Coordination and Learning in Human Dialogue

Raquel Fernández

Institute for Logic, Language & Computation
University of Amsterdam

LDLMD @ NIPS 2016

Language learning through interaction

Focus today: child language acquisition

$\hookrightarrow$ asymmetric setting: agents with different linguistic abilities

Outline:

- multimodal word learning from child-directed input
- coordination in child-adult dialogue
- impact of corrective feedback on child language learning

Motivation:

- relevant to computational models of learning in dialogue
- computational models as tools for gaining further insights

How children learn word meaning

```
MOT: here's the pig look at
   the piggie oink oink
```

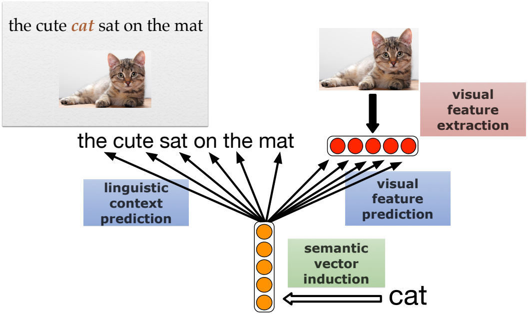
[Rollins corpus, CHILDES database]

- before production
- situated environment
- incrementally
- from few exposures

Multimodal semantic learning

Model that operates on images and child-directed speech to induce word-object associations (Lazaridou, Chrupała, Fernández & Baroni, NAACL-2016)

- based on the multimodal skip-gram model
  (Lazaridou, Pham & Baroni, NAACL-2015)



- trained on very limited amount of child-directed linguistic input
- referential uncertainty: multiple words co-occur with multiple objects

The Frank et al. corpus

(Frank, Goodman, & Tenenbaum, NIPS-2008)

http://langcog.stanford.edu/materials/nipsmaterials.html

```
*mot: let me have that
%ref: RING
*mot: ahhah what's this
%ref: RING HAT
*mot: what does mom look like with the hat on
%ref: RING HAT
*mot: do i look pretty good with the hat on
%ref: RING HAT
*mot: hmm
```

- 2 transcribed dialogues, 10 minutes each, 2,533 words in total
- visible objects manually annotated with arbitrary symbolic labels
- gold-standard lexicon: 36 words paired with 17 object labels

The Frank et al. corpus: Our version

(Lazaridou, Chrupała, Fernández & Baroni, NAACL-2016)

let me have that 

ahhah whats this  

what does mom look like with the hat on  

do i look pretty good with the hat on  

hmm  

Object identification after a single exposure

word	*gold object*	*17 objects*		*5.1K objects*	
		nearest	*r*	*nearest*	*r*
bunny	bunny	bunny	**1**	bunny	**1**
cows	cow	cow	**1**	lea	**7**
duck	duck	duck	**1**	mallard	**4**
duckie	duck	duck	**1**	mallard	**3**
kitty	kitty	book	**2**	bookcase	**66**
lambie	lamb	lamb	**1**	lamb	**1**
moocows	cow	cow	**1**	ranch	**4**
rattle	rattle	rattle	**1**	rattle	**1**

- the model associates words with relevant visual concepts
- like children, the model can get word meaning right based on a single exposure to the word in a situated context

Linguistic interaction in child-adult dialogue

input vs. interaction

sensitivity to statistical regularities in the child-directed input

sensitivity to when & how the input if offered in interaction

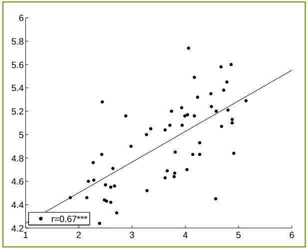
```
Adult: Help me put your toys away, darling.
Child: I'm going to Colin's and I need some toys.
Adult: You don't need a lot of toys.
Child: Only a little bit toys.
Adult: You only need a few.
Child: Yes, a few toys.
```

Interest in studying patterns of local coordination in interaction and their effect on learning [now abstracting away from embodiment]

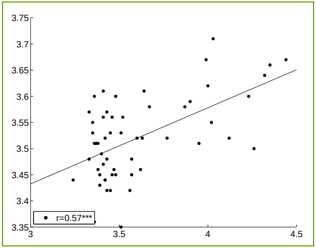
Linguistic interaction in child-adult dialogue

Coordination: there are robust correlations between the complexity of the adult's and the child's speech.

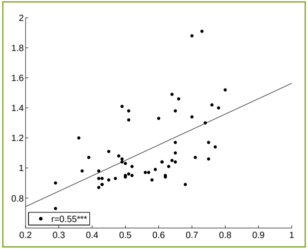
utterance length



word length



word types



379 child-adult dialogues from 3 children over a period of ∼3 years.

(Kunert, Fernández & Zuidema, SemDial-2011)

Turn-based cross-recurrence quantification analysis

Fernández & Grimm (CogSci-2014), inspired by Dale & Spivey's (2006) use of RQA

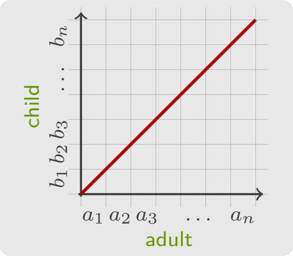
Two-party dialogue transcript

A_1: which one do you want first
B_1: that one
A_2: you like this one
B_2: yeah, give me
⋮
A_n: ...
B_n: ...

Recurrence score for each (i, j)
(lexical, syntactic,...)

$\Rightarrow$

Cross-recurrence plot: each cell corresponds to a pair of turns (i, j)

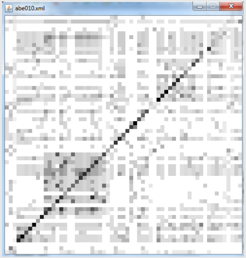


- global recurrence: average coordination over all turn pairs
- local recurrence: recurrence in (semi-)adjacent turns, separated by at most distance $d < n$ (diagonal line of incidence: $d = 1$)

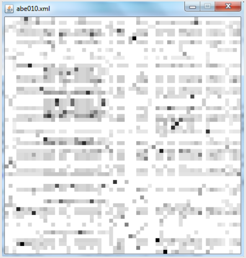
Cross-recurrence plot

dialogue with Abe (2.5 years old)

original dialogue



order of turns shuffled



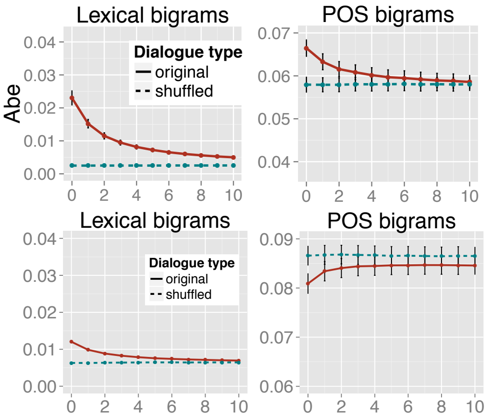
Clear pattern of local lexical recurrence when the temporal development of the interaction is preserved.

Child-adult vs. adult-adult dialogue

child-adult

adult-adult

(Switchboard)



- significantly more local coordination
- both coordinate more at local levels, but the adult recurs with the child significantly more (not shown by graphs)
- different coordination patterns in adults: local syntactic divergence

Negative input?

Local syntactic recurrence may be related to reformulations or recasts (Brown, 1973; Chouinard & Clark, 2003; Saxton et al., 2005)

```
Child:  you're good to sharing.
Mother:    I'm good at sharing?
```

Large scale data-driven analysis to test the influence of corrective feedback on language learning (Hiller & Fernández, CoNLL-2016)

Definition: Child-adult utterance pair meeting all these constraints

1. The child's utterance contains a grammatical anomaly.
2. There is some overlap between the adult and child utterances.
3. There is some contrast: the adult's utterance is not a mere repetition.
4. This contrast offers a correct counterpart of the child's erroneous form.

Corpus Study

4 children, 4-6 transcripts per child, 2,627 candidate CF exchanges
(child-adult utterance pairs with partial overlap)

subject, omission:

```
CHI: don't want to.
MOT: you don't want to?
```

irregular past, substitution:

```
CHI: he falled out and bumped his head.
MOT: he fell out and bumped his head.
```

auxiliary verb, addition:

```
CHI: I'm read it.
DAD: you read it to mummy.
```

Focus: subject omission errors (SOE)

	Om	*Add*	*Sub*	**Total**
Syntax				
subject	**171**	–	1	172
verb	90	1	–	91
object	13	–	–	13
N morph				
poss -'s	4	1	–	5
regular pl	–	3	–	3
irregular pl	–	–	3	3
V morph				
3rd person	4	–	–	4
regular past	10	1	–	11
irregular past	1	–	4	5
Unb. morph				
det	79	–	6	85
prep	21	1	12	34
aux verb	114	5	1	120
progressive	9	0	0	9
Other	4	2	19	25
Total	520	14	46	580

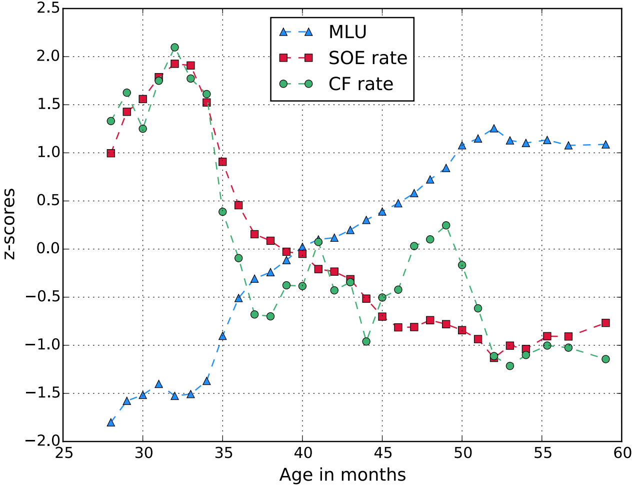
Automatic detection

High-precision automatic classifiers for SOE and CF on SOE to enable an analysis of the entire dataset:

- 25 children, 1,683 transcripts, 1,598,838 utterances, 136,152 candidate CF
- manually annotated data for training (2,627 candidate CF)
- 5-fold cross validation for feature tuning

Detection of	Classifier	Precision	Recall	Total #
SOE	rule-based	0.83	0.8	**287,309**
CF on SOE	SVM	0.89	0.36	**31,080**

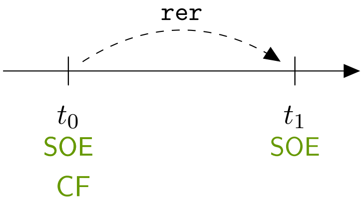
Adam, Brown corpus



MLU: mean length of utterance in words
SOE: subject omission errors
CF: corrective feedback on subject omission errors

Corrective feedback and learning

Relative error reduction (`rer`) of subject omission errors:



$$\texttt{rer}(t_0, t_1) = \frac{SOE_{t_0} - SOE_{t_1}}{SOE_{t_0}}$$

control variables

- child age
- child / adult MLU
- child / adult vocabulary size
- adult subject omissions
- proportion of child speech

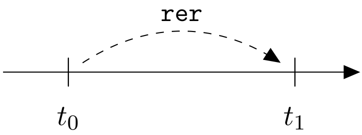
Linear regression models

- with `rer` as dependent variable
- including / excluding CF

3 experimental settings

- t_0: starting age
- $d(t_0, t_1)$: time lag

Results

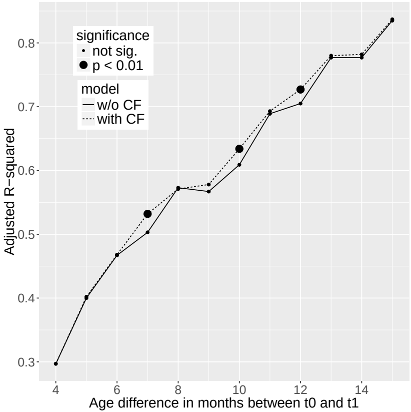


Setting 1: any t_0 and any $d(t_0, t_1) \geq 1$ month

- Positive correlation between CF_{t_0} and $\texttt{rer}(\texttt{t}_0, \texttt{t}_1)$
  $r = 0.29, p < 0.001$
- Linear regression model: CF explains a significant proportion of `rer`, independently of other predictors

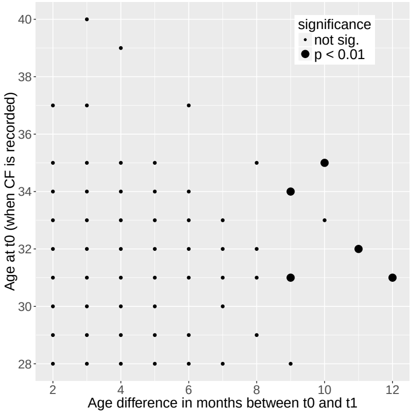
Results

Setting 2: any t_0 and fixed $d(t_0, t_1)$



CF has an impact after a time lag of 7–12 months. . .

Setting 3: fixed t_0 and fixed $d(t_0, t_1)$



. . . for all starting ages for which there is data available.

Concluding remarks

Language learning through interaction

- incremental multimodal word learning from few exposures
- coordination in child-adult dialogue
- local interaction can function as negative input:
  - ↪ corrective feedback contributes to learning of subject inclusion in English, after a lag of at least 7–9 months

How can we model this interactive process for automated learners?

- learning to communicate with each other
- exploiting implicit supervision
- in a situated environment

[work in progress with Afra Alishahi, Grzegorz Chrupała & Lieke Gerderloos]

thank you

Collaborators & students: Marco Baroni, Grzegorz Chrupała, Robert Grimm, Sarah Hiller, Richard Kunert, Angeliki Lazaridou, Willem Zuidema