The journey, not the endstate: finding innovation in the dynamics of L1A

Naomi Lee & Ailís Cournane
New York University

Unraveling approaches to learning morphosyntax link diachrony and acquisition of a language’s diachronic path towards economical adult grammars.
Unraveling approaches to learning morphosyntax link diachrony and acquisition

a language’s diachronic path towards economical adult grammars

children’s dynamic acquisition paths through input-divergent interim grammars
Unraveling approaches to learning morphosyntax link diachrony and acquisition

- a language’s diachronic path towards economical adult grammars

? children’s dynamic acquisition paths through input-divergent interim grammars
Unraveling approaches to learning morphosyntax link diachrony and acquisition

a language’s diachronic path towards economical adult grammars

unraveling

children’s dynamic acquisition paths through input-divergent interim grammars
1. Child analyses as innovations
diachrony from learning patterns, not vice versa
Most generative approaches to diachrony implicate children as innovators

“The changes in language fulfil themselves in the individual...the main period for the exercise of such influence [from others] is the time of the first acquisition - of the learning of language”
- Paul 1920 [1880]:15

 “[in a child’s] task of acquiring its first language...the child will be driven to a...grammar...[with] a degree of indeterminacy [that is required by] the view of syntactic change outlined here”
- Lightfoot 1977:194

“We assume that parameter change is an aspect of the process of parameter setting [and] a change is initiated when (a population of) learners converge on a grammatical system which differs in at least one parameter value from the system internalized by the speakers whose linguistic behaviour provides the input to the learners”
- Roberts & Roussou 2003:11

“the claim [is] that imperfect language learning, even by children acquiring their first language can be a source of language change”
- Kroch 2005:25

“I argue that cycles are the result of reanalysis by the language learner” and are due to “internal principles that bias the learner toward certain structures”, “emphasiz[ing] the role of the child learning the language”
- van Gelderen 2011: 4,6
However, proposed innovations have been back-engineered by comparing adult statesWithin well-established cycles of directionally recurring changes, children are taken to abandon ‘less economical’ (van Gelderen 2011) or ‘more elaborate’ (Roberts & Roussou 2003) adult analyses in reanalysis.
Syntactic change is therefore credited to simplifying economy principles that direct learning...

- The economy of the "simpler" adult state is taken to drive learning and cause change:

Adapted from Roberts & Roussou 2003: 135 - example (5), building on analysis given by Giusti 2001 and summarized in Roberts & Roussou 2003: 134-6
Syntactic change is therefore credited to simplifying economy principles that direct learning...

- The economy of the **“simpler” adult state** is taken to drive learning and cause change:
  - Minimize Feature Syncretisms (Roberts & Roussou 2003)

1. # of structural positions with PF realization of more than one formal feature

Adapted from Roberts & Roussou 2003: 135 - example (5), building on analysis given by Giusti 2001 and summarized in Roberts & Roussou 2003: 134-6
Syntactic change is therefore credited to simplifying economy principles that direct learning - that emerge under specific input

- The economy of the “simpler” adult state is taken to drive learning and cause change:
  - Minimize Feature Syncretisms (Roberts & Roussou 2003)

- Economical innovations arise under exceptional, insufficient input data
  - Relevant child is an innovator in a haystack
In fact, all children’s grammars are input-divergent through the learning process.

Speakers faced with the same input data go through “intermediate, transitory stages of the acquisition process” (Hale 2003:346)...and learning principles help determine which stages consistently follow others.
In fact, all children’s grammars are input-divergent through the learning process. Speakers faced with the same input data go through “intermediate, transitory stages of the acquisition process” (Hale 2003:346)...and learning principles help determine which stages consistently follow others.
In fact, all children’s grammars are input-divergent through the learning process.

Speakers faced with the same input data go through “intermediate, transitory stages of the acquisition process” (Hale 2003:346)...and learning principles help determine which stages consistently follow others.

This diagram draws inspiration from Hale 2003 & 2007.
In fact, all children’s grammars are input-divergent through the learning process.

Speakers faced with the same input data go through “intermediate, transitory stages of the acquisition process” (Hale 2003:346)...and learning principles help determine which stages consistently follow others.

This diagram draws inspiration from Hale 2003 & 2007.

Staying in one of these earlier input-divergent grammars constitutes innovation (Cournane 2017).
Roadmap

1. Child-first diachrony

2. Unraveling nominal morphosyntax

3. Empirical coverage of unraveling:
   ○ Canonical, economical: Dem to D ........................................... Latin > Romance
   ○ Non-canonical: Gender & Num .............................................. Proto-Cushitic > Somali

4. Conclusions
2. Unraveling nominal morphosyntax

recentering the learning task
The goal of morphosyntactic learning is to map forms to meanings (Roots or features).

**Input utterances**

- Adult Khoekhoe speaker
  - \(k^h\text{oè}k^h\text{ò}\) you two male people
  - \(k^h\text{oè}r\text{ò}\) you two female people
  - \(k^h\text{oè}k\text{ó}\) you many male people

**Segmentation**

**Vocabulary Items (form-meaning mapping)**

- PF spellout \(\sqrt{PERSON}\)
  - \(k^h\text{oè}\)
    - masculine
  - \(k^{(h)}\)
    - feminine
  - \(r\)
    - singular
  - \(\text{ò}\)
    - dual
  - \(\text{ó}\)
    - plural

1. Lee 2019 (fieldwork)
Let’s take as given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”¹...

(even so, generative approaches to the syntactic organization of features still differ - many models exist)
The unraveling approach induces the syntactic organization of those features
rather than assuming they exist in innately known (cartographic) structures
Less innate knowledge means the learning task determines children’s interim states

Given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”\(^1\)...

<table>
<thead>
<tr>
<th>Innatist</th>
<th>Inductive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full configurational knowledge</td>
<td>No syntactic knowledge about features</td>
</tr>
</tbody>
</table>

How much knowledge of the syntactic organization of features does the child start with?

1. Roberts & Roussou 2003:6-7
Less innate knowledge means the learning task determines children’s interim states

Given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”¹...

How much knowledge of the syntactic organization of features does the child start with?

**Innatist**

Full configurational knowledge

- Num°
- sg
- pl
- n°
- fem
- masc
- √Root

**Inductive**

No syntactic knowledge about features

1. Roberts & Roussou 2003:6-7
Less innate knowledge means the learning task determines children’s interim states

Given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”¹...

How much knowledge of the syntactic organization of features does the child start with?

What is the child’s learning task?

**Innatist**
- Full configurational knowledge
- Match a known structure to pronounced pieces

**Inductive**
- No syntactic knowledge about features

---

1. Roberts & Roussou 2003:6-7
Less innate knowledge means the learning task determines children’s interim states

Given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”¹...

How much knowledge of the syntactic organization of features does the child start with?

**Innatist**

Full configurational knowledge

- \text{Num}^* (sg, pl)
- \text{N}^* (fem, masc)
- √Root

**Inductive**

No syntactic knowledge about features

- sg, fem
- pl, masc

What is the child’s learning task?

- Match a known structure to pronounced pieces

---

¹ Roberts & Roussou 2003:6-7
Less innate knowledge means the learning task determines children’s interim states

Given that Merge, “the features themselves, ...and the nature of the movement operation are innately given as aspects of UG”...

1. Roberts & Roussou 2003:6-7
The unraveling child initially groups all features together in one syntactic position

**Innatist** - e.g. cartography

1. XP
2. [Y] YP
3. [X] WP
4. [Z] ZP
5. ... VP
6. √Root

* Possibility: Merge/Move/both

**Inductive** - e.g. unraveling

1. FP
2. [X, Y, Z, W, ...] √Root

- **Posit single feature bundle - the most specific VI** - pair all features and all non-Root phonological material
- Learn by **generalizing** and comparing overlapping VIs
- **Unravel** features into separate syntactic positions
The child generalizes over repeated occurrences of features to unravel that initial bundle.

chicos
male, plural

chicas
female, plural

chicos
\( \text{chicos} \)

F
plural, masc

\( \sqrt{\text{CHILD}} \)

-\( \text{as} \)

chicas
\( \text{chicas} \)

F
plural, fem

\( \sqrt{\text{CHILD}} \)

"∅"
The child generalizes over repeated occurrences of features to unravel that initial bundle

chicos
male, plural

chicas
female, plural

chica
female, singular
The child generalizes over repeated occurrences of features to unravel that initial bundle

- **chicos**: male, plural
- **chicas**: female, plural
- **chica**: female, singular
- **chicas**: female, plural
An unraveling inductive approach often predicts economical interim child analyses

- Starting with a maximally specific analysis of non-Root lexical material as bundled features means children base-generate PF material “higher” than in adult analyses (Cournane 2016)

- As a consequence of staying “earlier” along the acquisition path, changes will appear simpler and more economical because complexity builds in acquisition
  - Directional diachronic patterns are a product of the necessary ordering of interim grammars in acquisition

Pannemann 2007
3a. A canonical case study

from demonstrative to determiner:
Spec-head reanalysis
Latin’s *ille* demonstrative occupied the specifier of a D head (Giusti 1998, Lyons 1999)

```
[ illos          fortissimos      viros ]_{DP}
DEM.m.pl.acc   strongest.m.pl.acc   man.m.pl.acc
"those strongest men"
M. Tullius Cicero 63 BCE: De Lege Agraria: 1, 2, 7

[ vir  ille        fortissimus ]_{DP}
man.m.sg.nom  DEM.m.sg.nom  strongest.m.sg.nom
"those strongest men"
Seneca 40-45 CE: Ad Marciam de consolatione: 1.7

[ vir        summo  ingenio ]_{DP}
man.m.sg       greatest    capacity
"the man of greatest capacity"
M. T. Cicero 63 BCE: Oration XVI, Sect. VI
```

“Considerable freedom of position [word order], like adjectives”

Occur optionally in definite contexts
Latin’s *ille* demonstrative occupied the specifier of a D head (Giusti 1998, Lyons 1999)

1. Lyons 1999:118
2. Lightfoot 1979, 1999

---

**Diagram:**

```
DP
   DemP
      ille
       D
          ∅
```

**Example 1:**

```
[ illos fortissimos viros ]_dp
DEM.m.pl.acc strongest.m.pl.acc man.m.pl.acc
"those strongest men"
M. Tullius Cicero 63 BCE: De Lege Agraria: 1, 2, 7
```

**Example 2:**

```
[ vir ille fortissimus ]_dp
man.m.sg.nom DEM.m.sg.nom strongest.m.sg.nom
"those strongest men"
Seneca 40-45 CE: Ad Marciam de consolatione: 1.7
```

**Example 3:**

```
[ vir summo ingenio ]_dp
man.m.sg greatest capacity
"the man of greatest capacity"
M. T. Cicero 63 BCE: Oration XVI, Sect. VI
```

---

1. Considerable freedom of position [word order], like adjectives

---

640/644 in Caesar

4/644 in Caesar

Occur optionally in definite contexts
Given Latin input, an “unraveling” learner would straightforwardly posit the higher head to start

illos viros

I abstract away from Case features in this schematic treatment, showing different Case forms only as a way that Latin-learning children might have begun to segment their input.
Given Latin input, an “unraveling” learner would straightforwardly posit the higher head to start.

Initially posit a [definite] & [deix]-including, fully bundled terminal in head position.

I abstract away from Case features in this schematic treatment, showing different Case forms only as a way that Latin-learning children might have begun to segment their input.
Given Latin input, an “unraveling” learner would straightforwardly posit the higher head to start

Initially posit a [definite] & [deix]-including, fully bundled terminal in head position

I abstract away from Case features in this schematic treatment, showing different Case forms only as a way that Latin-learning children might have begun to segment their input.
Given Latin input, an “unraveling” learner would straightforwardly posit the higher head to start.

Initially posit a [**definite**] & [**deix**]-including, fully bundled terminal in head position

Later evidence required for a non-head analysis (not shown)

I abstract away from Case features in this schematic treatment, showing different Case forms only as a way that Latin-learning children might have begun to segment their input.
Given Latin input, an “unraveling” learner would straightforwardly posit the higher head to start

Initially posit a [definite] & [deix]-including, fully bundled terminal in head position

Later evidence required for a non-head analysis (not shown)

➢ The children who innovate do not consider and discard a more complex analysis, but never consider it at all

➢ The resulting grammar is simpler, not due to economy, but from children’s tactics for approaching the inductive learning task

I abstract away from Case features in this schematic treatment, showing different Case forms only as a way that Latin-learning children might have begun to segment their input
But accepted synthesis (Roberts & Roussou 2003, van Gelderen 2011) is that Dem > D was economy-driven.

Descriptive characterizations of the change are ultimately **economy-driven**:

- **Merge > Move**: Move requires one more feature syncretism (in the base-generated position) than Merge
- **Head Preference Principle**: specifiers acting as goals bear interpretable features, and don’t keep the derivation going - that is less economical than heads with uninterpretable features

---

1. Roberts & Roussou 2003  
2. van Gelderen 2011
...and requires cartographic knowledge of the syntax of Dem & D heads and their defining features

- Ultimately, **economy** is calculated by comparing derivations
- If children’s initial analyses are “economical” without transderivational comparison, their character must be due to acquisition itself

1. Giusti 1998 uses [referential] as the defining feature of D, and [deixis] as that of Dem
2. Lyons 1999 uses [definite] as the defining feature of D
With unraveling, innovation consists of staying at a **commonplace** earlier analysis.

Rather than comparing derivations, all children simply **proceed as best they can** in analyzing their input.

**Innovation** is any analysis inconsistent with input basis.

**Actuation** is the persistence of a learning phase through peer-to-peer reinforcement.

---

1. not just those with special exposure to specially insufficient input data
With unraveling, innovation consists of staying at a **commonplace** earlier analysis.

Rather than comparing derivations, all children simply **proceed as best they can** in analyzing their input.

**Innovation** is any analysis inconsistent with input basis.

**Actuation** is the persistence of a learning phase through peer-to-peer reinforcement.

---

**How can we test** our diachronic account - based on theorizing about acquisition of a dead language?

- Language with **pre-reanalysis properties**: Russian
- Investigate whether most children\(^1\) learning it go through a phase representative of the “**next step**” on the cycle?: head stage

---

1. not just those with special exposure to specially insufficient input data
3b. a non-canonical case study

number and gender features: never unraveled
Before unraveling, interim analyses bundle otherwise unexpected features together

- **Somali gender “polarity”**: heads bearing both number and gender features (Lecarme 2002, Kramer 2015)
- Pre-unraveling interim grammars **always bundle features** (of “different categories”) together
- Unraveling accounts for the possibility of the Somali system - whereas feature (syncretism) economy should prohibit it
Somali nouns famously exhibit so-called “gender polarity” in the plural.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite article: ta/da</td>
<td>woman the naag ta</td>
<td></td>
</tr>
<tr>
<td>masculine definite article: ga/ha</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Somali nouns famously exhibit so-called “gender polarity” in the plural.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite article: ta/da</td>
<td>woman the</td>
<td>naag ta</td>
</tr>
<tr>
<td>masculine definite article: ga/ha</td>
<td>bull the</td>
<td>dibi ga</td>
</tr>
</tbody>
</table>
Somali nouns famously exhibit so-called “gender polarity” in the plural.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite</td>
<td>woman the</td>
<td>woman-PL1 the</td>
</tr>
<tr>
<td>article: ta/da</td>
<td>naag ta</td>
<td>naag-o ha</td>
</tr>
<tr>
<td>masculine definite</td>
<td>bull the</td>
<td></td>
</tr>
<tr>
<td>article: ga/ha</td>
<td>díbi ga</td>
<td></td>
</tr>
</tbody>
</table>
Somali nouns famously exhibit so-called “gender polarity” in the plural.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite article: ta/da</td>
<td>woman the naag ta</td>
<td>bull-PL2 the dibi da</td>
</tr>
<tr>
<td>masculine definite article: ga/ha</td>
<td>bull the dibi ga</td>
<td>woman-PL1 the naag-o ha</td>
</tr>
</tbody>
</table>
However, “polarity” is better analyzed as gender specific to different pluralization strategies.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite article</td>
<td>woman the</td>
<td>naag ta</td>
</tr>
<tr>
<td></td>
<td>woman-PL1 the</td>
<td>naag-ó ha</td>
</tr>
<tr>
<td>masculine definite article</td>
<td>thief the</td>
<td>túug ga</td>
</tr>
<tr>
<td></td>
<td>thief-PL1 the</td>
<td>tuug-ó ga</td>
</tr>
</tbody>
</table>

PL1 imposes masculine gender.
Somali speakers analyzed terminals as bearing both number features and gender features.
However, “polarity” is better analyzed as gender specific to different pluralization strategies.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular form</th>
<th>Plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>feminine definite</td>
<td>bull the</td>
<td>bull-PL2 the</td>
</tr>
<tr>
<td>article: ta/da</td>
<td>díbi ga</td>
<td>díbi da</td>
</tr>
<tr>
<td>masculine definite</td>
<td>thief the</td>
<td>thief-PL2 the</td>
</tr>
<tr>
<td>article: ga/ha</td>
<td>túug ga</td>
<td>túug da</td>
</tr>
</tbody>
</table>
Somali speakers analyzed terminals as bearing both number features and gender features.
These innovations in Somali constitute the introduction of new overt feature syncretisms.

- PL1 [pl, m] -ó
- PL2 [pl, f] Tone shift

- Emerges naturally from the bundles that unraveling approach starts with
- But would constitute the introduction of very uneconomical new feature syncretisms
5. conclusions

Finding innovation in the L1A process
Unraveling approaches to learning morphosyntax link diachrony and acquisition

- The learning model family that best unites actual learning pathways with unidirectional diachronic patterns are **unraveling models** (Pannemann 2007, i.a.), capturing:
  - canonical changes: Dem>Det (Latin)
  - emergence of cross-categorial bundling of features: number-gender bundling (Somali)

- Innovations actuate into languages when **interim analyses** > ultimate analyses
  - Later input-divergent systems, persisting to the age of peer (not parent) social alignment, are most likely actuation candidates (Cournane 2017, cf. Labov 1989, 2001)
We can and must directly study the learning process to assess explanation in diachrony

We can because:

- All children are innovators, and studying any child is fruitful, with our linking hypothesis:

  ⇒ Experimental (Cournane 2014; Cournane & Perez-Leroux, in revision) and modeling (Lee 2019) studies can explore how children’s earlier analyses are actually built and assess diachronic directional hypotheses

We must because child innovators are the explanatory core of our theory
Sources


Cournane, Ailís & Ana Teresa Pérez-Leroux, in revision. “Leaving obligations behind: Epistemic incrementation in preschool English”


Sources


Thank you!

Please reach out!
Naomi Lee (naomilee@nyu.edu) & Ailís Cournane (cournane@nyu.edu)


Thanks to the members of NYU Child Language Lab, and to Maria Gouskova, Richie Kayne, Ruth Kramer, Laurel MacKenzie, Alec Marantz, Ana Teresa Pérez-Leroux, John Singler, and Gary Thoms for feedback and guidance.
appendix
Modal Mappings in L1A

**MaxCat1st Modal Mapping would proceeds as follows:**

**Initial setting:** All verbal expressions are INFL*

- Search for evidence of mapping rules:
  - *may:* \( \text{may}_W \leftrightarrow \text{INFL} \)
  - *going:* \( \text{going}_W \leftrightarrow \text{INFL} \quad \text{going}_W \leftrightarrow v \)
  - *want:* \( \text{want}_W \leftrightarrow \text{INFL} \quad \text{want}_W \leftrightarrow v \quad \text{want}_W \leftrightarrow V \)
But if you miss the target...

Your analysis will be UP; with **BE** and **TO** omitted, V/v-modals look like **INFL-modals**:  

\[
\text{I am going to go}
\]

\[
\text{I \_ going \_ go = I will go}
\]

MaxCat1st: Omission patterns reveal child **representations** that diverge from the input grammars in an upward manner
Case Study: Modal Omissions

“Sarah” (Brown 1973). 2;03 to 5;01. 37,021 utterances

Extracted all utterances with **INFL** or **v** modals, and **V_{want}**: 

- 1215 **INFL**-modals (INFL) *better, can, could, may, might, must, shall, should, will, would*
- 621 functional-modals (**v**) *got, have, ought, be+going, be+supposed*
- 503 instances of premodal (**V**)want (only VP/clausal complements)

Coded all complements

- bare V (*have go*), infinitival V (*have to go*), infinitival_reduced V (*hafta go*), or clausal (*want Daddy go*)
Case Study: Modal Omissions

Results 1

Infinitival to omissions disappear for V-modals, but persist past 5yrs for v-modals. Sarah’s v-modals pattern with input INFL-modals ("I have go")
Case Study: Modal Omissions

Results 2

Figure 3: BE-omissions with v-modals

BE-omissions with v-modals persist longer than other instances of BE-omissions in child language (Brown, 1973; Becker, 2002).
Ongoing computational modeling demonstrates efficacy of an unraveling inductive learner (Lee 2019)
Ongoing computational modeling demonstrates efficacy of an unraveling inductive learner (Lee 2019)