Computational Semantics and Pragmatics

Raquel Fernández

Institute for Logic, Language & Computation
University of Amsterdam

Autumn 2016

Outline

- timing coordination – turn taking
- meaning coordination – dialogue acts
- meaning coordination – grounding
- style coordination - alignment and adaptation
- language *acquisition* in interaction

Outline

Today:

- Main theories of first language acquisition.
  - Nativist
  - Empiricist
  - Interactive
- Interaction view: two examples of my own work:
  - language coordination in child-adult interaction
  - corrective feedback

Next Tuesday:

Discussion of a recent paper on language learning in artificial agents:

> Wang, Liang & Manning. *ACL 2016.*
> Learning Language Games through Interaction

The nativist view

Knowledge of grammar is innate, in the form of a Universal Grammar that is the initial state of the language faculty.

> "Language learning is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation"
>
> (Chomsky 1993, p. 519)

Main motivation:

- Acquisition is fast and easy,
- in spite of inadequate input (poverty of stimulus),
- and happens without direct instruction (no negative evidence).

None of these claims is well supported empirically.

The nativist view: counter evidence

- *Fast?*
  Children are exposed to language around 10 hours per day (millions of words/sentence in the first 5 years).
- *Easy?*
  Children go through learning stages and make errors over several years (meaning extension, morphological regularisation, word order).
- *Poor input?*
  Child-directed speech is simpler, clearer, and more well formed than adult-adult speech.
- *No negative evidence?*
  Typically no explicit correction, but plenty of implicit feedback (more later).

The empiricist vs. interaction views

input *vs.* interaction

sensitivity to statistical regularities in the input ignoring interaction

sensitivity to when & how the input is 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.
```

---

child $\longrightarrow$ adult *language learning*

child $\longleftarrow$ adult *child-directed speech*

The interactive view

"Relevant input" — *joint attention, engagement, topic continuity, contingent replies* ... — has been shown to be a positive predictor of language development (Tamis-LeMonda et al. 2001; Hoff & Naigles, 2002; Rollins, 2003; Mazur et al. 2005; Hoff, 2006; a.o.)

McGillion et al. (2013): what sort of responsiveness matters?

- *semantic responsiveness*: related to the child's focus of attentions
- *temporal responsiveness*: temporally contingent with an act produced by the child.

$\rightsquigarrow$ *combined measure only significant predictor of vocabulary growth*

Open question: use computational modelling to investigate how these aspects relate to the learning mechanisms employed by the child – and what this can tell us about theories of dialogue.

Examples today: recent work on methodologies for studying *interaction* and *contingent responsiveness* in corpus data.

Two examples of concrete work

Ways of investigating how speakers pick up on each other's language (*coordinate*) at different degrees of locality.

R. Fernández & R. Grimm. Quantifying Categorical and Conceptual Convergence in Child-Adult Dialogue, *36th Annual Conference of the Cognitive Science Society*. 2014.

Empirical study on impact of one particular interactive phenomenon on learning:

S. Hiller & R. Fernández (2016) A Data-driven Investigation of Corrective Feedback on Subject Omission Errors in First Language Acquisition. In *Proceedings of CoNLL*.

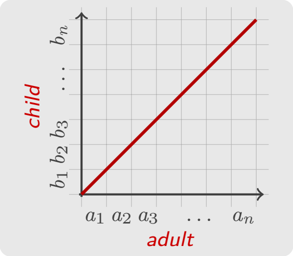
Turn-based Cross-Recurrence Plots

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: ...

⇒

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



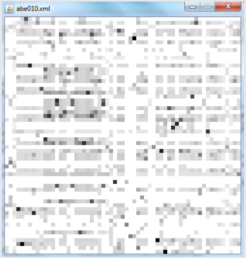
Recurrence (coordination) score for each (i, j)

- *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)
- *upper recurrence*: child's turn comes after adult's *adult* $\leftarrow$ *child*
- *lower recurrence*: adult's turn comes after child's *child* $\leftarrow$ *adult*

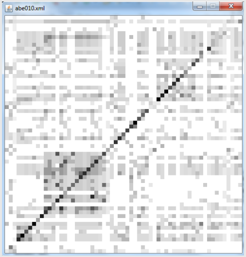
Turn-based Cross-Recurrence Plots

CRP of a dialogue with Abe (2.5 years old):

order of turns shuffled



original dialogue



Same *global* recurrence but very different *local* recurrence

⇝ global: chance recurrence regardless of temporal development of interaction

Linguistic Measures of Recurrence

Syntactic recurrence: number of shared part-of-speech bigrams factoring out lexical identity, normalised by length of longest turn.

Lexical recurrence: shared lexeme unigrams / biagrams factoring out lexical identity, normalised by length of longest turn.

```
Adult: you are pressing a button and what happens ?
       PRO|you AUX|be PART|press DET|a N|buttton CJ|and PRO|what V|happen
Child: what happens the horse tail
       PRO|what V|happen DET|the N|horse N|tail
```

Conceptual recurrence: semantic similarity, e.g., $\langle \texttt{N|dog} \rangle \approx \langle \texttt{V|bark} \rangle$

- distributional semantic model: 2-billion-word WaCuk corpus and the DISSECT toolkit (Dinu, Pham & Baroni, 2013)
- one vector per turn by adding up the lexical vectors
- cosine of a turn pair (i, j) as the convergence score

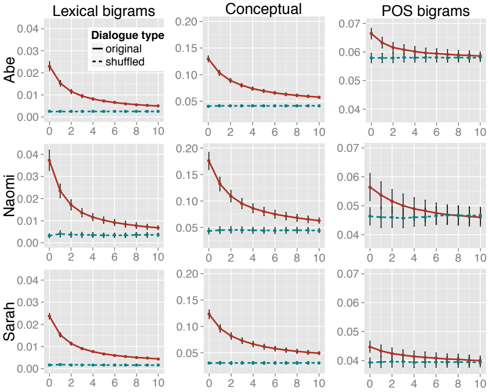
Data

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

corpus	age range	# dialogues	av. # turns/dialogue
Abe	2;5 – 5;0	210	191 (sd=74)
Sarah	2;6 – 5;1	107	340 (sd=84)
Naomi	1;11 – 4;9	62	152 (sd=100)

We generate a *CRP* for each dialogue, computing convergence values for all turn pairs (i, j) for each of the linguistic convergence measures: *lexical*, *syntactic*, *conceptual*.

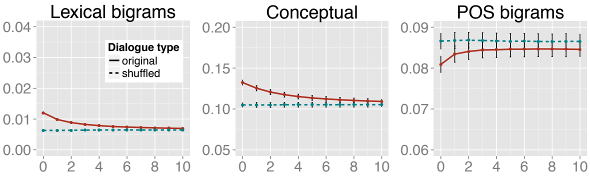
Results: child-adult dialogue



- *local vs. global*: significantly more local coordination.
- *directionality*: both coordinate more at local levels, but the adult recurs with the child significantly more.

Results: adult-adult dialogue

For comparison: $\sim$1000 *adult-adult dialogues* from Switchboard. We ignore backchannels (*"uh huh"*) since they are not considered proper turns (19% of all utterances).



- Semantic lexical/conceptual measures, same trend: above-chance convergence in close-by turns.
- Syntactic measure: very different coordination patterns, with adults showing syntactic *divergence* at adjacent turns:
  $\rightsquigarrow$ less recurrence than expected by chance.

Why?

Contrast with previous evidence of syntactic alignment in adult-adult dialogue (e.g., Pickering & Ferreira 2008), but not surprising

⇝ advancing a conversation requires *different dialogue acts* with distinct syntactic patterns.

Why is there syntactic recurrence in child-adult dialogue?

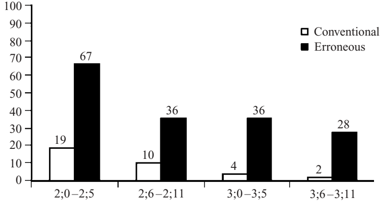
- *feedback mechanism* to ratify linguistic constructions?
- possibly related to *reformulations* / *recasts* / *corrective feedback*

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

Reformulations

M. Chouinard & E. Clark (2003) Adult reformulations of child errors as negative evidence, *Journal of Child Language*.

- Adults check up on the meaning intended by the child.
- 3 English and 2 French children (longitudinal data)
- Around 2/3 of erroneous utterances are reformulated by the adult.
- All types of errors (phonology, morphology, lexicon, syntax).
- Children attend to and respond to the reformulations



% of Abe's conventional utterances replayed and erroneous utterances reformulated.

S. Hiller & R. Fernández (2016) A Data-driven Investigation of Corrective Feedback on Subject Omission Errors in First Language Acquisition. In *Proceedings of CoNLL*.

Aim: large scale data-driven analysis to test the influence of corrective feedback on language learning

Outline of the approach:

Operationalize the phenomenon

- Definition and taxonomy of corrective feedback (CF)

Corpus study

- Identify frequencies of different kinds of CF
- In a manually annotated subset of the data

Investigate the influence of CF on language learning

- Focus on subject omission errors (SOE)
- Automatically detect errors and corrections in a larger dataset
- Test whether CF can predict decrease in SOE, when controlling for other predictors

Corrective Feedback

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

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.

Data Selection and Preprocessing

All relevant files from the English part of the CHILDES database

25 children	Total
transcripts	1,683
utterances	1,598,838
candidate CF	**136,152**
(exchanges with partial overlap)	

Additional information added automatically:

- Morphological decomposition, POS tags (CLAN)
- Syntactic dependency parsing (MEGRASP)
- Information on overlap between child-adult utterance pairs (CHIP)

Data Selection and Preprocessing

CHI: I climb up daddy .	
– **POS & morph**	`%mor: pro.sub|I v|climb prep|up n|daddy`
– **dependency**	`%gra: 1|2|SUBJ 2|0|ROOT 3|2|JCT 4|3|POBJ`
DAD: you did climb over daddy .	
– **POS & morph**	`%mor: pro|you v|do.PAST v|climb prep|over n|daddy`
– **dependency**	`%gra: 1|2|SUBJ 2|0|ROOT 3|2|OBJ 4|3|JCT 5|4|POBJ`
– **overlap**	`%adu: $EXA:climb $EXA:daddy $ADD:you did $ADD:over $DEL:i $DEL:up $REP=0.40`
manual annotation	`%cof: $CF $ERR=umorph:prep; $TYP=subst`

Corpus Study

4 children, 4-6 transcripts per child, 2,627 candidate CF exchanges.

Examples

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.
```

	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

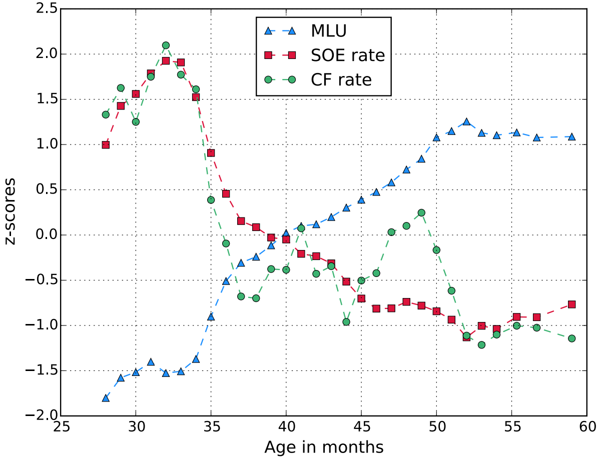
Focus: subject omission errors (SOE)

Automatic Detection

- Find *high-precision automatic classifiers* for SOE and CF on SOE
- To enable an analysis of the *whole dataset*
- Using the manually annotated data as *training set*
- *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

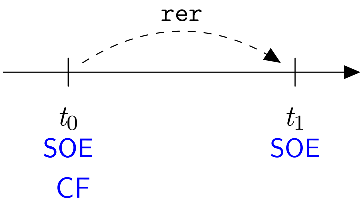
Adam, Brown corpus



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

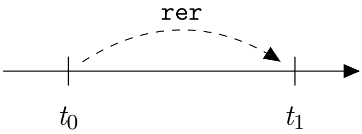
- child age
- child / adult MLU
- child / adult vocabulary size
- adult subject omissions
- proportion of child speech

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

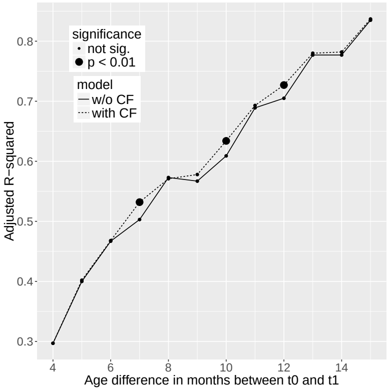


Setting 1: any t_0 and any $d(t_0, t_1) \geq 1$ month

- *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

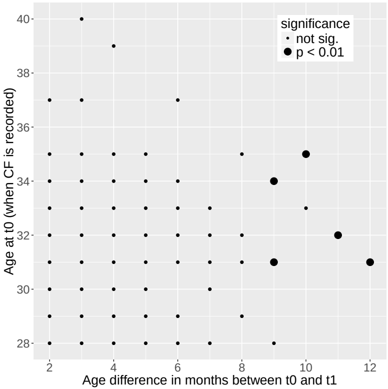
Results

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.

Conclusions of this study

- Local interaction can function as negative input and contribute to language learning
- Our analysis shows that *CF contributes to learning* of subject inclusion in English, after a lag of at least 7–9 months

- *Large scale data-driven* analysis using *automatic classifiers*
- Caution required regarding possible bias introduced by classification errors

Possible next steps:

- Extend the analysis to other kinds of errors
- How can we model this interactive process for automated learners?