Learning grammar: what can
development in different languages tell us?

Elena Lieven
ESRC International Centre for Language and
Communicative Development
(LuCiD)
University of Manchester
Road map

• The constructivist/emergentist approach
  – Uneven development
  – Lexical effects

• Morphology

• Implications for studying non-WEIRD language learning
Dalabon, Arnhem Land, Australia

“Wekemarnûmolkkûndokan”

We -ke -marnû - molkkûn
APPRehensive -3du.DISHARMONIC -BENEFACTIVE - unbeknownst
- doka -n
- go[respect.form] -NonPast

“I’m afraid that the two of them, who are in odd-numbered generations with respect to one another, might be sneaking around (i.e. going around unbeknownst to someone who should know); by choosing the form of words that I do, I hereby indicate that one of those I am referring to is a mother-in law’s brother or comparable relative”

Adapted from N.Evans, 2009
The usage-based approach [emergentist, constructivist]

Children learn holophrases: (individual words but also unanalysed strings of words or morphemes:

- I’m hitting it
- I’m kicking it
- I’m eating it
- Where’s Mummy
- Where’s Daddy
- More juice
- More biscuit

Low-scope schemas with slots are formed:

- I’m X-ing it?
- Where’s Y?
- More Z?

Slots have properties of items that appeared in same position in source utterances
→ gain phonological, semantic, syntactic properties → more abstract grammar

The greater the overlap between slot and potential filler
- the more easily the resulting utterance is comprehended and produced
- the greater the extent to which it is grammatically ‘correct’
Levels of analysis

• We start by counting at the level of specific form and string:
  – is/are
  – I’m X-ing/You’re Y-ing
  – What do X?/What can X?

• We only count at a more abstract level, when there is evidence for it

• We do not credit the child with pre-given, abstract linguistic categories from the outset
Abstraction and Productivity

Abstraction

• Children abstract from the beginning
• A concept is an abstraction, so is a word
• The issue is not whether children abstract but
  – What do they abstract over?
  – How specifically linguistic is the nature of the abstraction process?

Productivity

• Slot-and-frame patterns can be highly productive
• But the scope of the patterns increases with development
Abstraction  √

Productivity  √

But predictions are:

- Children will be less productive than adults for the same construction and the same forms
- Children will show uneven development for the same construction
• To say that the child knows that \textit{l goes} sounds wrong does not mean that ‘have agreement’

→ this is an underlying assumption
Spanish verb inflections

Nottingham corpus

• Lucia: 22 hours: 2;2.25 – 2;7.14
• Juan: 31 hours: 1;1-.21 – 2;5.28
Rates of Subject-Verb Agreement Error in Juan and Lucía’s present tense verbs data broken down by Person and Number

<table>
<thead>
<tr>
<th>Inflection</th>
<th>Contexts</th>
<th>Error Rate</th>
<th>Contexts</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>3152</td>
<td>4.5%</td>
<td>1672</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

Juan: 2;2.25 - 2;7.14
Lucía 1;1.21 - 2;5.28
<table>
<thead>
<tr>
<th>Inflection</th>
<th>Juan: 2;2.25 - 2;7.14</th>
<th>Lucía 1;1.21 - 2;5.28</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contexts</td>
<td>Error Rate</td>
</tr>
<tr>
<td>Overall</td>
<td>3152</td>
<td>4.5%</td>
</tr>
<tr>
<td>1sg</td>
<td>693</td>
<td>4.9%</td>
</tr>
<tr>
<td>2sg</td>
<td>147</td>
<td>10.2%</td>
</tr>
<tr>
<td>3sg</td>
<td>1997</td>
<td>0.7%</td>
</tr>
<tr>
<td>1pl</td>
<td>61</td>
<td>0</td>
</tr>
<tr>
<td>3pl</td>
<td>251</td>
<td>31.5%</td>
</tr>
</tbody>
</table>
Frequencies of verbs in person-number forms
Rates of Subject-Verb Agreement Error in Juan’s Data as a function of Lexical Frequency

First Person Singular
Second Person Singular
Third Person Plural
Low overall error rate reflects

– Children’s knowledge of a relatively small number of high frequency forms
– Children’s use of most frequent inflection (3rd person singular) as a ‘default’ when they don’t know what to do

So development is uneven and the ability to generalise (and show abstraction) builds over development
Uneven development

The transitive: Who does what to whom

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb&lt;sub&gt;transitive&lt;/sub&gt;</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent</td>
<td>Action</td>
<td>Patient</td>
</tr>
</tbody>
</table>

Dog bites man  
Man bites dog

We know that children aged 21 months can find the correct match between two scenes with agent and patient reversed
The monkey is tamming the frog. Find tamming!
The frog is weefing the monkey. Find weefing!
But this effect is fragile and development is uneven despite the same abstract argument structure.

Success is developmentally later:
- With non-prototypical transitives e.g. with an inanimate subject,
- If the two actions are identical with roles reversed

Young children treat the first noun as agent in conjoined subject intransitives:

The boy and the girl are gorping → The boy is gorping the girl

In elicited production, comprehension, weird word order correction, and priming, studies:
Children do significantly better with pronoun frames:

*He’s meeking it, It got X-ed by it*

And the language matters:

Children’s reliance on the first noun as Subject:
English > German > Chinese
Lexical effects

• Optional infinitives in Dutch, German and Spanish
  – Learning model trained on CDS can account for different rates AND for which verbs appear as OIs
    [Freudenthal et al., 2007]

• Accusative for Nominative errors in English
  – Children whose caregivers produce more complex sentences containing me+verb (Let me do it, Watch me dancing), produce more errors (Me do it, Me dancing). Specific verbs in CDS constructions more likely to occur in children’s errors
    [Kirjavainen, Theakston & Lieven, 2009]

• Long distance wh-questions [What, do you think [t, is in the box]]
  – Imitation experiment (5;0 – 6;0)
    \[\text{vv} \text{What do you think the old man really hopes?}\]
    \[\text{XX} \text{What does the old man really hope you think?}\]
    [Dąbrowska, Rowland & Theakston 2009]
Learning inflectional morphology
Previous studies

• Not many systematic experimental studies on inflection in morphologically complex languages

  – “Most of the children studied acquired those rules quickly and made no consistent errors.”
  – “…the full system is productive in many children before 2 years-of-age.”

  Smoczyńska (1985)
**Morphology: Productivity develops**

- **Polish**
- **Finnish**
- **Spanish**

**Compare child and adult on same stems and inflections:**
Young children use significantly fewer inflections per stem
Krajewski, Lieven & Theakston, 2012
Aguado-Orea and Pine, 2015

- **Chintang**

**Entropy measures:**
Children approach adult CDS levels around 3;0
Stoll, Mazara & Bickel, 2015

- **9 languages**

**Mean size of paradigm in child-directed speech:**
Correlated with speed of development of paradigms
Xanthos et al., 2011
Polish, Finnish and Estonian

• Morphologically complex languages with rich verbal and nominal inflection paradigms

• Comparing languages with very different inflectional systems

Joanna Kolak  Sonia Granlund  Virve Vihman
Questions

1. Accuracy
• How does the frequency of the form in the child’s input influence his/her performance?
• How does the number of phonological neighbours of a given form affect children’s performance?

2. Errors
• What do the errors that children make tell us about the process of learning?
POLISH VERB INFLECTION

- Present tense
- 6 person/number contexts
- 4 conjugation classes
- Multiple subgroups based phonological properties of the stem

<table>
<thead>
<tr>
<th>CLASS</th>
<th>GROUP</th>
<th>TYPE FREQ (Dictionary)</th>
<th>INFINITIVE</th>
<th>1ST &amp; 2ND PERS SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>4</td>
<td>25.76%</td>
<td>malować</td>
<td>maluję, malujesz</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>5.84%</td>
<td>mazać</td>
<td>możę, mążesz (alveolar or post-alveolar fricative or affricate in infinitive becomes palatalised)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.57%</td>
<td>ciągnąć</td>
<td>ciągnę, ciągniesz</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>&lt;5%</td>
<td>tanieć</td>
<td>tanieję, taniejesz</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>&lt;5%</td>
<td>like 4, but different stem in Past Tense</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10a,b</td>
<td>&lt;5%</td>
<td>pić</td>
<td>piję, pijesz</td>
</tr>
<tr>
<td></td>
<td>10c</td>
<td>&lt;5%</td>
<td>like 5, but different stem in Past Tense</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>&lt;5%</td>
<td>wieźć</td>
<td>wiezę, wieziesz</td>
</tr>
<tr>
<td>II</td>
<td>6</td>
<td>22.86%</td>
<td>robić, wierzyć</td>
<td>robię, robisz/ wierzę, wierząz</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.59%</td>
<td>widzieć</td>
<td>widzę, widzisz (palatal fricative or affricate in Infinitive is fronted in some forms)</td>
</tr>
<tr>
<td>III</td>
<td>1</td>
<td>25.72%</td>
<td>czytać</td>
<td>czytam, czytasz</td>
</tr>
<tr>
<td>IV</td>
<td>2</td>
<td>0.04%</td>
<td>umieć</td>
<td>umiem, umiesz</td>
</tr>
</tbody>
</table>
Verbs: Design

• Children see videos of different people doing different activities

• They have to produce an inflected form of the verb, e.g.
  “She sings”, “They dance”

• Tested 80 3- & 4-yos in each language
Can the acquisition of verb inflection be simulated by a single mechanism for two morphologically complex and dissimilar languages?

- 3-layer neural networks trained on CDS
- Corpora of person- and number inflected present tense forms
- Finnish 1787 forms from 802 verbs; Polish 3175 forms from 1306 verbs
- Phoneme representations of verbs stems presented in a right justified syllable template together with a code for one target person/number context
- Network trained to output the correct form
Questions (and answers!)

1. Is it possible to acquire the inflection system just by example? **Yes.**

2. Are the errors affected by the same factors as for the kids (form frequency and phonological neighbourhood density)? **Yes.**
In addition to this:

3. We can see the full learning trajectory for different parts of the paradigm.

4. We can selectively switch off differences in frequency or in neighbourhood size.
MODELLING — GENERALISATION

From 51 verbs, one form each was withheld from training

Different generalisation by subgroup

- I.11: small & stem change
- I.4: Large and regular
- I.9
- II.6
- II.7
- III.1

Proportion correct vs. Epoch
Modelling - summary

• Error types, frequency effects and generalisation performance differ across morphophonological subgroups and P/N contexts
• More errors for low frequency P/N contexts
• Suffix overgeneralisation
• Substitution of high-frequency for low-frequency forms
Nouns in Finnish, Polish and Estonian

Affixes
• mostly regular in FINNISH and ESTONIAN
• complex and differ across classes in POLISH
• FINNISH more agglutinative (one form=one meaning)
• POLISH more fusional

Stem changes
• more important in EST than FIN
• phonologically conditioned in FIN and therefore more predictable

Distribution of declension classes
FINNISH: 75% of all nouns in first class
ESTONIAN and POLISH: nouns more evenly distributed across classes
Nouns: Design

• Children see the characters interacting with different objects in different case contexts

• They hear the beginning of the context sentence and have to complete it with the correctly inflected noun

• Tested 45 3- & 4-yo children in each language
Elicited production of case marking

This is a ball.
The fox does not see the ball.
The fox takes the ball.
The fox lives in the ball.
The fox is coming out of the ball.
The fox is going into the ball.
The fox is waving at the ball.

NOM: Tämä on pallo.
PAR: Kettu ei näe palloa.
GEN: Kettu ottaa pallon.
INE: Kettu asuu pallossa.
ELA: Kettu tulee ulos pallosta.
ILL: Kettu menee sisään palloon.
ALL: Kettu vilkuttaa palollole.
Summary of results

• Error rates: 8% (verbs) and 12-18% (nouns)

• In both studies, across languages:
  – children were more accurate with forms that occur more frequently in the input
  – the more phonological neighbours a given word had, the easier it was for children to learn

• For different languages, there were also other factors influencing children’s performance (distinctiveness of the inflectional ending, animacy, morphological complexity)
Chintang, Tibeto-Burman, East Nepal

\[ yo=go=i=yã \quad luŋghek \quad thitta \]

DEM.ACROSS=NMLZ=LOC=ABL stone one

\[ kob-u-thapt-o-kh-o! \]
pick.up-3O-V2:move.across.TR-3O-V2:see[-IMP.2sA]-3[s]O

‘Bring over one stone from there!’

[CLLDCh4R02S02b.412] \approx 3;4

Stoll, Mazara & Bickel, 2015
Extending this work to non-WEIRD cultures and languages?

Difficulties:

• Conducting experiments
• Collecting corpora
• Depth of semantics and pragmatics

There’s a lot we can do within WEIRD cultures
So why bother?

• How does the universal become particular?
  – Early communication and intention reading
  – Building on primate cognition: from cognition to semantics

• What is the range of challenges set by different languages?
  – Important to select languages (e.g. Stoll: ACQDIV)

• What is the range of challenges set by different contexts?
Particular thanks to …

The LuCiD Directors

- Julian Pine  
  *U of Liverpool*

- Anna Theakston  
  *U of Manchester*

- Caroline Rowland  
  *U of Liverpool*  
  *Max Plank-Nijmegen*

- Padraic Monaghan  
  *Lancaster University*

Wonderful colleagues

- Michael Tomasello
- Ben Ambridge
- Danielle Matthews
- Minna Kirjavainen
- Sabine Stoll
- Kirsten Abbot-Smith
- Thea Cameron-Faulkner
- Silke Brandt
- Alissa Ferry
- Bahar Köymen

Our funder

- ESRC Economic & Social Research Council

Grant number ES/L008955/1

Cristia.cnrs.may2017
Find out more about LuCiD:

www.lucid.ac.uk

info@lucid.ac.uk

lucidcentre

lucid_centre
The end

Thank you!