Children’s production of determiners: a test case for innate syntactic categories?

Catriona Silvey
Christos Christodouloupolous
Evolutionary question

• What do humans bring to the task of acquiring language?

• What can we learn from children’s early spontaneous productions?

• Ongoing debate: children’s production of determiner+noun combinations
What do humans bring to the task of acquiring language?

DP \rightarrow \text{DET} + \text{N}

the cat
the baby

Valian (1986); Pine & Lieven (1997); Valian, Solt & Stewart (2009); Yang (2013); Pine et al. (2013)
What do humans bring to the task of acquiring language?

Valian (1986); Pine & Lieven (1997); Valian, Solt & Stewart (2009); Yang (2013); Pine et al. (2013)
Overlap

\[
\frac{\text{# of nouns used with both the and a}}{\text{# of nouns used with either the or a}}
\]

Pine & Lieven (1997)
Overlap

# of nouns used with both *the* and *a*  

# of nouns used with either *the* or *a*  

- a baby
- the baby
- a cat
- a cat
- a cat
- the ball
- the ball
- a ball

Pine & Lieven (1997)
Overlap

\[
\frac{\text{\# of nouns used with both \textit{the} and \textit{a}}}{\text{\# of nouns used with either \textit{the} or \textit{a}}}
\]

- a baby
- the baby
- a cat
- a cat
- a cat
- the ball
- the ball
- a ball

Pine & Lieven (1997)
Overlap

\[
\frac{\# \text{ of nouns used with both } \text{the} \text{ and } a}{\# \text{ of nouns used with either } \text{the} \text{ or } a}
\]

a baby
the baby

a cat
a cat
a cat

the ball
the ball
a ball

\[\frac{2}{3} = 67\%\]

Pine & Lieven (1997)
Previous findings

• Pine et al. (1997, 2013)
  – child overlap initially lower than parent overlap
  – children gradually abstract syntactic categories

• Valian et al. (2009, 2014)
  – child overlap no different from parent overlap
  – children have innate syntactic categories
Yang (2013)

DP $\rightarrow$ DET + N

GIVE DRINK
GIVE BALL
MORE DRINK
Yang (2013)

DP $\rightarrow$ DET + N

GIVE DRINK
GIVE BALL
MORE DRINK
Yang (2013)

- Overlap depends on noun frequency

  a ball  a flower
  the ball a flower
  the ball
  the ball
  a ball
  a ball
  the ball
  the ball

- Determiners & nouns should freely combine within frequency constraints
Yang (2013)’s model

• Predicts overlap from 3 main parameters:
  – Zipfian probability of each noun
  – Zipfian probability of each determiner
  – Sample size (number of det+noun pairs)

• Predicted & empirical overlap values for 6 children (1;1-5;1) from CHILDES

MacWhinney (2000); Zipf (1949)
Yang (2013) results

![Graph showing empirical values vs model predictions for different language data sets (child language, early child language, and Brown Corpus).]
Child data: replication
Free combinations

Data

a baby
the baby

a cat
a cat
a cat

the ball
the ball
a ball
Free combinations

Data

- a
- the  baby  baby
- a  cat
- a  cat
- a  cat
- the  ball  ball
- the  ball  ball

Frequencies

- a  5
- the  3
- baby  2
- cat  3
- ball  3
Free combinations

Data

- baby
- cat
- ball

Shuffled Data

- baby
- cat
- ball

Frequencies

- a: 5
- the: 3
- baby: 2
- cat: 3
- ball: 3
Child data: free combinations
Child data: implications

• Model underestimates overlap under freely combinatorial rule

• This holds for simulated Zipfian samples

• Why does the model fit the real data?
Real data do not combine freely

a cookie
a cookie
a cookie
a cookie

the door
the door
the door
the door
the door
the door
the door
the door

Lana Dandan (Flickr)
Frederik Ranninger (Flickr)
Should a generativist theory predict free combinations?

• Children don’t freely combine determiners and nouns

• And they shouldn’t!

• Regularities in discourse context constrain combinations beyond marginal frequencies
Should a constructivist theory predict early constrained combinations?

- Children produce nouns alone before det +noun combinations (Clark, 2003)

- Evidence from input that ‘the’ and ‘a’ can combine with many nouns

- By the time children produce combinations, ample evidence from which to construct a rule
Other approaches

• Bayesian modelling (Meylan, Frank & Levy, 2013)

• Denser sampling e.g. Human Speechome Project (Roy et al., 2006)

• Experimental studies (Maratsos, 1974; Warden, 1976; Karmiloff-Smith, 1979)
More broadly

• Taking a broader perspective

• Determiners + nouns
  – Historical change
  – Invention without input
On the cultural level

- Rules change during transmission and interaction

Greenberg (1978); Beckner & Bybee (2009); De Mulder & Carlier (2012); Smith, Fehér & Ritt (2014)
On the individual level

• A learner with no input still generates rules

[RGENCY point at penny] point at self

[penny that] me

‘(You) (give) me that penny.’

• Only after abstract noun category appears

Goldin-Meadow (2003); Hunsicker & Goldin-Meadow (2012)
Back to our original question

• What do humans bring to the task of acquiring language?
  – Propensity to infer (or create) combinatorial rules
  – Rules emerge via individual & historical reanalysis

• What can we learn from children’s early spontaneous productions?
  – Different theories may not make different predictions
  – Combinations will be semantically constrained
Thanks!

Thanks to Charles Yang for his assistance with our replication.
Funding: NIH P01HD40605, R01-HD054448

Code available at:
https://github.com/christos-c/noun-det-diversity

References


Pictures: Lana Dandan [https://www.flickr.com/photos/lanadandan/346204320/](https://www.flickr.com/photos/lanadandan/346204320/) and Frederik Ranninger [https://www.flickr.com/photos/130218015@N02/16461029851/](https://www.flickr.com/photos/130218015@N02/16461029851/)
Yang (2013) model

\[ E_r = 1 - (1 - p_r)^S - \sum_{i=1}^{D} \left[ (d_i p_r + 1 - p_r)^S - (1 - p_r)^S \right] \]

Probability noun not sampled

Probability noun sampled exclusively with one determiner
Simplified model

\[ E_r = 1 - \sum_{i=1}^{D} d_i^{(p_r,S)} \]

Probability noun sampled exclusively with one determiner
Brown (1973)

- World knowledge
- Knowledge of what others know
- Understanding of connected discourse
- Part-whole entailment
- Fictitious/hypothetical reference
SARAH: I want to open the door.
MOTHER: what door?

GLORIA: he’s going on the fox’s tail.
...
EVE: he on a fox’s nose.
Free combination is not the goal

go to the kitchen and get me a cookie.

I’m going to have a bath.

answer the phone!
Free combination is not the goal

go to a kitchen and get me the cookie.

I’m going to have the bath.

answer a phone!
Nim

2-sign combinations

MORE Nim
MORE Nim
MORE Nim

GIVE drink
GIVE drink
MORE drink
GIVE drink
MORE drink

Terrace (1979)
Nim: replication
Nim: free combinations

![Graph showing expected overlap versus empirical overlap with data points and trend line.]

- Real data
- Free combinations

Expected overlap

Empirical overlap
Nim: implications

• Model overestimates overlap under freely combinatorial rule

• Nim’s sample not strictly Zipfian – low-ranked signs less frequent than predicted

• Nim data and child data not comparable using this model
Children

Peter 1;9.08 - 3;1.20
Adam 2;3.04 - 5;2.12
Sarah 2;3.05 - 5;1.06
Eve 1;6 - 2;3
Naomi 1;2.29 - 4;9.03
Nina 1;11.16 - 3;3.21