Emergent parameters: hierarchies and phylogenies

Theresa Biberauer\textsuperscript{1,2} & Ian Roberts\textsuperscript{1}

University of Cambridge\textsuperscript{1} & Stellenbosch University\textsuperscript{2}

ReCoS Project (University of Cambridge)

http://www.mml.cam.ac.uk/dtal/research/recos/
The Rethinking Comparative Syntax Team
The ReCoS objective

- Rethinking **comparative** syntax = rethinking **parametric** syntax
  - ✗ richly specified UG with a fixed menu of parameters
  - ✓ steady-state (adult) grammars expressing (systematic) syntactic knowledge in parametric terms
    - success in capturing and predicting (unexpected) aspects of crosslinguistic variation (typology), diachronic change, and (maybe?) acquisition
    - new possibilities in the domain of phylogenetic studies
The case for parameters

• **Typology**: the existence of *structured variation*
  - Some hypothetical possibilities are unattested (*gaps*), yielding *asymmetries* whereby one option is common while its reverse is unattested/subject to stringent constraints.
  - In other cases, large numbers of unrelated languages share features (*commonalities*).

• **Diachrony**:
  - Evidence suggesting *pathways of change* (Biberauer & Roberts 2008, 2009, Biberauer, Holmberg & Roberts 2007 *et seq.*).
  - **Constant Rate Effects** (Kroch 1989): what changes in frequency during a syntactic change is language users’ overall tendency to choose one *abstract grammatical option* over another in their language production.
The case for parameters

<table>
<thead>
<tr>
<th>Dates</th>
<th>Negative Declaratives</th>
<th>Negative Questions</th>
<th>Affirmative Transitive Questions</th>
<th>Affirmative Intransitive Questions</th>
<th>Affirmative wh- object Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% do N</td>
<td>% do N</td>
<td>% do N</td>
<td>% do N</td>
<td>% do N</td>
</tr>
<tr>
<td>1400-1425</td>
<td>0 177</td>
<td>11.7 17</td>
<td>0 3</td>
<td>0 7</td>
<td>0 1</td>
</tr>
<tr>
<td>1426-1475</td>
<td>1.2 903</td>
<td>8.0 25</td>
<td>10.7 56</td>
<td>0 86</td>
<td>0 27</td>
</tr>
<tr>
<td>1476-1500</td>
<td>4.8 693</td>
<td>11.1 27</td>
<td>13.5 74</td>
<td>0 68</td>
<td>2.0 51</td>
</tr>
<tr>
<td>1501-1525</td>
<td>7.8 605</td>
<td>59.0 78</td>
<td>24.2 91</td>
<td>21.1 90</td>
<td>11.3 62</td>
</tr>
<tr>
<td>1526-1535</td>
<td>13.7 651</td>
<td>60.7 56</td>
<td>69.2 26</td>
<td>19.7 76</td>
<td>9.5 63</td>
</tr>
<tr>
<td>1536-1550</td>
<td>27.9 735</td>
<td>75.0 84</td>
<td>61.5 91</td>
<td>31.9 116</td>
<td>11.0 73</td>
</tr>
<tr>
<td>1551-1575</td>
<td>38.0 313</td>
<td>85.4 48</td>
<td>73.7 57</td>
<td>42.3 71</td>
<td>36.0 75</td>
</tr>
</tbody>
</table>

Table 1: Frequency of periphrastic ‘do’ sentences by context.

<table>
<thead>
<tr>
<th>Dates</th>
<th>% adverb–V</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1400-1425</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>1426-1475</td>
<td>23.5</td>
<td>154</td>
</tr>
<tr>
<td>1476-1500</td>
<td>34.7</td>
<td>186</td>
</tr>
<tr>
<td>1501-1525</td>
<td>69.4</td>
<td>109</td>
</tr>
<tr>
<td>1526-1535</td>
<td>88.8</td>
<td>170</td>
</tr>
<tr>
<td>1536-1550</td>
<td>89.8</td>
<td>152</td>
</tr>
<tr>
<td>1551-1575</td>
<td>89.2</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 2: Frequency of ‘never’–V word order in sentences with tensed main verbs.

Kroch (1989)
The unfeasibility of “classic” parameters

- Plausibility as part of an innate specification?
  - Macroparameters, e.g. the **HEAD PARAMETER**: Heads **PRECEDE/FOLLOW** their complements
  - Mesoparameters, e.g. the **NULL SUBJECT PARAMETER**: pro is **licensed by an appropriately specified INFL**
  - Microparameters, e.g. a **PRO-PREDICATE DO PARAMETER**:

  *Is John coming? He said he might (*US do).*

The unfeasibility of “classic” parameters

- The size of the variation space?

\[ n = |\text{parameters}|, \ L = \text{linguistic systems} \]

**Scenario A**: \[ |L| = 2^n = 8 \]

\[ \text{if } n = 30, \ |L| = 1073741824 \]

```
Parameter 1
  \text{N} \quad \text{Y}
```

```
Parameter 2
  \text{N} \quad \text{Y}
```

```
Parameter 3
  \text{N} \quad \text{Y}
```

**Scenario B**: \[ |L| = n+1 = 4 \]

\[ \text{if } n = 30, \ |L| = 31 \]

```
Parameter 1
  \text{N} \quad \text{Y}
```

```
Parameter 2
  \text{N} \quad \text{Y}
```

```
Parameter 3
  \text{N} \quad \text{Y}
```

- Acquisition sequence?
The unfeasibility of "classic" parameters


- **Empirical difficulties** (cf. i.a. Newmeyer 2005)
  - the predicted “slew” of empirical reflexes associated with individual parameters doesn’t seem to be “out there” in the empirical domain
The unfeasibility of “classic” parameters

• Most commonly cited case of a “failed” GB parameter: the Null Subject Parameter: pro is licensed by an appropriately specified INFL

  connection between null pronoun subjects and (rich) agreement

  plus: null pronouns > Free Inversion (1), and no that-trace effects (2)

  (1) Ha telefonato Gianni (vs Gianni ha telefonato)

  (2) a. Who did you say (that) you met t_who ?
          b. Who did you say (*that) t_who met you?  [non-null subject setting]
          c. Who did you say (that) t_who met you?  [null subject setting]

  no positive evidence for (b)
Minimalist parameters

- Minimalist parameters:
  - at least 2 potential loci (cf. also Richards 2009):
      - languages may differ as to i.a.:
        (a) the features they grammaticalise,
        (b) the way in which these features are bundled,
        (c) the loci in which features are overtly realised
    - the **PF interface** (cf. Berwick & Chomsky’s (2011) Externalisation proposal)
      - cf. (c) above
      (seemingly less plausible: parameters at the LF interface ...)

**Diagram:***

- **LEXICON**
- **SYNTAX**
- **PHONOLOGICAL FORM (PF)**
- **LOGICAL FORM (LF)**

---

10
Emergent Parameters

Proposals

• It is not necessary for parameters to be **hard-wired** for the notion of ‘parameter’ to be meaningful.
• It is also not necessary for parameters to “steer” acquisition in the manner traditionally assumed for them to play a role in addressing the logical problem of language acquisition.
• **Emergent** parameters, which capture the crosslinguistically parallel systematicities of adult grammars and whose parallel emergence across learners can be accounted for on a principled basis, validate a revised parametric approach:
  ▪ they may allow us to understand the typological and diachronic (and phylogenetic) patterns/regularities observed
  ▪ they may facilitate insight into the logical problem of language acquisition ... and possibly into new questions like the quantification of syntactic complexity
Emergent parameters

- A principled explanation for the emergence of minimalist parameters?
- Emergent parameters arise as a result of the interaction of Chomsky’s (2005) 3 factors:
  - **Factor 1**: the biological endowment for language, i.e. UG
  - **Factor 2**: the environmental input, i.e. the Primary Linguistic Data/PLD
  - **Factor 3**: non-language-specific considerations, which include:
    1. principles of efficient computation, and
    2. principles of data analysis employed in acquisition
       - e.g. Input Generalisation, Feature Economy
Emergent Parameters

What’s left in UG?

• The content of UG should be **minimal** (≠ empty)

• Chomsky (2001:10) on formal features and UG:
  ▪ "FL specifies the features F that are available to fix each particular language L"
  ▪ "We adopt the conventional assumption that L makes a one-time selection [FL] from F. These are the features that enter into L; others can be disregarded in use of L"

• Gianollo, Guardiano & Longobardi (2008): **parameter schemata**
Emergent Parameters

• Zeijlstra (2008; cf. also i.a. Evers & van Kampen 2008, Biberauer 2011):

Formal features = syntactically visible/grammaticalized features
What’s left in UG?

• Biberauer (2011, 2013), Biberauer & Roberts (2013):

  - UG-specified: only formal features not intersecting with the set of semantic features
    - “intersecting” features can be acquired from the data: doubling/agreement, movement, etc.

  - UG may contain: [uF]/[iF] template, ^ (=movement diacritic), certain formal features like [Person], ![Case]; Merge, Agree (Internal and External Select, as in Rizzi 2009?)
Outline of the rest of the talk

1. ReCoS thinking on the **shape and role** of hierarchies
2. Case study 1: a hierarchy for **word order**
3. [Case study 2: a hierarchy for **null arguments**]
4. Case study 3: a hierarchy for **negation**
5. Emergent hierarchies and diachrony: the matter of **stability**
6. Emergent hierarchies and **complexity**
ReCoS parameter hierarchies

ReCoS parameter hierarchies: the role of Factor 1

• Parametric variation emerges where UG doesn’t mind (Biberauer & Richards 2006)

• Parameters arise from **underspecification** of formal features in UG in relation to i.a:
  a. association of formal features with (functional) heads
     e.g. scattering/syncretism (Giorgi & Pianesi 1997), presence/absence of [gender] or [tense]
  b. values of formal features, triggering Agree
     e.g. whether agreement inflection is [uF] (English) or [iF] (Italian), whether a negation element is [uF] or [iF] (Negative Concord vs “Double Negation” systems)
  c. features triggering movement (Internal Merge)
     e.g. V-to-l, wh-movement, etc.
ReCoS parameter hierarchies: the role of Factor 1

• What UG gives the acquirer:
  ▪ the “expectation” that language is structured around not just phonological [P] and semantic [S] features, but also structurally crucial formal features [F]
  ▪ the [uF]/[iF] template as the basis for analysing the input and representing speakers’ formal knowledge
  ▪ certain prespecified formal features [F] to distribute among the lexical items in the input – e.g. the movement trigger, [Person], (?) [Case]

• Consequence: we expect regularities across I-languages reflecting this shared, UG-given formal structuring imperative
ReCoS parameter hierarchies: the role of Factors 2 & 3

- A given formal feature $F$ may associate with a different set of heads (including the empty set) in different languages.
- How does the child acquire these (parametric) specifications?
- **Proposed learning procedure:**
  1. default assumption: $\neg \exists h [ F(h)]$
  2. if $F(h)$ is detected, generalise $F$ to all relevant cases
     $$ (\exists h [ F(h)] \rightarrow \forall h [ F(h)] ) $$
  3. if $\exists h \neg [ F(h)]$ is detected, restrict $h$ and go back to (i)
  4. if no further $F(h)$ is detected, stop.
ReCoS parameter hierarchies: the role of Factor 3

• Why this procedure?

• 2 central acquisition-oriented Factor 3 considerations (“principles of data analysis”):

a. Feature Economy (FE) (Roberts & Roussou 2003: 201)

Given two structural representations R and R’ for a substring of input text S, R is less marked than R’ iff R contains fewer formal features than R’.

b. Input Generalisation (IG) (Roberts 2007)

If a functional head F sets parameter $P_j$ to value $v_i$ then there is a preference for (similar) functional heads to set $P_j$ to value $v_i$. 
ReCoS parameter hierarchies: their shape

- Together, FE and IG form a **minimax search/optimisation algorithm** (FE: minimise features; IG: maximise detected features).
- They also naturally result in hierarchies with the following general “shape”:

```
F present?
  /       \
 /         \
NO    YES: All heads?
  /       \
 /         \
YES  NO: Which subset of heads?
          [postulate a new [F]]
```
ReCoS parameter hierarchies: their shape

- **NO > ALL > SOME** acquisition sequence

- **NO** = default as the acquirer doesn’t receive (systematic) input pointing to need for FE & IG respected

- **ALL**: plausibly follows from the acquirer’s initial “ignorance” (Biberauer 2011, Branigan 2012) → **IG respected & FE minimally violated** (“Make maximal use of minimal means” – cf. Biberauer 2011)

- **SOME**: both IG & FE violated, but FE will dictate that the violation should be minimal

- “Recovery” from **superset traps** (Berwick 1985) possible as the acquirer must postulate appropriately specified heads to capture the data, with the [uF]/[iF] template in combination with IG and FE gradually reducing the acquirer’s “ignorance”
ReCoS parameter hierarchies: the significance of their shape

- True macroparameters sit at the top of the network. As we move successively down:
  - **Typology**: Systems become more marked (in featural terms)
    - Parameters have a longer description (the conjunction of all the “nodes”)
    - Parameters become meso then micro then nano
    - Complexity increases as we “go down”
  - **Acquisition**: Lower parameters are further along a learning path
  - **Diachrony**: Systems become less stable
Parameter hierarchies give rise to a rough taxonomy of parameter types (Biberauer & Roberts 2012, 2013):

For a given value $v_i$ of a parametrically variant feature $F$:

a. **Macroparameters**: all heads of the relevant type share $v_i$;

b. **Mesoparameters**: all heads of a given naturally definable class, a subset of the full class of heads of the relevant type, e.g. $[+V]$, share $v_i$;

c. **Microparameters**: a small subclass of functional heads (e.g. modal auxiliaries, pronouns) shows $v_i$;

d. **Nanoparameters**: one or more individual lexical items is/are specified for $v_i$. 

• **Hierarchy 1: Word order**

Is head-final present?

- No: **head-initial**
- Yes: present on all heads?
  - Yes: **head-final**
  - No: present on [+V] heads?
    - Yes: head-final in the clause only
    - No: present on ...

See Biberauer, Holmberg, Roberts & Sheehan (2010), Biberauer, Holmberg & Roberts (2013), Biberauer, Roberts & Sheehan (2013) for empirically motivated argumentation in favour of the idea that head-finality is marked, i.e. requires the presence of a feature absent in head-initial structures. The feature in question: ^, the UG-given movt diacritic.
ReCoS parameter hierarchies: WORD ORDER

- Word order = a parameter that’s set very early (cf. Wexler’s VEPs)
  - (maximally) robustly attested in the PLD
  - Typologically, strong tendency towards “harmony” (Greenberg 1963, Hawkins 1983): heads in the system show consistent ordering
  - Full vs partial harmony: insights from the Final-over-Final Constraint (FOFC) (Biberauer, Holmberg & Roberts 2007 et seq.)

(1) If $\alpha$ and $\beta$ are part of the same Extended Projection and $\alpha$ is a head-initial phrase, with $\beta$ dominating $\alpha$, then $\beta$ must be head-initial.

(If $\alpha$ is a head-final phrase, and $\beta$ is a phrase immediately dominating $\alpha$, then $\beta$ can be head-initial or head-final.)
The GB Head Parameter cannot cover these facts without further stipulation.
ReCoS parameter hierarchies: WORD ORDER

FOFC-compatible systems

a. VP [+V]
   D/PP [-V]  V
   D/P  NP  [German, Dutch...]

b. VP [+V]
   V  D/PP [-V]
   D/P  NP  [English, French ...]

c. VP [+V]
   D/PP [-V]  V
   NP  D/P  [Japanese, Malayalam ...]

d. VP [+V]
   D/PP [-V]
   NP  D/P  [Gungbe, Thai ...]

(b) = NO head-initial; (c) = ALL head-final; (a) only [V] head-final;
(d) = only [N] head-final
ReCoS parameter hierarchies: WORD ORDER

• FOFC-incompatible systems:

- No VO plus final auxiliary
- No VO plus final complementiser (subordinator) (Hawkins 1983, Dryer 2009)
• FOFC implies that there will be **pathways for word-order change**:
  
  - Change from head-final to head-initial order must proceed “top-down”:

    (1) \[[[O V I] C] \rightarrow [C [[O V] I]] \rightarrow [C [I [O V]]] \rightarrow [C [I [V O]]]\]

  - Change from head-initial to head-final order will have to proceed ‘bottom-up’:

    \[ [C [I [V O]]] \rightarrow [C [I [O V]]] \rightarrow [C [[O V] I]] \rightarrow [[[O V] I] C]\]

  Thus: **directionality** as there is a pathway of diachronic change determined by synchronically impossible stages
The pathway holds in relation to both “internal” and contact-induced change:

- Icelandic conforms (cf. Hróarsdóttir 2000)
- Yiddish conforms (cf. Diesing 1992 and Beatrice Santorini, p.c.)
- Kaapse Afrikaans conforms:

(1) a. ... dat ek **het** [\_VP gekry R1400]  
    that I have got  R1400 = ‘... that I received R1400’  
    **Aux-VO**

b. ... dat ek [\_VP R1400 gekry] **het**  
    **OV-Aux**

c. ... dat ek **het** [\_VP R1400 gekry]  
    **Aux-OV**

d. * dat ek [\_VP gekry R1400 van die Revenue] **het**  
    *VO-Aux
ReCoS parameter hierarchies: WORD ORDER

• Permitted patterns (\(^\wedge\) signals “head-final”):

1 (a) V\(^\wedge\) - Aux\(^\wedge\) - C\(^\wedge\) (consistently head-final: Japanese)
(b) V\(^\wedge\) - Aux\(^\wedge\) - C (head-final barring subordinator: German)
(c) V\(^\wedge\) - Aux - C (head-final in VP only: some Mande languages)

• Barred patterns:

2 (a) V - Aux\(^\wedge\) - C
(b) V - Aux\(^\wedge\) - C\(^\wedge\)
(c) V - Aux - C\(^\wedge\)

Head-finality starts at the bottom of an Extended Projection, and once it’s stopped, it’s stopped for good.
NB: the partial harmony patterns in (1).
What guarantees the FOFC pattern?
Crucial ingredients:

- acquirer’s initial focus on **lexical categories** \((V, N) = \text{bottom of EPs}\): most salient in the input, with clear semantic and phonological contrasts between V and N (cf. i.a. Christophe, Nespor, Guasti and van Ooyen 2003, Bion, Höhle and Schmitz 2007, Gervain & Werker 2008, 2012, Gervain & Mehler 2010, etc. on phonological properties; and Braine 1963, much work by Dedre Gentner and Ellen Markman, Lebeaux 1988, 2000, Roeper 2011, 2012 on semantic properties)

- the UG-given \([uF]/[iF]\) template and UG-given \(^\text{\footnotesize{movt diacritic}}\) \(F1\)

- acquirer’s FE- and IG-driven approach to the input \(F3\)

- acquirer’s “greedy” approach to the use of \(^{\text{\footnotesize{F3}}}\)
ReCoS parameter hierarchies: WORD ORDER

• In FOFC-compatible languages:
  ▪ verbs are [V] (initial) and higher clausal elements are [V] or not specified as specifically verbal (i.e. [V]; plausibly, the case for non-inflecting Asp, Tense and Mood particles); or
  ▪ verbs are [V^] (final) and higher clausal elements are either [V^] or [V]
  ▪ Key: if the acquirer doesn’t associate ^ with category of [V]-specified heads, (s)he can’t subsequently create a sub-category of [V]-specified heads which is [V^]

    ➢ In formal-feature [F] terms: sub-categories of an existing category require the postulation of an additional substantive feature [uF]/[iF] (and you could potentially further lose ^)

    ➢ so ^ can only be introduced at the bottom of an EP defined by the categorial feature of the bottommost head (cf. Grimshaw 1990 et seq.)
• **Hierarchy 2: Null arguments**

\[ \Phi = \text{[Person]}, \text{[Number]} \text{ and } \text{[Gender]}, \text{i.e. agreement features} \]

Are \( \varphi \)-features present in the system (on probes)?

- No
- Yes

**Radical pro-drop**

Are \( \varphi \)-features present on all probes?

- Yes
- No

**Pronominal arguments**

Are \( \varphi \)-features fully specified on some probes?

- No
- Yes

**Non-null subject**

Are the \( \varphi \)-features of fully specified on \( T \)?

**Canonical null subject** (Italian, etc.)
Typically distinguished types of negation systems (cf. Biberauer & Zeijlstra 2012a,b):

- **DOUBLE NEGATION/DN** languages (e.g. English)
  - sentential negator and negative indefinites: [iNEG]

- **STRICT NEGATIVE CONCORD/NC** languages (e.g. Czech)
  - sentential negator and negative indefinites: [uNEG] ([iNEG] = abstract OP – Ladusaw 1992)

- **NON-STRICK NC** languages (e.g. Italian)
  - sentential negator: [iNEG]. and negative indefinites: [uNEG]
A typological gap

- A typological gap, barring standard Afrikaans: PARTIAL NC languages

- sentential negator: [uNEG], and negative indefinites: [iNEG]

(1) Dit is nie[uNEG] moontlik nie[uNEG]  
  it is not possible not  
  ‘It is not possible’

(2) Niemand[NEG] het niks[NEG] nie[uNEG]  
  no-one has nothing not  
  ‘No-one has nothing’, i.e. Everyone has something
Why the gap?

• What Afrikaans speakers do colloquially:

(3) \text{Niemand}_{\text{uNEG}} \text{ het } \text{niks}_{\text{uNEG}} \text{ nie}_{\text{uNEG}}

no-one has nothing not

‘No-one has anything’, i.e. Strict NC where all negative elements = [uNEG]

• Why?
  - Absence of \textbf{unambiguous PLD}: [iNEG] negative indefinites \( \Rightarrow \) DN readings ([iNEG] + [iNEG] = 2x[NEG])
  - BUT: DN readings also possible in NC systems, with the same characteristic intonation contour (Liberman & Sag 1974)
Factor 3 overrides Factor 2

- DN: $[\text{Foc} \ [\text{iNEG}]] + [\text{iNEG}] = 2x[\text{NEG}]$
- NC: $\{[\text{Op} \ \text{iNEG}]\} [\text{Foc} \{[\text{Op} \ \text{iNEG}]\}[\text{uNEG}]] [\text{uNEG}]$
  $= 2x[\text{NEG}]$

- Thus: **F3**, Input Generalisation (here: “Assume all negative elements to be [uNEG] if you have clear evidence that [uNEG] is part of the system”), and **F2**, the absence of PLD supporting the IG-violating, but theoretically available possibility that some element may be [iNEG] conspire to rule out Partial NC systems (these require prescriptive support).
A parametric hierarchy for negation

Is [NEG] a formal feature?

N
- Standard Dutch (?)

Y
- DN parameter: Are all negative elements [iNEG]?
  N
  - Non-strict NC parameter: Are all sentential negators [iNEG]?
    Y
    - Italian
    N
    - Strict NC parameter: Are all NIs [iNEG]?
      Y
      - OUT
      N
      - Czech

Adapted from Biberauer & Zeijlstra (2012c)
No-choice parameters

• Because the hierarchies are systematically constructed on the basis of ever more specific questions about categories and/or feature distributions, with each question being binary (Y/N), we expect no-choice parameters to arise.

  ▪ Further constrain the space of variation

  ![Diagram showing decision process involving parameters A and B, each with binary (Y/N) outcomes.]

  ▪ cf. also Biberauer (2013), Biberauer, Roberts & Sheehan (2013) and Sheehan (2013) for discussion of other no-choice scenarios
ReCoS hierarchies and diachrony

• Diachronic correlates (Biberauer & Roberts 2012b, Branigan 2012):
  a. Macroparameters are **stable over millennia**
     e.g. (i) (Multiple) Incorporation in the Algonquian languages (Branigan 2012):
     • Proto-Algonquian spoken 2000-3000 years ago. Since then, numerous structural, lexical and phonological features have changed, but incorporation has remained as a “signature” property of the family.
     • Assume (for concreteness) a new generation of native speakers emerges every 25 years.
       → 3000 years = 120 iterations of the learning cycle
  (ii) Harmonic head-final order in Dravidian (Steever 1998:31)
  • Proto-Dravidian dated to 6000 years ago,
    → roughly 240 iterations of the learning cycle
b. Mesoparameters are **somewhat stable**, often characterising genera.
   e.g. (i) null subjects in (most) Romance
   • Pronouns: Strong subject pronouns > weak subject pronouns > syntactic clitics
   • *Pro*: Full NSL > restricted NSL > non-NSL > NSL again
     Latin/rest of Rom > OF/Med NIDS > literary Fr/16\(^{th}\) C NIDs > “advanced” Fr/NIDs
   • NIDs and OF suggest the intermediate stages were short-lived – probably 200 years for each stage post-full NSL, i.e. around 8 learning cycles. Full NSLs appear to be stable, except in the face of contact (cf. the Romania Nova)

(ii) V2 in (most) Germanic
   • V2 has remained remarkably stable across nearly all North and West Germanic varieties, with English changing in the 15\(^{th}\) century for obscure reasons (contact again involved)
c. Microparameters are somewhat unstable.
   e.g. Aux-movement in Modern English (recent (post-1600), idiosyncratic both typologically and genetically)
   - the class of modals arose via a microparametric change (grammaticalisation is always micro- or nanosyntactic parametric change)
   - the modals seem to have started to change in the 18th century, 200 years ago, i.e. a mere 8 iterations of the learning cycle, after their creation as a separate class

d. Nanoparameters are highly unstable.
   e.g. the ongoing collapse of the aux-system in contemporary varieties of English (Biberauer & Roberts 2012a)
ReCoS hierarchies and Complexity (?)

• Biberauer, Holmberg, Roberts & Sheehan/BHRS (2013):
  
  ▪ Using the 5 “ReCoS” hierarchies in (1), and taking into account the number of Y/N questions the acquirer has to postulate, we can attempt an initial comparison of syntactic complexity.

(1) H1: Word Order (^ distribution)
H2: Null Arguments (φ distribution)
H3: Word-structure (head-movement)
H4: Discourse Configurationality (A’-movement)
H5: Alignment (of case and agreement marking)
ReCoS hierarchies and Complexity (?)

Probability of a given output of the hierarchy = \(0.5^n\) where \(n\) is the level of embedding
• Considering the 5 hierarchies in BHRS, the least-marked possible system would have to be:

a. Harmonically head-initial
b. Radical pro-drop
c. High analyticity
d. No A’-movement (i.e. no mechanism of focussing, topicalisation, \textit{wh}-movement, scrambling)
e. Accusative alignment (or no Case if Parameter 5 is expanded upwards)

• Some creoles come close (they tend to have \textit{wh}-movement)
ReCoS hierarchies and Complexity (?): quantifying probability

• Languages tested: English, Japanese, Mohawk, Basque and Chinese

• **English:** (basically) harmonically head-initial (0.5 on H1), non-pro-drop (0.125 on H2), Aux but not V-movement (0.03125 on H3), *wh*-movement but no scrambling (0.03125 on H4), and is accusative (0.5 on H5).
  
  product of the 5 probabilities = **0.003%**
  i.e. English is relatively complex

• **Mohawk:** default head-initial (0.5 on H1), has pronominal arguments (0.25 on H2), polysynthesis (0.25 on H3), free word order (0.25 on H4), and split-S alignment (0.25 on H5).
  
  product of the 5 probabilities = **0.195%**

... And so on for Japanese, Basque and Chinese
The smaller the probability product is, the more complex (and thus less probable) the grammar of the language.

Table 1: Parametric probability summary
ReCoS hierarchies and Complexity (?): number of parametric choices

Worth noting is that initial N “choices” (i.e. defaults) aren’t real choices and so wouldn’t count towards the choice tally.
### ReCoS hierarchies and Complexity (?): number of parametric choices

<table>
<thead>
<tr>
<th>Language</th>
<th>Probability</th>
<th>Average number of choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>0.0003</td>
<td>3</td>
</tr>
<tr>
<td>Mohawk</td>
<td>0.195</td>
<td>1.8</td>
</tr>
<tr>
<td>Basque</td>
<td>0.098</td>
<td>2</td>
</tr>
<tr>
<td>Japanese</td>
<td>0.391</td>
<td>1.6</td>
</tr>
<tr>
<td>Chinese</td>
<td>0.098</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Comparison of the two methods of quantifying complexity

A consistent complexity ranking

**English > Basque & Chinese > Mohawk > Japanese**
Conclusions

• It is not necessary for parameters to be **hard-wired** for the notion of ‘parameter’ to be meaningful.
• It is also not necessary for parameters to “steer” acquisition in the manner traditionally assumed for them to play a role in addressing the logical problem of language acquisition.
• **Emergent** parameters, which capture the crosslinguistically parallel systematicities of adult grammars and whose parallel emergence across learners can be accounted for on a principled basis, validate a revised parametric approach:
  ▪ they may allow us to understand the typological and diachronic (and phylogenetic) patterns/regularities observed
  ▪ they may facilitate insight into the logical problem of language acquisition ... and possibly into questions like the quantification of syntactic complexity