

Modeling the Emergence of Natural Language Lexicons

TONIO DE LA CONTRACTION DEL CONTRACTION DE LA CO

Russell Richie¹, Charles Yang², & Marie Coppola³

- ¹ University of Connecticut, Department of Psychology, Language Plasticity IGERT, igert.cogsci.uconn.edu
- ² University of Pennsylvania, Departments of Linguistics, Computer Science, and Psychology
- ³ University of Connecticut, Departments of Psychology and Linguistics, Cognitive Science Program, Language Plasticity IGERT

KEFEKENCE

1.Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language.

Proceedings of the National Academy of Sciences, 104(12), 5241-5245.

Calantucci B. Garrod S. and Roberts G. (2012) Experimental Semiotics. Language and Linguistics Comp.

2.Galantucci, B., Garrod, S., and Roberts, G. (2012) Experimental Semiotics. *Language and Linguistics Compass*, 6(8), 477-493.

3.Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28(6), 937-962.

4.Senghas, R. J. (1997). An 'unspeakable, unwriteable' language: Deaf identity, language & personhood among the first cohorts of Nicaraguan signers. (unpublished doctoral dissertation). University of Rochester, NY.

ACKNOWLEDGMENTS

Thanks to our Deaf and Hearing Nicaraguan participants, and to members of the Language Creation Lab. This work was supported in part by NSF IGERT Grant 1144399 to the University of Connecticut, and by NIH grant P30 DC010751 to Marie Coppola and Diane Lillo-Martin.

Homesigners converge with respective

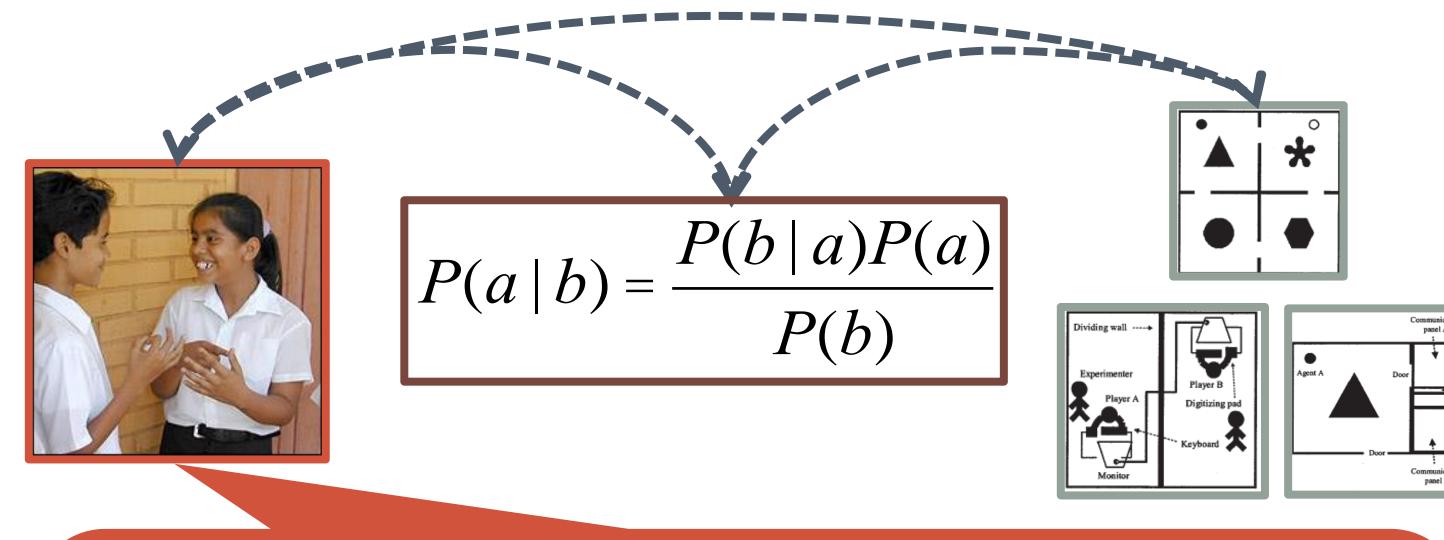
partners, but not fully

INTRODUCTION

From what kinds of learners, and interactions among learners, does language emerge?

Most data on the question come from computational models¹ and experiments² -- rarely from naturalistic data.

Weak integration of these approaches.



Systems created by Deaf individuals, either by:

- Deaf communities -- give rise to natural signed languages, some of which emerged recently (e.g. ~35 y.o. Nicaraguan Sign Language, NSL⁴; cf. ~200 y.o. American Sign Language)
- Homesigners -- deaf individuals, isolated from other Deaf, who create gestural systems for use with hearing families

For example, w.r.t. how linguistic conventions emerge among interacting individuals:

CAT!

- Many agent-based models³, few links to experiments (e.g. above)
- And no direct, longitudinal data from naturally emerging systems on conventionalization

METHODS

| Family 1 | Family 2 | Family 3 | Family 4 |
|----------------|-----------------|----------------|--------------------|
| Homesigner (M) | Homesigner (F) | Homesigner (M) | Homesigner (M) |
| Friend (M) | Mother | Mother | Younger brother |
| Mother | Younger brother | | Younger sister |
| Older brother | Younger sister | Hearing fa | milies and friends |

Procedure

- Elicited gesture responses to 22 cross-linguistically and –culturally basic objects
- Tested 2002, '04, '06, '11

Coding

- Multi-gesture responses common (62% of all responses)
- Coded each gesture for Conceptual Component (CC), or aspect of item's meaning gesture represents





Measuring conventionalization

- Response is binary vector of CC's
- Conventionalization is distance between two responses

| 'cow' | HORNS | MILKING | DRINK | Distance from HSer |
|------------|-------|---------|-------|--------------------|
| Homesigner | 1 | 1 | 0 | n/a |
| Sister | 1 | 1 | 1 | 1 |
| Mother | 0 | 1 | 1 | 2 |

RESULTS (SEE FIGURE AT TOP RIGHT)

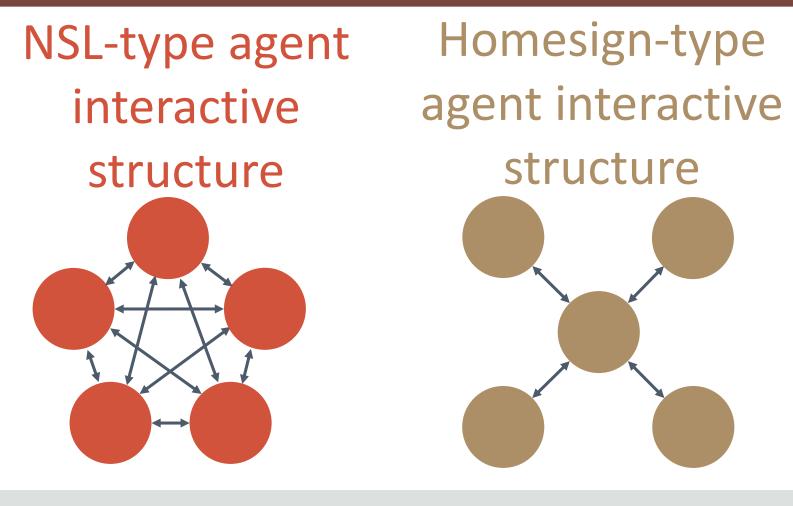
- Gradual convergence between homesigners and partners.
 - All 9 slopes of HS-CP distance = $f(year \ of \ testing)$ are negative (Wilcoxon Signed Rank, W=0, p<.01).
- But convergence not complete
 - 2011 HS-CP distances still significantly greater than 0; all tests significant (W's \geq 91, p's \leq .001).
- Cf. NSL, which conventionalized in <15 years⁴
 - 1978 Deaf community formed in Managua
 - 1993 'standardization seminars' held in rest of Nicaragua to spread signs developed in Managua
- Because of NSL users' richer patterns of interaction?
 - Use agent-based model to test this hypothesis

AGENT-BASED MODEL

2004

2006

2002



binary random variables:
E.g. CC#1 is 1 with p=.9, 0 with p=.1; CC#2 is 1 with p=.3, 0 with p=.7...

belief about word

2011

Agents have probabilistic

pronunciation: vector of

-Mom

-O. Bro

-Friend

–Y. Bro

-Mom

-Y. Sis

-Mom

-Y. Sis

-Y. Bro

 $p_1 = .9$

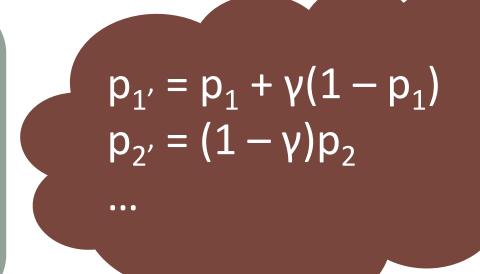
 $p_2 = .3$

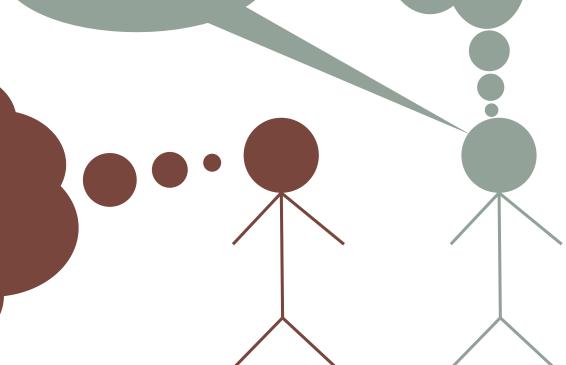
Avg. # interaxns required(% simulations reaching convent w/in 2M interaxns)

260K(100%) 698K (80%)

10...

Convergence faster in NSL-type interactive structure





CONCLUSION

Behavioral results: Homesign systems' lexicons are conventionalizing slower than did NSL's

Modeling results: Multi-agent reinforcement learning model suggests richer patterns of interaction among agents hasten conventionalization

More integration of computational, experimental, and naturalistic data will lead to unprecedented insight into language emergence

PRESENT STUDY

- Show conventionalization in Nicaraguan homesign systems, and compare to prior, indirect evidence of conventionalization in NSL
- Use agent-based model of conventionalization to test hypothesis that differences in patterns of interaction among users (partially) explain differences in homesign and NSL conventionalization