A web-based (iterated) language learning paradigm with human participants

T. Florian Jaeger¹, Harry J. Tily², Michael C. Frank³, Jacqueline Gutman¹ and Andrew Watts¹

¹University of Rochester, ²Massachusetts Institute of Technology, ³Stanford University
Artificial Language Learning

• Miniature language taught to participants to study the time course, mechanisms, and biases of (first and second) language acquisition
  [Morgan, Meyer, and Newport, 1987; Hudson Kam and Newport, 2005]

• Exposure may be as short as 2 minutes
  [e.g. for word segmentation tasks; Saffran, Aslin, and Newport, 1996; Frank, Mansinghka Gibson, and Tenenbaum, 2006]

  or be distributed over 4-7 sessions, distributed over 1-2 weeks
  [e.g. for acquisition of syntax; Hudson Kam and Newport, 2005, 2009; Wonnacut, Newport, and Tanenhaus, 2008; Fedzechkina, Jaeger, and Newport, 2011]
Example

<table>
<thead>
<tr>
<th>Case</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOV</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>OSV</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>

[taken from Fedzechkina, Jaeger, and Newport, in prep]
A powerful methodology

- Has provided insights into acquisition of:
  - Lexicon
  - Morphology
  - Syntax
  differences and commonalities between adult and child
  or even infant learners
  biases on acquisition
  - complementing quantitative typology and historical
    linguistics

- But also: a time-consuming methodology
60-80 exposures

60-80 trials

x 4 days

and:
Production data must be scored

[taken from Fedzechkina et al., 2011]
A possible solution: Crowdsourcing

- And in comes the knight in shining armor

[Dr. Hal Tily, enjoying his daily ice cream]
A Flash applet (by Hal Tily)

video + sound + (optionally) written stimuli at top
Noun learning

tobbat

Progress
Learn 1
1 out of 12
Understand 1
Learn 2
Understand 2
Learn 3
Understand 3
Learn 4
Speak 1
Speak 2
Sentence learning
Comprehension test

melnawg bliffen flugit

[Replay] [Continue]

Progress
- Learn 1
- Understand 1
- Learn 2
- Understand 2
  - 3 out of 8
- Learn 3
- Understand 3
- Learn 4
- Speak 1
- Speak 2
Production
Putting the applet and crowdsourcing to the test

- Hudson Kam & Newport (2005): Adults match frequency of alternating determiner in input language

  - Manipulated:
    - Proportion of nouns preceded by determiner
    - Determiner type by noun (arbitrary vs. count/mass)

[taken from Tily et al., 2011]
Experiment 1: Attempt to replicate
[Tily et al., 2011]

- Constructed language contains:
  - 6 animate (3 male, 3 female) and 2 inanimate referents
  - 2 intransitive and 4 transitive actions
- Each subject taught a different randomly generated language, varying:
  - Proportion of nouns preceded by determiner (none, 1/3, 2/3, or all)
  - Determiner type by noun (arbitrary vs. natural gender)
[taken from Tily et al., 2011]
Comparison

• Hudson Kam and Newport (2005)
  – 40 participants
  – Took weeks/months to run the experiment
  – Required lab manager, RAs, and graduate student to schedule and run participants and to score data

• Tily, Frank, and Jaeger (2011):
  – 1-2 days to get 134 participants ($0.50 to $1 per participant)
  – No subject scheduling; no RA time required to run experiment
  – Automatic scoring possible
Experiment 2: Word order acquisition bias

Dryer, 2011
Testing word order universals

- Greenberg's (1963) Universal 1: In a dominant word order, the subject precedes the object

- Greenberg's (1963) Universal 3: Languages with dominant VSO order are always prepositional

- Greenberg's (1963) Universal 4: Languages with normal SOV order are usually postpositional
Design

• 12 languages, randomly assigned to participants

• Basic word order (6):
  – SOV / SVO / VSO / OSV / OVS / OSV

  crossed with

• Determiner-noun order (2):
  – Det N / N Det
• Participants: 285 in XXX days.
Argument x Determiner Order
[taken from Tily et al., 2011]
Pushing onward
[Gutman, 2011 with help from Watts]

• **Iterative** artificial language learning (IALL)
  [e.g. Kirby et al., 2008; Smith and Wonnacut, 2010]
Bias against unnecessary complexity?

• Recent information theoretically inspired work on morphological paradigms suggests that there is a bias against high conditional entropy.

• Very recent IALL work has provided evidence that learners tend to reduce conditional entropy (i.e. highly unpredictable variability) over generations: e.g. plural markers [Smith and Wonnacut, 2010]

• Q: Are learners biased against high conditional entropy in case-marking systems?
Input language to generation 1: $L_0$

- Lexicon:
  - 6 nouns (4 humans, 2 inanimates)
  - 4 verbs (transitive)
- SOV-biased: 63% SOV, 37% OSV
- Object always case-marked
  - 2 Case-markers: $kah$ vs. $zub$
  - Equally frequent and unconditional on context: maximum conditional entropy
Conditional Entropy of Case Marker given Action

Conditional Entropy of Case Marker given Agent

Conditional Entropy of Case Marker given Theme

Conditional Entropy of Case Marker given Theme Animacy and Gender

[taken from Gutman, 2011]
Comparison

  - 40 participants, 10 in each chain
  - It takes about 3-12 months to run studies like this

- Gutman (2011):
  - 40 participants, 10 in each chain ($2.50 each)
  - 1-2 weeks
  - Scoring mostly automatic
Conclusion

- **ALL** and IALL can be conducted over the web
  - Reducing time necessary to run these studies
  - Reducing costs

- Expanding the paradigm

- Current limitation:
  - No sound recording, but we’re working on that
Thanks!
If number correct is ≤ 10 out of 12, then loop

**Learn 1**
Lexical Learning
12 trials [6 nouns]

**Understand 1**
Lexical Discrimination
12 trials

**Learn 2**
Sentence Learning
8 trials
[1 verb, 2 agents, 6 themes]

**Understand 2**
Sentence Discrimination
8 trials

**Learn 3**
Sentence Learning
8 trials
[1 verb, 2 agents, 6 themes]

**Understand 3**
Sentence Discrimination
8 trials

**Learn 4**
Sentence Learning
24 trials
[4 verbs, 4 agents, 6 themes]

**Speak 1**
Sentence Production
20 trials

**Speak 2**
Sentence Production
20 trials