SWITCHBOARD corpus) and found the same kind of effect, but non-categorially:

Table 1.2: English person/role by voice

action	# Active	# Passive	% Active	% Passive
1,2 → 1,2	179	0	100	0
1,2 → 3	6246	0	100	0
3 → 3	3110	39	98.8	1.2
3 → 1,2	472	14	97.1	2.9

Although passives are rare in conversational English, Bresnan et al report that the pattern here is significant ($\chi^2 = 115.8$, $p < 0.001$).

Why should this be? There is presumably a discourse functional explanation for the different frequencies found in the English data: first and second person pronouns are the crucial participants in a discourse, and so they are best as subjects. Passives demote the subject, and so first and second persons are not good in the by phrase in a passive.

The explanation Bresnan et al. give is couched in Stochastic optimality Theory. They propose a constraint which penalizes structures where there is a third person subject and a first/second (i.e. local) object. This constraint is always wins in Lummi, but it is present in English as a depressing factor. The way the grammar works is to specify constraints with values on the continuous scale of real numbers. A random value drawn from a normal distribution is added to this at evaluation, and this new scale is used for the actual evaluation. So the grammar ends up constraining but underdetermining the output. The randomness is added to take account of various non-syntactic factors.

An alternative would be to say that the grammar of English allows both outputs, but that the choice of which output depends on how well they suit the interface constraints that are imposed on them. So say that there is indeed a 'person hierarchy', but this is a discourse preference. Then we expect to see this preference having a stochastic effect. In some languages, the grammar has essentially co-opted the discourse principle (that is, the notions of speaker and hearer, and aboutness topic have been reified in the grammar as features which have to enter into syntactic relations; features such as [person] [addressee] [nominative] and relations such as [specifier of TP], [complement of V] etc).

1.3.2 I-language

Theories of I-language

The Rule-format Approach

UG provides a certain format for rules, which the child uses together with the primary linguistic data to construct the actual rules of her/his language. e.g. phrase structure rules, rewrite rules in phonology, transformational rules. (e.g ?, ?)

The Principles and Parameters Approach: First Stab—filtering

UG provides a restricted range of very general rules which operate optionally within a modular architecture, and a set of constraints on the application of these rules (e.g. subjacency) or on their output (e.g. the ECP). Output constraints are very much part of narrow syntax, and need have no external motivation. (e.g. ?, ?)

The Principles and Parameters Approach: Second Stab—WYSIWYG

UG provides a restricted range of very general rules which are, in most cases, triggered by properties of lexical items (especially functional categories, but also perhaps selection). Output constraints (Interface Constraints) are very much to be motivated externally and they connect elements of the vocabulary of syntax to elements of semantics/prosody etc. (e.g. ?)

Incorporating variation into formal rule systems

Labov's original approach: **variable rules**—attach probabilities to rules. Example from ?):

(25) que → 0 / ⟨[+sib], [+cns, -sib], [-cns]⟩ ## ___ ## ⟨[+sib], [+cns, -sib], [-cns]⟩

Table 1.3: variable rule

Preceding Environment	[+sib] __	[+cns, -sib]__	[-cns] _
Effect	1	.85	.37
Following Environment	[+sib] __	[+cns, -sib]__	[-cns] _
Effect	1	.50	.10
Occupational Class	Workers	Professionals	
Effect	1	.35	
Sex	Women	Men	
Effect	1	1	

Within a Filtering version of P & P, optional movements/operations are unproblematic, since they are the elsewhere case. It is more difficult to see how probabilities might be attached to the options, as they are in the variable rules paradigm.

Minimalist Perplexions

(26) Move is triggered

(27) optionality of triggering features is assumed to have an effect on the output

? suggests that EPP features on a head are only available when such a feature contributes to some outcome at the semantic interface, e.g. scopal, or discourse related effects. The EPP feature is the formal mechanism that implements the function of providing a new interpretation.

(the idea is to capture the following generalization: If movement to a position is optional, then it is associated with an interpretative effect; if it is obligatory, then it is not. See ?, ?); ?); ?))

Why then do we have apparent optional movements with no interpretative effects?

Competing Grammars

? argues, on the basis of a range of historical texts, that distinct grammars may coexist, but that they will contain 'doublets'. Doublets are words which are different forms which are functionally equivalent. For example:

(28) warm; warmth; ??warmness

If a system contains doublets that have different grammatical effects (for example, triggering movement) but the same semantic function, then such a system is unstable, and diachronic change will take place.

(29) a. How great and greuous tribulations suffered the Holy Appostels?
b. How great tribulations did the Holy Apostles suffer?

(30) a. spoile him of his riches by sondrie fraudes, whiche he perceiueth not.
b. which he does not perceive.

(31) a. Quene Ester looked never with swich an eye.
b. Queen Esther never looked with such an eye.

Table 1.4: Frequency of periphrastic do sentences by context.

dates	negative decl	negative Qs	affirm trans Qs	affirm intrans Qs	affirm whobj -Qs
	% do N	% do N	% do N	% do N	% do N
1400-1425	0 177	11.7 17	0 3	0 7	0 1
1426-1475	1.2 903	8.0 25	10.7 56	0 86	0 27
1476-1500	4.8 693	11.1 27	13.5 74	0 68	2.0 51
1501-1525	7.8 605	59.0 78	24.2 91	21.1 90	11.3 62
1526-1535	13.7 651	60.7 56	69.2 26	19.7 76	9.5 63
1536-1550	27.9 735	75.0 84	61.5 91	31.9 116	11.0 73
1551-1575	38.0 313	85.4 48	73.7 57	42.3 71	36.0 75

Kroch shows that the rate of change is the same for all contexts. This is the 'constant rate effect'

We take its [the CRE - dja] general validity to indicate that what changes in frequency in the course of time during a syntactic change is language users overall tendency to choose one abstract grammatical option over another in their language production; and it is this changing tendency that produces the changes in the individual surface contexts where usage frequencies can be measured. Note that the underlying option may be reflected in different surface contexts in different ways, due to interaction with other grammatical factors. Kroch 1994:3

Kroch models this in terms of competing grammars. A single grammar cannot contain two functional heads with the same semantic function but different morphological/spellout effects, since this is ruled out by blocking. So we have two grammars, which agree on everything except that one contains one head and the other contains the other. The grammars compete to analyse sentences, with the more successful grammar being rewarded by being more 'prominent'. This forces diachronic change.

Kroch's approach is one way to get variation at the intrapersonal level. An alternative is to extend the notion of parameter.

Parameters

Two views of Parameters: Macro and Micro

Macro-parameters—?); ?)

Parametric differences are stated at a high level and have an effect on the 'design' of a language:

The Morphological Visibility Condition

- (32) An XP is visible for θ -role assignment from a head Y only if it is coindexed with a morpheme in the word containing Y via (a) Agreement (gives non-configurationality) or (b) Movement (incorporation)

Gives **Polysynthesis Parameter**: yes: Mohawk, Nahuatl no: English, Chichewa

- (33) [_w Aff_i + Y] XP_i (Y assigns a θ -role to XP because XP is coindexed with the affix)
- (34) [_w X_i + Y] [XP_i <X_i>] (Y assigns a θ -role to XP because XP is coindexed with its head X, and X has incorporated into the word containing Y)

Note that this is global in nature, since it refers to any morpheme. Baker argues that it's not appropriate to encode this as properties of lexical items hence the globality of the parameter.

- brings together noun incorporation and rich agreement. Note that they could be separated, so that a language might be non-configurational and headmarking without being polysynthetic (Warlpiri; Navajo etc).
- predicts no verb-incorporation of transitive verbs, since the object of the lower verb will not be able to get a θ -role from that verb (unless the object incorporates first)

- (35) [[Cause + V] [_{VP} <V> NP]

- Predicts no infinitives without incorporation.
- Note also that in Baker's analysis theta roles are still assigned to NP arguments by virtue of their being coindexed with agreement morphemes (so pro-drop still has to be stated independently), rather than theta-roles being assigned within a word to agreement morphemes.

Lexico-parameters—?)

From this perspective, all variation is tied down to properties of functional heads. There are no large structural differences in languages that are not ultimately the result of the specific properties of functional elements. Various ideas have been put forward as to what these properties might be: uninterpretability, strength, propensity to incorporate etc. From ?), Chapter 9:

Table 1.5: A typology of binary parameter values

	tense on Aux	tense on <i>v</i>	EPP on T	Decl on T	[top] on C	wh on C
English	strong	weak	strong	weak	optional	strong
French	strong	strong	strong	weak	?	optional
Swedish	weak	weak	strong	strong	strong	strong
Scottish Gaelic	strong	strong	weak	weak	weak	strong
German	strong	strong	strong	strong	strong	strong
Japanese	strong	strong	strong	weak	weak	weak

This now, once again, raises a question of explanatory adequacy: what limits the variation here? The nature of the probe-goal relation in the syntax? Feature geometries which are themselves the result of Interface Constraints?

Microparametric variation within a dialect:

?):

- (36) %Read you that.
 (37) Go you away.
 (38) %Give you me quickly that paper.
 (39) %Throw us you your end there.

Two grammars - sociolinguistically distinct. The older generation accept all the examples, the middle aged and younger generation accept only unaccusatives.

However, all speakers who accept the inverted versions, also accept the non-inverted ones:

(40) You read that

1. The (phonetically null) imperative morpheme which occurs in the C position is optionally strong in Belfast English (triggering verb raising to C in imperatives) in one dialect.
2. The subject in Belfast English is not forced to raise to spec Agr; therefore in imperatives, which lack Tense, it is not forced to raise at all, and remains in situ in its D-structure position.

G1:

(41) Read you that: [CP read [TP [vP you ⟨read⟩ that]]]

G2:

(42) *Read you that: [CP [TP [vP ⟨you⟩ read that]]]

(43) Go you away: [CP [TP [vP go you] away]]

Many unaccusatives appear with particles in the data:

(44) Go you away.

(45) Sit you down.

A standard VARBRUL analysis might lead one to believe that the presence of a particle favours inversion. But note that no speaker has such a grammar. The presence of a particle is independent. Moreover, the direction of change appears to be towards the unaccusative-only grammar. Henry argues that this would not be evident if the data were 'pooled'

Various problems with this story: however, it argues that a single speech community (in some sense anyway) can consist of speakers with different grammars. Moreover, Henry suggests that there is variation within the grammars (e.g. in the 'optionally strong' statement). Further evidence:

- (46) a. There should have been lots of students working on their assignments
 b. There should have lots of students been working on their assignments
 c. There should lots of students have been working on their assignments

For Henry, the variation found within a dialect, or within a language should be of the same 'sort' as that found between languages. In a sense, micro-parametric variation is not about especially small parameters, it's just about parameters tout court. However, note that we don't have a Blocking Effect here: there's true optionality with no apparent change. Moreover, there appears to be no interpretive effect.

?)

In the case of core linguistics, there is a need to recognize that parameters exist only because variation is the norm within human languages. The type, range, and limits of parameters, along with our understanding of their relationship with the core principles of universal grammar, can surely be fully worked out only by exploring the limits of language variation in both time and space, to borrow the phrase of Labov 1994 (and this allows for more than one method of analysis of such variation).

There's not a great difference between these two approaches. Henry proposes multiple parametric options encoded as properties of lexical items, while Kroch says that whenever you have such a situation, there is competition because the Blocking Principle makes such doublets diachronically unstable. Kroch's view seems to be a little at odds with certain sociolinguistic findings. For example, (?) found the following range of effects in an investigation of negative concord vs negative polarity, which suggests that something more complicated than a simple change of frequency across the generations is going on:

Table 1.6: From Smith 2001

Old	90%	.80	87
Middle	39%	.19	72
Young	71%	.47	178

Moreover, for other variables in her corpus, Smith finds only a slight, or no effect of age, gender etc. In fact, what we seem to have in the Buckie data is a socially homogeneous speech community that still shows a lot of variation, but much of it is diachronically stable. This suggests that the functional category specifications responsible for the variants are not affected by the Blocking Principle.

1.4 Acquiring a variable grammar

Learning a language (?):

$$(47) \quad L(S_0, E) \rightarrow S_t$$

- GG has been interested in S_0 and S_t , with little role for L or E
- Sociolinguistics has been mainly interested in E

?) and ?) suggest a genetic algorithm approach:

$$(48) \quad L \text{ assigns probabilities to grammars on the basis of the grammars abilities to analyse elements of E. If G is successful, it is rewarded, otherwise it is penalised. (similar to Kroch's approach)}$$

What is a G?

$$(49) \quad \langle \text{SynOp, IntCon, Parameters} \rangle$$

$$(50) \quad \text{SynOp is invariant (Internal and External Merge)}$$

$$(51) \quad \text{IntCon is invariant (Interface constraints: e.g. The Functional Sequence, Binding, Predication etc)}$$

$$(52) \quad \text{Parameters is a finite set of parameters with values 0 or 1.}$$

Simplifying, we could say that each grammar is an n-tuple of parameters with on/off settings. This will give us a set of grammars that is available to the child at S_0 :

$$(53) \quad G = \langle F_1[0, 1], \dots, F_n[0, 1] \rangle, \text{ where each } F_i \text{ has the specification } [0] \text{ or } [1]$$

As each sentence is heard, the child attached probabilities to members of G. If a member of G can analyse the S, it is rewarded, and those which can't are punished. Eventually (by the time the critical period for language acquisition is over), we have just a single G, or a small set of Gs, which are capable of analysing the elements of E. In the latter case,

Table 1.7: From Alison Henry's Unpublished Work

	CHILD		CAREGIVERS	
	-agr	+agr	-agr	+agr
Stuart	3	33	2	17
Barbara	4	16	2	40
Conor	15	36	20	51
Michelle	0	25	2	22
Courtney	2	12	21	12
David	0	2	4	34
Rachel	0	4	3	57
Johnny	0	3	1	5
TOTAL	24	131	55	218
% -agr		15%		20%

we have multiple grammars, and, moreover, we predict that the child should be mirroring the proportions of the variants found in E. This seems to be the case:

I think we can also think of this just in terms of lexical items. When a child encounters S, she needs to construct an analysis using SynOp, IntCon and a set of lexical items. Some of these lexical items are functional, and may have syntactic effects. If there are variants in E, then for a single LF-analysis (distribution of interpretable features), there will be more than one functional lexical item that can lead to this analysis, since both will be rewarded, while others will be penalized.

(54) go you away

(55) you go away

(56) T[imperative, EPP] ; T[imperative]

Notice that lexical items are precisely the things that we have choices about. Those (sometimes tacit, usually at least unreflecting) choices are dependent on sociolinguistic facts about us as speakers, about our expectations of our hearers, and they may be constrained by questions of lexical access (priming effects, frequency of hearing the word). Moreover, since they have grammatical meaning and grammatical function rather than encyclopedic meaning, they act essentially as parameters—they are the basic elements of the child's syntactic/semantic analysis of the sentences she hears. Children match frequencies because they have attached probabilities to the variants.

For example, take Henry's analysis of singular concord:

(57) [_{TP} The eggs [_T is] cracked]

(58) [_{AgrP} The eggs [_{Agr} are] [_{TP} [_T t_{are}] cracked]

Under the interpretation of variability suggested here, a speaker has a (possibly tacit) choice to make about whether to use an Agr head or not. Since number features on Agr are uninterpretable, they vanish at LF, and only the tense features of T are left.

The final LFs of both sentences are, as far as the Interface Constraints are concerned, identical. We get the effect of the Labovian definition of a sociolinguistic variable in terms of interpretable and uninterpretable features! More of this anon.

Chapter 2

Variability in the Agreement System

2.1 The data

- (1) a. He says 'I thocht you were a diver or somethin.' (7:262.41)
He said I thought you were a diver or something.
b. 'Aye, I thocht you was a scuba diver.' (7:259.21)
Yes, I thought you were a scuba diver.
- (2) a. There was one nicht we were lyin at anchor. (g:875.32)
There was one night we were lying at anchor.
b. We played on at beach til we was tired, sailin boaties, bilin whelks(b:254.15)
We played on that beach until we were tired, sailing boats, boiling whelks.
- (3) a. Was you ones so bad, like? (l:262.7)
b. NEED EXAMPLE
- (4) a. They were aie sort o pickin on me, like. (j:504.3)
They were always sort of picking on me.
b. They were still like partying hard. (j:635.28)
They were still partying hard.
- (5) a. Buckie boats were a bonny grant. (g:1066.0)
Buckie boats were all nicely grained.
b. The mothers was roaring at ye comin in. (b:256.34)
The mothers were shouting at you to come in.
- (6) a. There were a puckle thatched houses like that. (c:335.17)
There were a couple of thatched houses like that.
b. Oh, there was a lot of coopers at time. (c:13.45)
There were a lot of barrel makers at that time.

2.2 Variable vs Categorical Distribution

There are patterns in the variability. ? showed that the person and number features of the subject were crucial (and that other factors such as whether the verb was negative or positive, whether the subject was collective or not, whether the verb is a copula or an auxiliary, were irrelevant to determining the form.

Table 2.1: distribution of *was* in *were*

pronoun	percentage	N
second singular	69	161
first plural	67	368
second plural	10	10
third plural	0	435
existential	90	162
NP plural	56	187

Crucially we see here a pattern of variability and categoricity.

2.3 Historical Explanation

Table 2.2: Survey of the pronominal forms *be* in Middle English (Forsström, 1948) .

	Kent	Saxon	Southwest	East Midlands	Northeast	West
1st, 3rd sg	wes	was	was	was	was	was
2nd sg	were	ware, were	were	wore	was, wore (ware)	was [north]
Plural	were(n)	ware(n), were(n)	were(n)	wore	wore(n), ware/n/	

Note that, in Northern varieties, we see *you*[singular] *was* most of the time and *you*[pl] *was* only once (admittedly with low numbers). Smith suggests that this is a retention from the historical record. Similarly, she suggests that the categorical zero rate for *they* contrasting with the high rates for NP[plural] is a retention of the Northern Subject Rule reported by Murray 1873. The problematic case for this approach is *we was*, which isn't found in the historical record.

2.4 Functional Explanation

Looking at other varieties of English we can see that there is some correlation across dialects (data from Smith). Plural existentials generally have a high rate of *was*, and *they* tends to have a lower rate than NP[plural]. However, overall, the varieties are quite different.

Table 2.3: Distribution of was in were by person across varieties

Variety	existential	you	we	NP plural	they
Alabama	100	91	98	100	93
Buckie	90	91	73	81	0
Guysborough	86	80	75	65	50
North Preston	64	71	43	52	37
Guysborough village	66	7	6	33	0
York	63	0	2	9	4
The Fens		68	69	26	53
Sydney	89	32	11	17	9
Devon	97	65	43	79	57

2.5 Formal explanation

2.5.1 The notion of Agreement

Within current minimalist syntactic theory agreement phenomena are dealt with by matching features. The requirement that a feature matches with something else is thought to be derivative of whether that feature is associated with a semantic interpretation or not.

$$(7) \quad X[F] \dots Y[uF] \rightarrow X[F] \dots Y[\cancel{uF}]$$

2.5.2 A 2 grammar approach

2.5.3 A single grammar approach

2.5.4 A remaining Problem

2.6 Finishing off yesterday's discussion

$$(8) \quad \text{We was thegither, we were all thegither}$$

$$(9) \quad \text{Pronoun}[F:a] \dots \text{Be}+\text{past}+[F:] \rightarrow \text{Pronoun}[F:a] \dots \text{Be}+\text{past}+[F:a]$$

In G1 we have a version of T that is lexically [num:]; in G2 we have [pers:].

Table 2.4: multiple grammars

G1	G2
[num:pl] → were	[pers:+] → was
[num:sg] → was	[pers:-] → were

The evidence is overwhelming that **she were* is not in the grammar, so the child restricts G2:

Table 2.5: restricting G2

G1	G2
[num:pl] → were	[pers:+] → was
[num:sg] → was	[pers:-] → *

Once the child has retreated to the reduced version of G2, she can use both grammars, correctly predicting:

- (10) a. I was; *I were
 b. you[sg] was
 c. she was; *she were
 d. we were; we was
 e. you[pl] were; you[pl] was
 f. they were; *they was

Both grammars together will analyse all the data, and will generate all the data. Except for *you[sg] were*: variety of possible solutions

Table 2.6: Extending G1

G1	G2
[num:pl] → were	[pers:+] → was
[num:sg] → was	[pers:-] → *
[num:sg, addressee] → were	

Correctly predicts the data.

Table 2.7: Extending G2

G1	G2
[num:pl] → were	[pers:+] → was
[num:sg] → was	[pers:-] → *
	[pers:+, addressee] → were

Assuming the Elsewhere principle, which tells us to use the most specific thing in our lexicon, this second version predicts no variability for *you[pl]*.

However, there's still a problem:

- (11) a. Buckie boats were a bonny graint. (g:1066.0)
 Buckie boats were all nicely grained.

- b. The mothers was roaring at ye comin in. (b:256.34)
 The mothers were shouting at you to come in.
- (12) Sa provoserer Salomes mannshunger fortsatt dagens publikum
 So provokes Salomes man-hunger still the.days audience
 'Then Salomes hunger for men still provokes todays audience.
- (13) Sa provoserer fortsatt Salomes mannshunger dagens publikum
 So provokes still Salomes man-hunger the.days audience
- (14) Sa provoserer den fortsatt dagens publikum
 So provokes Salomes man-hunger still the.days audience
 'Then it still provokes todays audience.
- (15) ??Sa provoserer fortsatt den dagens publikum (unless subj is focused)
 So provokes still it the.days audience

Alternative:

- (16) [_{AgrP} pronoun [Agr [_{TP} DP T [_{AgrP} DP [Agr [_{TP} ⟨ DP ⟩ T * [_{TP} pronoun T
- (17) a. Be+[Agr, num:pl, past] → were
 b. Be+[Agr, num:sing, past] → was
 c. Be+[Agr, pers:-, past] → was
 d. Be+[T, past] → was

Note that the grammars here are not in competition, in that there are no lexical doublets. Nor do the structures we output violate the Fox/Reinhart Generalization, since we don't have optional features that attach to functional heads.

Variation then arises because we have a choice of lexical items. The features on these items do not enter the interpretative component, and therefore don't bear on meaning. We can attach social meanings to these with impunity. So we have true Labovian variants. Change in a grammar, then, isn't due to competition driven by the instability of doublets, it must be driven, essentially, by fashion (that is by social changes).

Chapter 3

Negative Concord - A case of dissolving variation

3.1 Basic data

From the Buckie Corpus (Smith 2001)

[... [+neg] ... NEG]

- (1) You did na think nothing about hard work that time. (r:381.17)
- (2) They do na get nae love or nothing. (l:73.52)
- (3) You wi na have nothing to open the morn. (w:653.5)
- (4) She'll nae get nobody better than him. (x:307.25)
- (5) There was na nobody singing. (a:384.20)

[... [-neg] ... NEG]

- (6) They had nae sanitation or water . (b:162.56)
- (7) They've nae choice but go to the high school. (!:47.27)
- (8) She's haen nae problems. (§:85.60)
- (9) I've nae experience, you-ken. (3:15.39)
- (10) And I've nae smell or taste, so that is annoying. (3:184.10)
- (11) She kens nothing. (n:12029)
- (12) Listen, I think you're getting nothing. (n:446.60)
- (13) But there was nothing wrong wi' it. (6:175.64)

*[NEG ... [+neg] ...]

- (14) *Nobody didna get it
- (15) *Nobody you ken was na there.
- (16) *Nobody has na been there.

[NEG ... [-neg] ...]

- (17) Nobody could help you (h:32.30)
 (18) Nobody drinks it (1:58.90)
 (19) Nobody takes the water. (1:59.11)
 (20) Nobody bothered you know (1:356.31)

3.1.1 Quantitative considerations

N=359				
config	[+Neg ... NEG]	[-Neg ... NEG]	[NEG ... +Neg]	[NEG ... -Neg]
N	220	128	0	11
%	61	36	0	3

3.1.2 NC is not NPI licensing

- Possible to have [... [-neg] ... NEG] but not [... [-neg] ... ANY].
- Different licensers:

- (21) They had nae sanitation or water. (b:162.56)
 (22) *They had ony sanitation or water
 (23) Do you know anybody fae Elgin (c:169.63)
 (24) *Do you know naebody fae Elgin
 (25) *He left wi'oot talking to naebody

3.2 The syntax of NC

3.2.1 Negative Concord in Italian

- (26) Paola **non** ha visto **nessuno**
 Paola NEG have-3SG seen noone
 'Paola hasn't seen anyone.'
- (27) *Paola ha visto **nessuno**
 Paola have-3SG seen noone
 'Paola hasn't seen anyone.'
- (28) *Gli studenti che **non** avevano telefonato avevano detto **niente**
 The students that NEG have telephoned have said nothing
 'The students who didn't telephone said nothing.'
- (29)

3. How is [...[-neg] ...NEG] possible at all (the *uneg* feature of NEG wouldn't be checked)?

The Negative Spellout Rule: If [Spec, NegP] is filled during the course of the derivation, Neg⁰ has no phonological realisation.

→ The subject has to traverse through [Spec, NegP] on its way to [Spec, TP].

(33) Naebody came.

(34) *Naebody didna come.

Other dialects

(35) No stranger ain't got to come (Howe and Walker 1999:124) (Early African American English)

(36) Nobody couldn't handle him. (Wolfram and Christian 1976:112) (Appalachian English)

NSR is parametrized.

3.3 Why the NSR? A different tack

Starke(2002) - Specifiers don't really exist. Everything is a head complement relation, and what look like specifiers are just XPs Merged to satisfy the Functional Sequence (fseq)C-T-v-V.

- (37)
- [eat which pasta] → Merge DP[v] (v=agentivity)
 - [these boys][eat which pasta] → Merge DP[T] (T=nominative)
 - [these boys] t [eat which pasta] → Merge DP[wh]
 - [which pasta] [these boys] t [eat t]

Follows that there is no Specifier-Head relation, so no NSR. Rather if fseq requires negation, either something bearing a Neg feature (interpretable or not) must be externally Merged or internally merged (Remerged).

- (38)
- V NEG → Merge DP[v] (v=agentivity)
 - DP V NEG → is there a +neg in the numeration?
 - yes: neg DP V NEG = Negative Concord
 - no: NEG DP V t = Negative Quantifier

Note that (38) (c) followed by NEG movement is *, as is (38) (d) followed by Merge of +neg, since that would give us two instances of negation syntactically for one element of fseq. We then derive the NSR. This essentially updates Ladusaw 1992, with the difference that n-words are not ambiguous between NQs and NPIs, rather their interpretation depends on their syntactic context.

So we predict a negative head should occur with negative objects if the proposition is to be negative. This means that in such a language:

(39) I saw noone

should not be sentential negation, which seems correct as we'll see below.

An example of (38) (d): Icelandic NegXP movement.

- (40) Strákarnir höfðu hent grjóti í bílana
 the.boys had thrown rock in the.cars
 'the boys had thrown rocks at the cars.'
- (41) *Strákarnir höfðu grjóti hent í bílana
 the.boys had rock thrown in the.cars

Icelandic is uncontroversially VO.

- (42) *Strákarnir höfðu hent engu grjóti í bílana
 the.boys had thrown no rock in the.cars
- (43) Strákarnir höfðu engu grjóti hent í bílana
 the.boys had no rock thrown in the.cars
 'the boys hadn't thrown any rocks at the cars.'

What about subjects?

- (44) a. $V DP \rightarrow \text{Merge } DP[\text{NEG}, v]$ ($v=\text{agentivity}$)
 b. $\text{NEG } V DP \rightarrow$ is there a +neg in the numeration?
 c. yes: neg NEG $V DP =$ but if Neg raises over [neg] we have a configuration which violates at least some views of locality: NegXP ... [neg] ... $\langle \text{NegXP} \rangle$.
 d. no: NEG $t V DP =$ Sentential negation, but no negative concord.

Predicts

- (45) *Nobody didna come
 (46) Nobody could help you (h:32.30)

What about the other dialects where a subject occurs with a negative head?

- (47) *Anybody did [NEG $\langle \text{anybody} \rangle$] arrive \leftarrow prohibition on stranding NEG

Variation in NegXP subjects with Neg head is really variation in whether you can strand NEG

- (48) a. n+naebody \leftarrow Buckie, Italian
 b. n...nobody \leftarrow Appalachian, Spanish
 c. ambiguous \leftarrow Catalan

Summary:

- fseq approach predicts negative concord for objects and not for subjects.
- It also predicts sentential negation readings should arise with subjects, but not with objects, unless there is a head in Neg or the object raises to Neg.

What about the variability we saw in the $[-\text{Neg}] \dots \text{NEG}$ configuration?

3.4 Apparent variation in Negative Concord

3.4.1 Existentials

[\exists ... [+neg] ... NEG]

(49) There was na nae downies, you- see. (r:154.34)

(50) There was na nobody singing. (a:384.20)

(51) There was na nae fuss. (b:236.30)

[\exists ... [-neg] ... NEG]

(52) There 's nae Jennifer 's in Finichty. (t:977.27)

(53) There was nae work for making barrels and that. (c:14.22)

(54) You've nae herring and makes you wonder why there 's none. (b:39.46)

(55) But there was nothing wrong wi' it. (6:175.64)

(56) But there's nae fish, there's nae white fish. (7:195.0)

(57) There was no pensions or nothing. (g:262.13)

(58) There was no pubs or nothing in Uigg. (c:247.31)

3.4.2 Quantitative considerations

N=90		
config	[+Neg ... NEG]	[-Neg ... NEG]
N	4	83
%	5	95

Why the difference in numbers between existentials and non-existentials?

3.4.3 Main Verbs with NC

(59) And I've nae smell or taste, so that is annoying. (3:184.10)

(60) She kens nothing. (n:12029)

Separating out main verb *have* from other main verbs we find that of 45 cases, 39 occur with this verb. The other 6 are all in some sense emphatic (they bear focal stress, or are information structurally odd in other ways).

(61) ??She kens nothing, does she?

(62) ??She kens nothing, doesn't she?

Suggestion: these are not negation, satisfying fseq, but rather have only a Generalized Quantifier interpretation:

(63) {things} \cap {x:she knows x} = 0

Perhaps these raise to a low VP scope position, but cannot raise from there to Neg.

(64)

	[[+Neg] ... NEG]	[[Neg] ... NEG]
Lexical main verbs	97% (198)	3% (6)
Existentials	5% (4)	95% (83)
Main <i>have</i>	32% (18)	68% (39)

3.4.4 Negative Movement and Holmberg's Generalization

(Svenonius 2000)

- (65) Knut leverte ingenting inn
 Knut handed nothing in
 'Knut didn't hand anything in'
- (66) *...hvis Knut leverte ingenting inn
 ...if Knut handed nothing in
 'Knut didn't hand anything in'
- (67) *Knut har levert ingenting inn
 Knut has handed nothing in
 'Knut hasn't hand anything in'

Contrast is due to Holmberg's generalization:

- (68) Object shift cannot cross overt material in VP
- (69) Knut leverte ingenting [_{VP} t_V t_{neg} inn]
 Knut handed nothing in
- (70) hvis Knut [_{VP} leverte ingenting inn]
 if Knut handed nothing in

3.5 Holmberg's Generalization and NC in Buckie

The auxiliaries *have* and *be* in Buckie can appear **higher** than the negative particle, while main verbs must appear **lower**, just as in standard English.

- (71) I was **na** there
- (72) I had **na** been there
- (73) I **na** ken him

Since the auxiliaries are structurally higher than negation, Negative Movement can take place without violating HG.

- (74) *She [_{NegP} nothing [_{VP} kens t_{NEG}]]. Violates HG
- (75) She kens nothing (no sentential negation reading)
- (76) She does [_{NegP} na [_{VP} ken nothing]]

Compare:

(77) I had [_{NegP} na [_{VP} t_{have} nae bairns that time]] (e:712.47)

(78) And I've [_{NegP} nae smell or taste [_{VP} t_{have} t_{NEG}]], so that is annoying. (3:184.10)

and:

(79) There was [_{NegP} na [_{VP} t_{be} nae downies]], you- see. (r:154.34)

(80) There was [_{NegP} nae work for making barrels [_{VP} t_{be} t_{NEG}]] and that. (c:14.22)

Further predictions:

(81) She's haen nae problems. (§:85.60) - only Focus Reading available

(82) I did na hav nae tools or nothing, (j:829.25) - sentential reading

(83) You wi na have nothing to open the morn. (w:653.5) - sentential reading

(84) He was na in nae trouble.

(85) He was na feart of nothing.

Potential issue:

(86) He was na nae doctor

(87) He was nae doctor

Note crucially that what is relevant is the syntactic position of the verb, rather than its status as main or auxiliary.