Models of Language Evolution
Session 8: The Iterated Learning Model

Roland Mühlenbernd

2016/12/14
## Project Topics - Frames

<table>
<thead>
<tr>
<th>Topic</th>
<th>Talk (3 CP)</th>
<th>Project (6 CP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>evolution of language structure</td>
<td></td>
<td>[Johannes, Severin]</td>
</tr>
<tr>
<td>language universals (Typology)</td>
<td>Sabrina Tabea</td>
<td>[Chih-Chun, Xiang, Daniel]</td>
</tr>
<tr>
<td>language change (Socioling.)</td>
<td>Max Paul Bastian</td>
<td>[Isabella, Ryan, Deniz]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Katrin, Verena]</td>
</tr>
<tr>
<td>grammatical change (Historical Ling.)</td>
<td></td>
<td>[Teslin, Lisa, Sarah]</td>
</tr>
<tr>
<td>language family reconstruction</td>
<td></td>
<td>[Mei-Shin]</td>
</tr>
</tbody>
</table>
“The most basic principle guiding [language] design is not communicative utility but reproduction – theirs and ours... Languages are social and cultural entities that have evolved with respect to the forces of selection imposed by human users. The structure of a language is under intense selection because in its reproduction from generation to generation, it must pass through a narrow bottleneck: children’s minds.”

(Deacon, 1997: 110)
Language Emergence on 3 Adaptive Systems

Language is a result of three complex adaptive systems:

- biological evolution (phylogeny)
- individual learning (ontogeny)
- language change/cultural evolution (glossogeny)
The Iterated Learning Model

Two forms of representation:

▶ I-language: Internal representation as pattern of neural connectivity (more abstract: grammar)
▶ E-language: External representation as actual sets of utterances (all possible grammatical expressions)

Two forms of interactions:

▶ language use: I-language → E-language
▶ language learning: E-language → I-language
Exercise 1 (Kirby & Hurford 2002)

The Iterated Learning Model has the following four basic components:

- a learning bottleneck
- a homogeneous population structure
- a meaning space ✓
- one or more language-using agents ✓
- a set of stable languages
- a signal space ✓
- one or more language-learning agents ✓
- one or more language-imitating agents
A Simple ILM

Learner as neural network

Meanings
- male/female
- related/unrelated
- older generation
- younger generation

Signals
- p/m
- u/a
- t/d
- a/o

Speaker production: \( s_{desired} = \arg \max_s C(m|s) \)
A Simple ILM

1. initial population: two randomly initializes networks for speaker and hearer each
2. certain number of randomly chosen meanings from 00000000 to 11111111 ($n$ of 256)
3. speaker produces signals for each of this meanings
4. hearer learns by back-propagation error learning (minimizing an error function)
5. remove speaker, hearer becomes speaker, new hearer is added (with random weights)
6. repeat circle
Exercise 2 (Kirby & Hurford 2002)

The simple ILM version (as presented on page 125/126) produces three types of behavior, dependent on the size of the training set. Allocate the following training set sizes to the type of the language that emerges.

A very small learning set
(20 random meanings) \(\iff\) inexpressive and unstable language

A very large learning set
(2000 random meanings) \(\iff\) completely expressive unstructured language

A medium-sized learning set
(50 random meanings) \(\iff\) expressive and highly structured language
A Simple ILM: Results

Structured:
- p/m → +og/-og
- u/a → +yg/-yg
- t/d → rel./unrel.
- a/o → female/male

Unstructured:
- pato → (grand)father/uncle
- muda → (grand)mother/aunt
- pata → young woman

... speaker/learner difference
- proportion of covered meaning space
Compositionality & Recursion

- **Compositionality**: A compositional signaling system is one in which the meaning of a signal is some function of the meaning of the parts of that signal and the way in which they are put together.
- But: sentence-meaning mapping is not only compositional, but recursive.
- Digital infinity: potentially infinite use of finite means by constructing syntactic structures that contain structures of the same type.
- Note: Simple ILM produced compositional (but not recursive) language for the medium-sized learning set.
ILM for Recursive Compositionality

Extended Model

- meaning space: simple variant of predicate logic
- signal space: string of characters
- example: \( \text{loves(mary,john)} \leftrightarrow \text{“marylovesjohn”} \)
- learning method enables the learning/parsing of rules
- production mechanism includes innovation

Results:

\[
(12) \quad \text{gj h f tej m}
\quad \text{John Mary admires}
\quad \text{“Mary admires John”}
\]

\[
(13) \quad \text{gj h f tej wp}
\quad \text{John Mary loves}
\quad \text{“Mary loves John”}
\]

\[
(14) \quad \text{gj qp f tej m}
\quad \text{Gavin Mary admires}
\quad \text{“Mary admires Gavin”}
\]

\[
(15) \quad \text{gj qp f h m}
\quad \text{Gavin John admires}
\quad \text{“John admires Gavin”}
\]

\[
(16) \quad \text{i h u i tej u gj qp f h m}
\quad \text{John knows Mary knows Gavin John admires}
\quad \text{“John knows that Mary knows that John admires Gavin”}
\]

\[
S/p(x,y) \rightarrow gj \quad A/y \neq A/x \quad B/p
\quad S/p(x,q) \rightarrow i \quad A/x \quad D/p \quad S/q
\]
Size & Expressivity of Grammars

![Diagram showing the relationship between size of internal representation (number of words + rules) and expressivity (number of communicable meanings). The diagram includes categories such as Redundant languages, Vocabulary-like languages, and Syntactic languages. The impossible languages are also indicated.]
Frequency and Irregularity

Frequency often correlates with Irregularity:

- top 10 verbs (Engl.): be, have, do, say, make, go, take, come, see, get

Changed Model:

- meaning space: (objects with) two properties $a$ and $b$
- meaning probability defined by index: if $i > j$, then $p(a_i) < p(a_j)$
- signal space: string of characters

Result: example of emerged language:

<table>
<thead>
<tr>
<th></th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_0$</td>
<td>g</td>
<td>s</td>
<td>kf</td>
<td>jf</td>
<td>uhlf</td>
</tr>
<tr>
<td>$b_1$</td>
<td>y</td>
<td>jgi</td>
<td>ki</td>
<td>ji</td>
<td>uhli</td>
</tr>
<tr>
<td>$b_2$</td>
<td>yq</td>
<td>jgq</td>
<td>kq</td>
<td>jq</td>
<td>uhlq</td>
</tr>
<tr>
<td>$b_3$</td>
<td>ybq</td>
<td>jgbq</td>
<td>kbq</td>
<td>jbgq</td>
<td>uhlbgq</td>
</tr>
<tr>
<td>$b_4$</td>
<td>yuqeg</td>
<td>jguqeg</td>
<td>kuqeg</td>
<td>jguqeg</td>
<td>uhluqeg</td>
</tr>
</tbody>
</table>

Note: Hurford (2000) simulates a ILM with a meaning space that combines predicate logic with frequency
Exercise 3 (Kirby & Hurford 2002)

Fill the gaps in the following quote from “The Emergence of Linguistic Structure”:

“In this simulation, an iterated learning model was implemented, with a population of agents starting with no language at all, and over time a language emerged in the community in which there were completely general compositional rules for expressing a range of meanings represented as formulae in predicate logic. A variation on the basic simulation was then implemented in which one particular meaning was used with vastly greater frequency than any of the other available meanings. This inflated frequency held throughout the simulated history of the community. The result was that, as before, a language emerged in the population characterized by general compositional rules, but in addition, all speakers also had one special idiosyncratic stored fact pertaining to the highly frequent meaning.”
Social Transmission favors Generalization

Note:

- the strength of coverage of generalization correlates with the survival potential of meaning-form pairs
- linguistic generalization is favored by social transmission (iterated learning)
- infant drive for internal generalization might be the prime mover in causing regularities in languages
- creatures with no such drive at all would produce no historical E-languages with persistent regular patterns
Exercise 4 (Kirby & Hurford 2002)

Kirby & Hurford discuss the force of generalization in human language evolution and mention that it is important to distinguish between

a) the evolutionary source and

b) the reason for the historical persistence

of generalization. Accordingly, complete the following expressions appropriately:

(a) The evolutionary source of generalization... ...

is the child’s innate capacity to generalize (a)

(b) The reason for the historical persistence of generalization...

is the inherent advantage of general patterns to be propagated across generations (b)
Coevolution in glossogeny and phylogeny

- **phylogeny**: fitness = capacity to communicate
- it is essential for a ‘newcomer’ to adapt to the code of the community
- glossogenetic evolution can put a significant drag on phylogenetic adaptation
- favorable result: an innate ‘universal grammar’ that is useful for each form of human language
- But what is “each form of human language”?
Exercise 5 (Kirby & Hurford 2002)

According to Kirby and Hurford “the primary pressure on the evolution of language…”

- is the necessity for coordination
- is the need to be learnt ✓
- is the fact that communicative skills involve reproductive success
Language Emergence on 3 Adaptive Systems

Language is a result of three complex adaptive systems:

- biological evolution (phylogeny)
- individual learning (ontogeny)
- language change/cultural evolution (glossogeny)
Timescale of Literature

1990 Pinker & Bloom: language evolution theory
1991
1992
1993
1994
1995 Bickerton: PL-fossils in form of language behavior
1996
1997
1998
1999 Jackendoff: PL-fossils, Nowak & Krakauer: The Evolution of Language
2000
2001
2002 Kirby, Hurford: The Emergence of Linguistic Structure
2003
2004 Jäger: Evolutionary Game Theory for Linguists. A primer
2005
2006
2007 Bickerton: perspective from linguistics, Kirby: LE-modelers perspectives
2008 Jäger: Applications of Game Theory in Linguistics
2009
2010
2011
2012
2013
2014 Mühlenbernd & Franke: Meaning, Evolution, and the Structure of Society