Exploring Language Evolution using Behavioral Simulations
-- What have been done and what to do

Tao GONG

Language Evolution and Computation Research Unit
University of Edinburgh

Email@LEC: tgong@staffmail.ed.ac.uk
Homepage@MPI EVA: http://www.eva.mpg.de/lingua/staff/gong/home.php

Oct. 2, 2009

Computational simulation and behavioral models

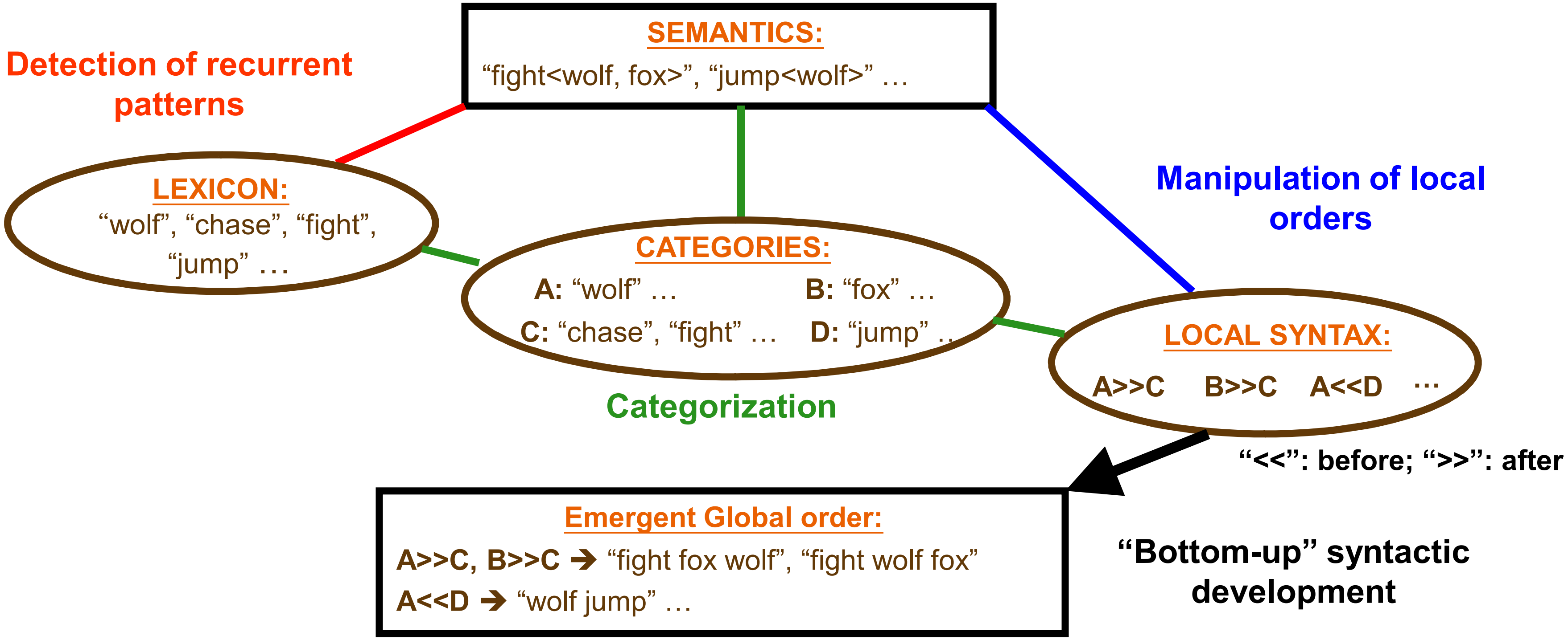
- Computational models ➔ "operational" hypotheses or theories;
- Simulation results ➔ empirical predictions of those theories or hypotheses.

- Two types of computational models:
  - Behavioral models: specify a set of actual behaviors of an individual or a group of individuals during language processing to explore their effects on language evolution. E.g., Iterated Learning Model, Fluent Construction Grammar.
  - Mathematical models: abstract the performances of linguistic behaviors into mathematical equations, and predict the tendencies of language evolution based on statistical analyses. E.g., Language competition models, Bayesian learning models.

- My work:
  - Explore if domain-general abilities, such as pattern extraction and sequential learning, could lead to compositionality, syntactic regularity, and categorization;
  - Design multi-agent, behavioral models to simulate linguistic communications;
  - Examine the effects of socio-cultural factors on language evolution.

The models developed: 1) the lexicon-syntax coevolution model

- Study the coevolutionary emergence of lexical items and simple word order;
- Semantic space:
  - Intransitive meanings: “Pr_1<Ag>”: e.g., “hop<deer>”;
  - Transitive meanings: “Pr_2<Ag, Pat>”: e.g., “chase<fox, wolf>”;



Local syntax: binary sequential order relation (before or after) between 2 lexical items;
Global word order (e.g., SVO, SOV) results from re-use of local syntax; S: subject; V: verb; O, object;

Key features: a) the item-based learning mechanisms

- "Learning from experience":
  - Pattern extraction to acquire lexical rules;
  - Sequential learning to acquire local orders and categories: Lexical rules having the same semantic role (Ag, Pat or Pr) and similarly used (local order) in M-U mappings can be categorized;

Newly acquired lexical rules

(a) "fox" ←→ /d/ (0.5)

(b) "run<#>" ←→ /m/ (0.5)

(c) "wolf" ←→ /a c/ (0.5)

(d) "chase<#,#>" ←→ /b/ (0.5)

M-U mappings (previous experience) in an individual's buffer

(1) "hop<fox>" ←→ /d h/

(2) "run<fox>" ←→ /d m/

(3) "run<wolf>" ←→ /a c m/

(4) "chase<wolf, deer>" ←→ /a c f b/

(5) "chase<fox, sheep>" ←→ /d b e/

Category 1 (S)

Category 2 (V)

Newly acquired categories

List of lexical rules

(a) "fox" ←→ /d/ (0.5) [0.5]

(c) "wolf" ←→ /a c/ (0.5) [0.5]

List of syntactic rules

Cat1 << Cat2(b)(S)(0.5)

List of lexical rules

(b) "run<#>" ←→ /m/ (0.5) [0.5]

(d) "chase<#, #>" ←→ /b/ (0.5) [0.5]

List of syntactic rules

Cat1 << Cat2(c)(S)(0.5)

Numbers within () are rule strengths; numbers within [] are association weights;

- Connections with LEC's work:
  - How many instances are necessary to trigger generalization? --- Anna's work
  - Whether human subjects are sensitive to local order regularity? --- Monica, Barbora & Andrew's work

Key features: b) the "Verb-island" hypothesis (Tomasello, 2003)

- Category merge:
  - Two categories with identical syntactic roles (S, V, or O) share some lexical member(s);
  - Some of their lexical members are similarly used (having the same local order) with respect to another lexical item;
  - The merged category comprises all lexical and syntactic members of the original categories;

M-U mappings in the buffer

(1) "fight<fox, wolf>" ←→ /d a c b/

(2) "fight<deer, gazelle>" ←→ /k b n/

Acquired lexical rule

"fight<# , #>" ←→ /b/ (0.5)

Available categories

Category 1 (S): *List of lexical rules:*

{(a) "fox"←→/d/ (0.6)} [0.5]

{(b) "wolf"←→/a c/ (0.6)} [0.7]

List of syntactic rules:

(I) Category 1 (S) << Category 3 (O) (0.6)

Category 2 (S): *List of lexical rules:*

{(b) "wolf"←→/a c/ (0.6)} [0.5]

{(c) "deer"←→/k/ (0.7)} [0.8]

List of syntactic rules:

(II) Category 2 (S) >> Category 4 (V) (0.7)

Merged category

Category 1 (S): *List of lexical rules:*

{(a) "fox"←→/d/ (0.6)} [0.5]

{(b) "wolf"←→/a c/ (0.6)} [0.7]

{(c) "deer"←→/k/ (0.7)} [0.8]

List of syntactic rules:

(I) Category 1 (S) << Category 3 (O) (0.6)

(II) Category 1 (S) >> Category 4 (V) (0.7)

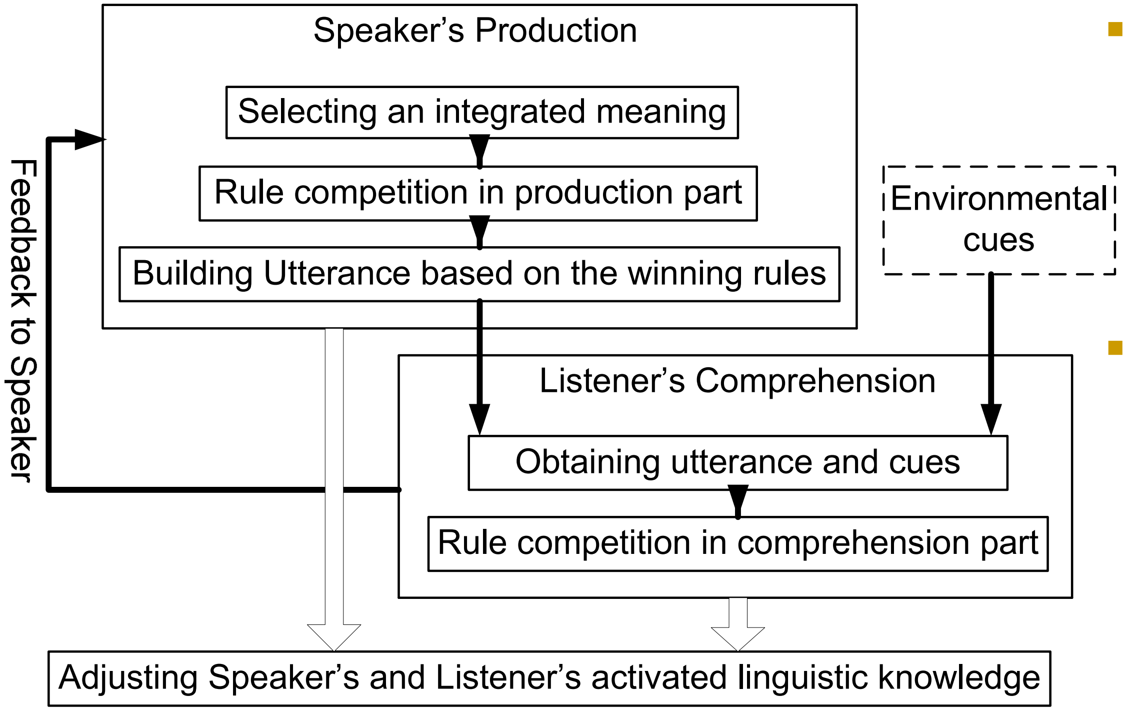
(III) Category 1 (S) << Lexical rule (d) (0.5)

Numbers within () are rule strengths; numbers within [] are association weights;

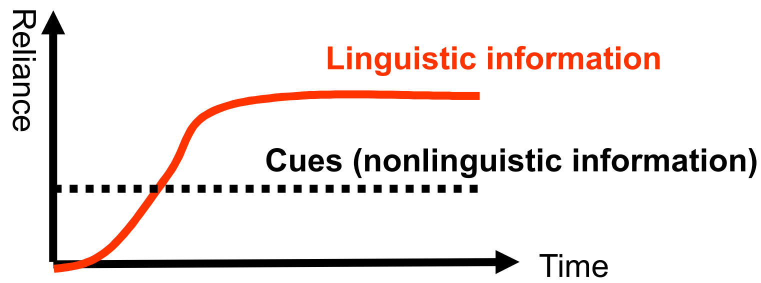
Tomasello, M. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Key features: c) "indirect meaning transfer"

- Nonlinguistic information (environmental cues) is involved in comprehension;
- Reliability of Cue ➔ the probability with which the speaker's intended meaning is in the cue available to the listener;



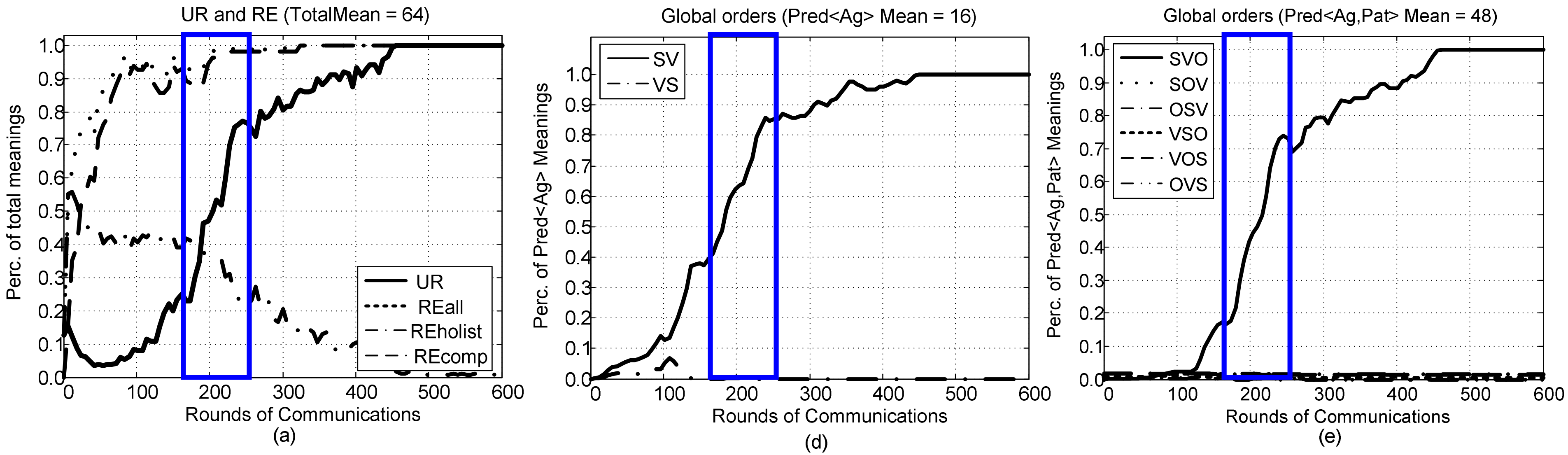
- "Indirect meaning transfer" (Smith, 2003);
  - No mind reading;
  - Confidence feedback;
  - Coordination of linguistic and nonlinguistic information in comprehension;
- The emergence of displacement (Hockett, 1960)



- Connections with LEC's work:
  - Intentionality sharing and coordination of multisensory channels --- Barbora's work;
  - Inferential capacities before language -- Thom's work;

Smith, A. D. M. 2003. Intelligent meaning creation in a clumpy world helps communication. *Artificial Life*, 9(2): 559–574.
Hockett, C. F. 1960. The origin of speech. *Scientific American*, 203: 88–96.

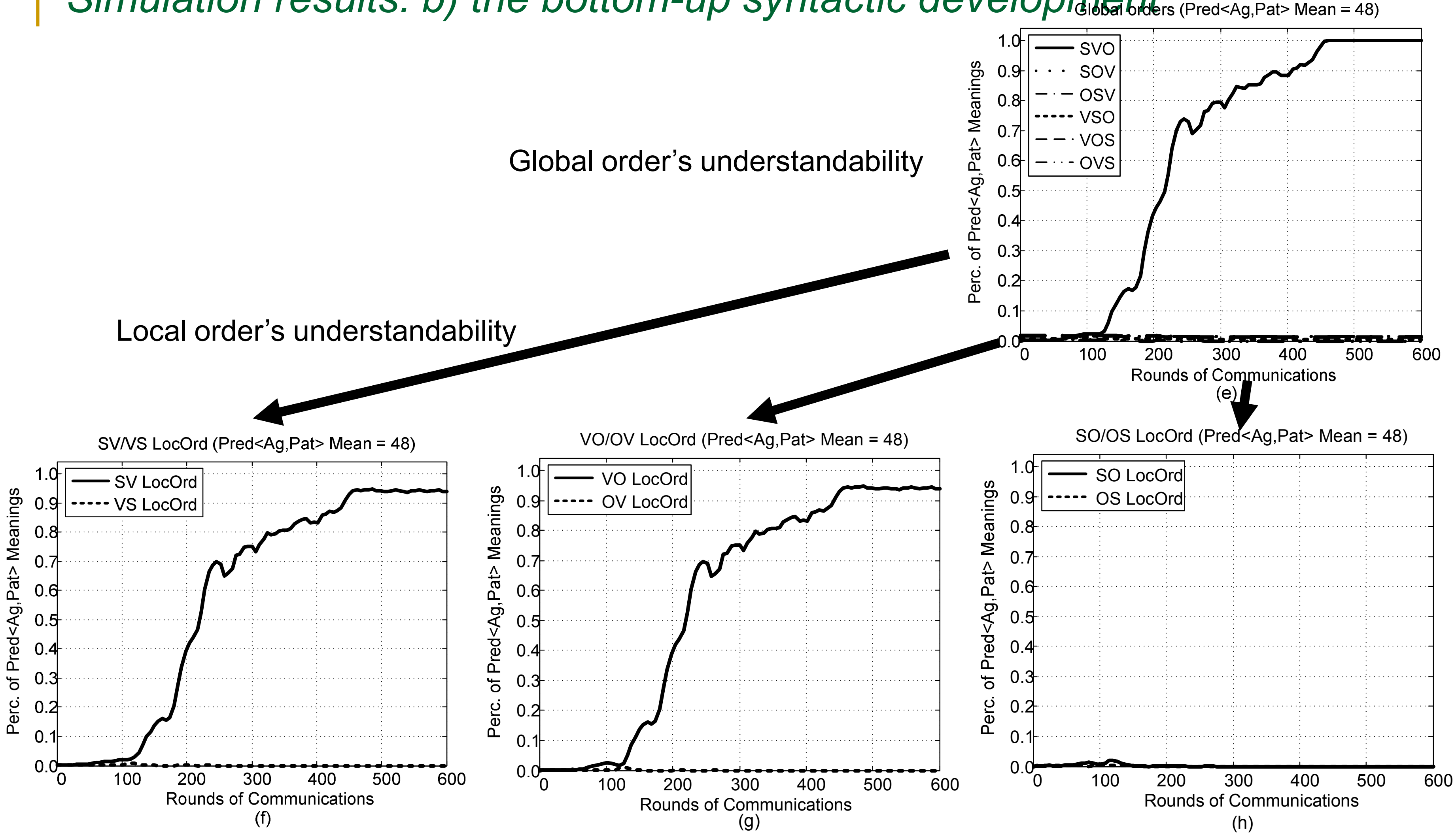
Simulation results: a) the lexicon-syntax coevolutionary emergence



- Simulation setup: 10 individuals, 64 meanings, 600 communications per individual, 20 utterance exchanges in each communication; initially 8 holistic rules, reliability of cue is 0.6, buffer size is 40;

- Emergence of compositionality:
  - *RE* ➔ Holistic signals → compositional expressions;
  - *UR* ➔ The emergence of a common compositional language (the S-shape sharp transition).
- Emergence of regularity: some global word orders become prevalent;
  - Prevalent order: the global / local orders that are frequently used in comprehensions;
  - e.g., SV for "Pr_1<Ag>" meanings; SVO for "Pr_2<Ag, Pat>" meanings;
- Coevolution of compositionality and syntactic regularity:
  - The emergence of common lexical items and that of prevalent global orders are synchronized during the transition from a holistic signalling system to a compositional language.

Simulation results: b) the bottom-up syntactic development



- The "bottom-up" syntactic development: the prevalent global word order reflects the local sequential information specified by the prevalent local orders;

Simulation results: c) the word order bias (Gong et al., 2009)

- With only transitive meanings (Pr<Ag, Pat>), there is no bias on global orders;
- With both intransitive (Pr<Ag>) and transitive meanings:
  - if SO is specified ➔ SVO and SOV become the biased global orders;
  - if OS is specified ➔ OVS and VOS become the biased global orders;

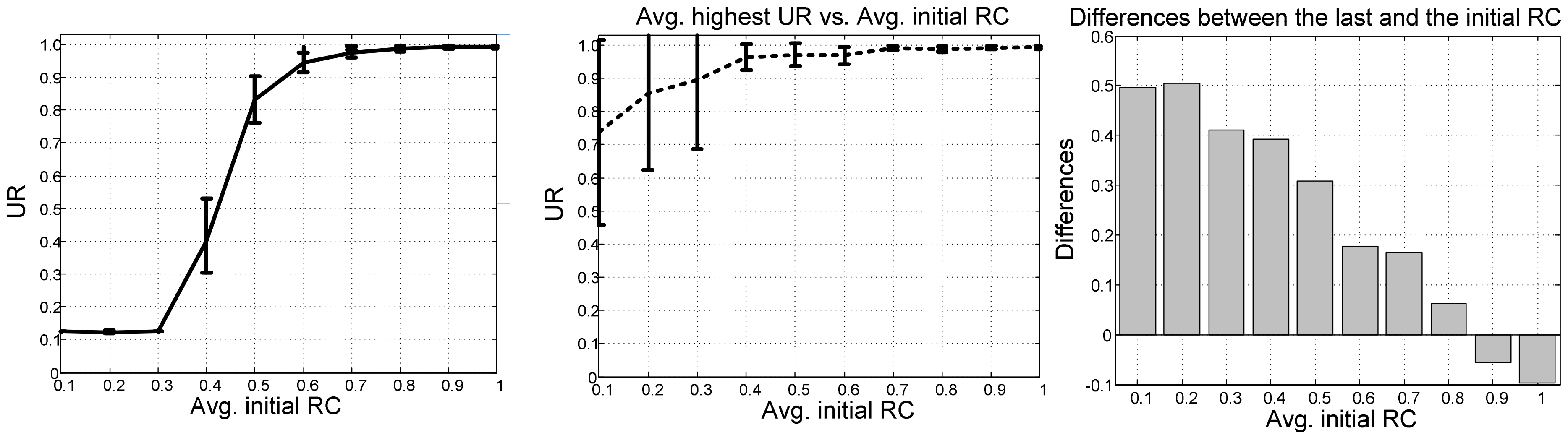
- Factors that trigger the word order bias:
  - The clarification requirement: To distinguish Ag and Pat in "Pr<Ag, Pat>" meanings;
  - The semantic structure: Intransitive meanings require SV or VS ➔ O-dominant syntaxes do not specify the order between S and V, so they are unbiased;

  - The "semantics driving syntax" hypothesis (Schoenemann 1997):
  - Syntactic structure arose from the need to convey semantic structures (Newmeyer, 2008);
  - The change in semantics causes changes in the emergent syntactic structures.
  - Self-organisation of multiple linguistic components during language processing:
  - The functional principles at the sentence level are trigged by lower-level constraints on local orders and semantics;

Gong, T., Minett, J. W., Wang, W. S-Y. 2009. A simulation study on word order bias. *Interaction Studies*, 10(1): 51-76.
Schoenemann, P. T. 2005. Conceptual complexity and the brain: Understanding language origins. In: Minett, J. W., Wang, W. S-Y. (Eds.), *Language acquisition, change and emergence: Essays in evolutionary linguistics* . Hong Kong: City University of Hong Kong Press, pp. 47-94.
Newmeyer, F. J. 2008. Conceptualization, communication, and the origins of grammar. In: Laks, B. (Ed.), *Origin and evolution of languages: Approaches, models, paradigms*. London: Equinox, pp. 112-132.

Simulation results: d) the roles of reliability of cue (RC)

- RC reflects the ability of intentionality sharing (Tomasello et al. 2005);
  - Speaker: RC ➔ the probability of expressing ongoing events in the common ground;
  - Listener: RC ➔ the probability of using ongoing events to interpret heard utterances;



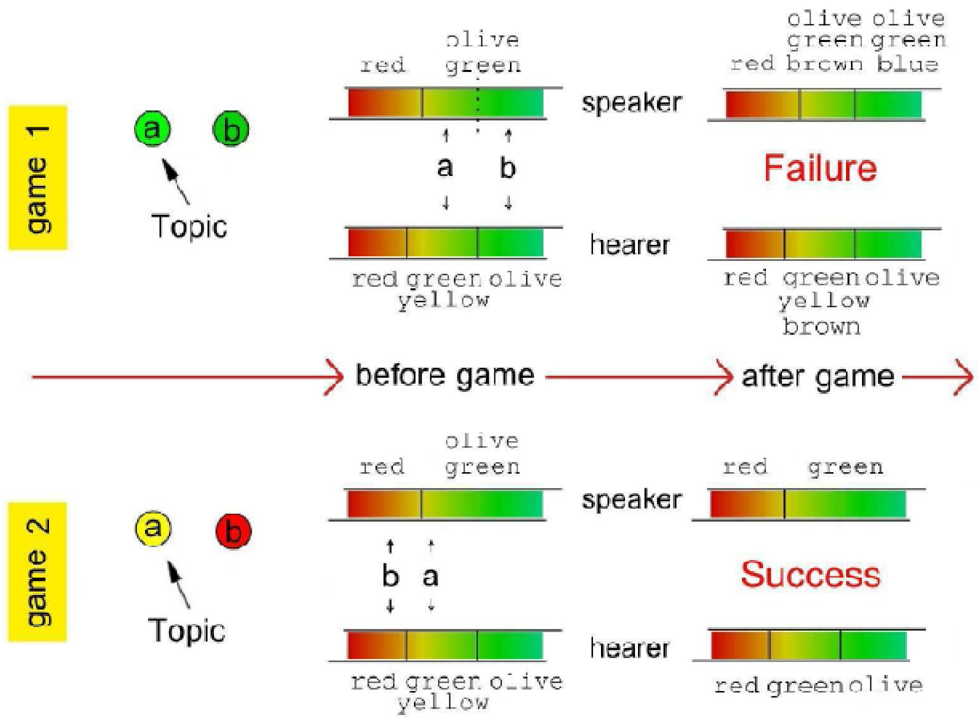
- Under a fixed initial RC ➔ the threshold of RC;
- Under a changeable RC ➔ An initially low RC can be increased after generations of communications, together with the emergence of a displaced language;
- The innateness of high RC in humans? (Tomasello, 2008): The high level of intentionality sharing in humans could be due to piggybacking on language.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences (BBS)*, 28(5): 675–691.

Tomasello, M. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.

The models developed: 2) the category game (Puglisi et al., 2008)

- Inspired from the naming game (Baronchelli et al., 2006);
- Study the coevolution of perceptual categories and their word labels;



d_{min} (just noticeable difference): the minimum distance between two stimuli that can be distinguished by agents;

Perceptual categories: a subdivision of the stimuli space [0.0 1.0];

Linguistic categories: perceptual categories that share the same last winning word labels;

Key Features:

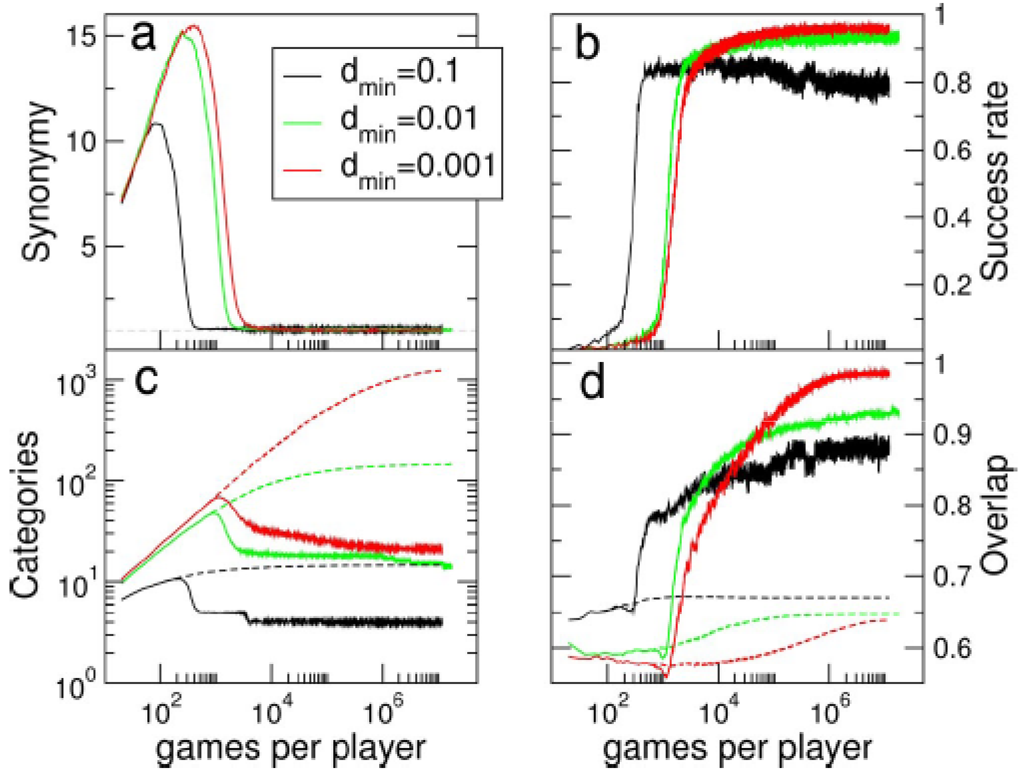
1) Continuous perceptual space;
2) No focal colors beforehand;
3) Few parameters & simple learning;

- Connections with LEC's work:
  - Categorization in 2-D space based on shapes and colors;
  - Categorization modeling;

Baronchelli, A., Felici, M., Caglioti, E., Loreto, V., and Steels, L. 2006. Sharp transition towards shared vocabularies in multiagent systems. *Journal of Statistical Mechanics* P06014.

Puglisi, A., Baronchelli, A., and Loreto, V. 2008. Cultural route to the emergence of linguistic categories. *PNAS,* 105: 7936.

Simulation results: coevolution of linguistic categories and their lexical labels;



N=100

Perceptual categories increase, while linguistic categories decrease and be stable

- Successful rate: the percentage of successful category games between all pairs of agents;
- Overlap (O): the alignment of perceptual/linguistic category boundaries;

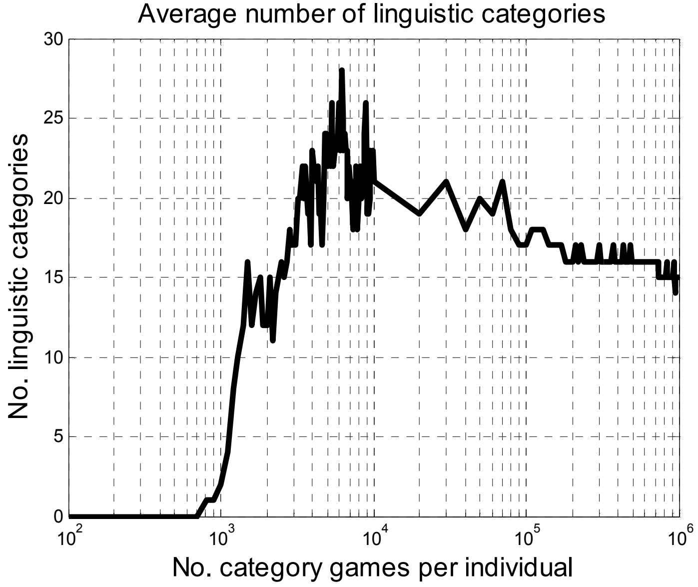
$$O = 2\sum_{i<j} o_{ij}/N(N-1)\ , \quad o_{ij} = \frac{2\sum_{c_{ij}}(lc_{ij})^2}{\sum_{c_i}(lc_i)^2 + \sum_{c_j}(lc_j)^2}$$

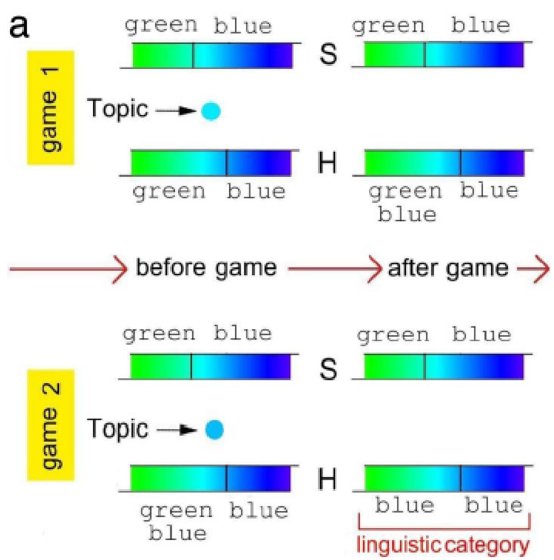
```
Game=5000000
NumComLingCat= 15;
LingCat [  1]: LastWinWord= 1112, Range[0.00 0.03)
LingCat [  2]: LastWinWord= 2054, Range[0.06 0.08)
LingCat [  3]: LastWinWord= 2475, Range[0.10 0.15)
LingCat [  4]: LastWinWord= 3224, Range[0.18 0.23)
LingCat [  5]: LastWinWord= 3634, Range[0.26 0.28)
LingCat [  6]: LastWinWord= 1200, Range[0.31 0.35)
LingCat [  7]: LastWinWord= 4034, Range[0.39 0.41)
LingCat [  8]: LastWinWord= 3860, Range[0.42 0.44)
LingCat [  9]: LastWinWord= 2464, Range[0.47 0.56)
LingCat [ 10]: LastWinWord= 2200, Range[0.60 0.67)
LingCat [ 11]: LastWinWord= 3585, Range[0.71 0.74)
LingCat [ 12]: LastWinWord= 1158, Range[0.76 0.80)
LingCat [ 13]: LastWinWord= 1549, Range[0.83 0.87)
LingCat [ 14]: LastWinWord=  649, Range[0.88 0.92)
LingCat [ 15]: LastWinWord= 1067, Range[0.94 1.00)
```

d_{min}=0.01

Simulation results: the evolution of linguistic categories

- Four phases of evolution (All agents initially have only a perceptual category [0, 1) with no words):
  - The increasing phase: the pressure for discrimination makes the number of perceptual categories and their labels increase.
  - The coarsening stage: when on average only one word is recognized by the whole population for each perceptual category, words expand their dominion across adjacent perceptual categories, joining these categories into *linguistic categories*.
  - The stable stage: coarsening of these categories becomes slower and slower, with a dynamical arrest analogous to the physical process in which super-cooled liquids close to the glass transition. The categorization pattern has a degree of sharing between 90% and 100% and remains stable.
  - The dropping stage: the number of linguistic categories drops down, caused by the slow diffusion of category boundaries that ultimately takes place due to small size effects.



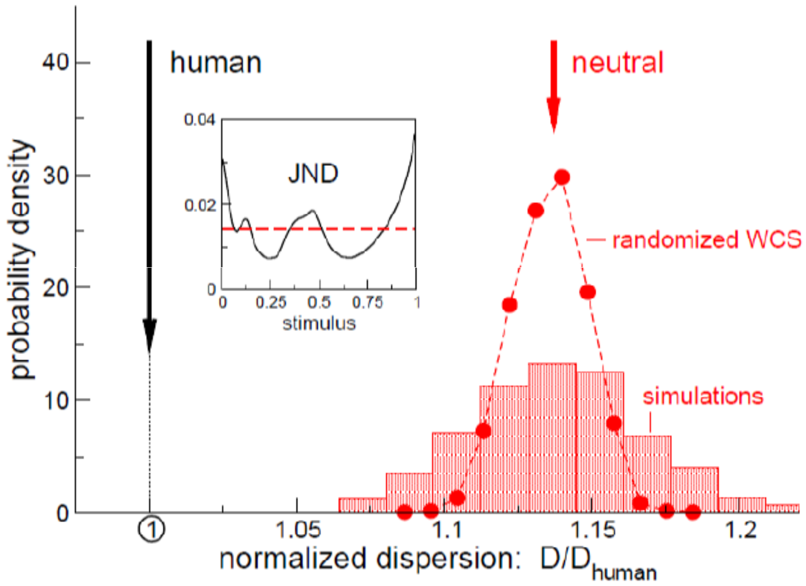


Simulation results: the human-d_{min} for color categorization (Baronchelli et al., submitted)

- Humans have a non-uniform perceptive ability (Long et al., 2006):
- Dispersion (Kay and Regier, 2003): *l* and *l** are two different languages, *c* and *c** are two basic color terms from languages *l* and *l** respectively and distance(*c*; *c**) is the distance between the focal points.

$$D_{S_0} = \sum_{l,l^* \in S_0} \sum_{c \in l} \min_{c^* \in l^*} \text{distance}(c, c^*)$$

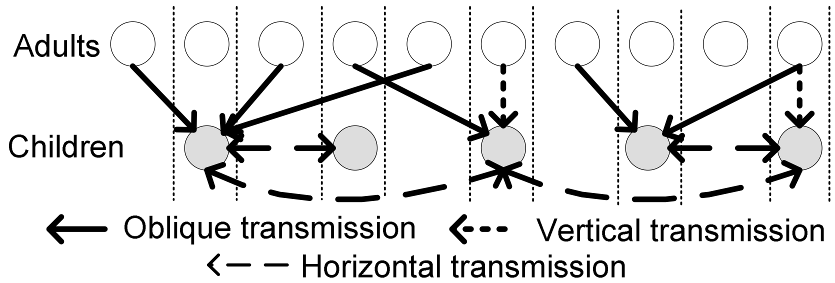
- The influence of human-d_{min} on categorization: Universal categorization pattern may be due to some bio-physiological factor taking effect during cultural transmission



Baronchelli, A., Gong, T., Puglisi, A., and Loreto, V. Modeling the emergence of universal categorization. PNAS, submitted.
Long, F, Yang, Z, and Purves, D. 2006. Spectral statistics in natural scenes predict hue, saturation, and brightness. *PNAS*, 103: 6013-6018.
Kay, P. and Regier, T. 2003. Resolving the question of color naming universals. *PNAS*, 100:9085-9089.

Work to do: 1) the roles of various forms of cultural transmission

- The acquisition framework:



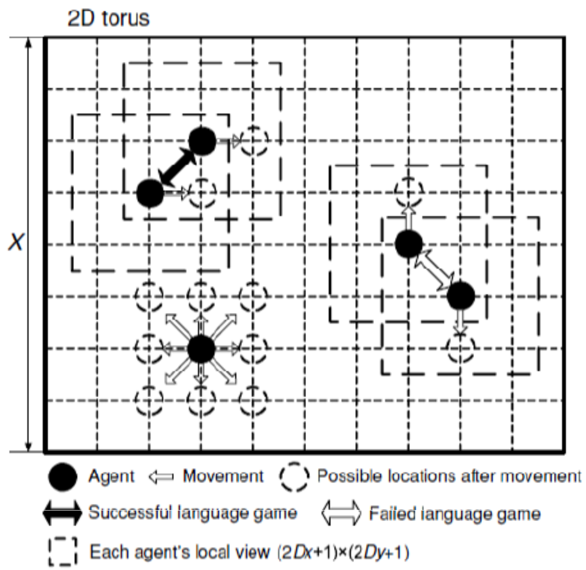
- Three forms of cultural transmissions:
  - Horizontal transmission: intra-generational;
  - Vertical transmission: inter-generational;
  - Oblique transmission: inter-generational;
  - Grand parent-grand child transmission;

- UR_{ser}: UR between individuals of Gen n and Gen n+1;
- UR_{ini}: UR between individuals of Gen 1 and Gen n;

- My aims:
  - Study under what combinations of these transmissions, a communal language can be efficiently trigged and well maintained across generations;
  - Language evolution is a dynamic equilibrium; in the short run, understandability is maintained across generations; in the long run, change is inevitable;

Gong, T., Minett, J. W., and Wang, W. S-Y. in press. A simulation study exploring the role of cultural transmission in language evolution. *Connection Science*.

Work to do: 2) The effects of social structures on language evolution



- Linguistic understandability may affect social connections and vice versa:
  - Mutual understanding of linguistic utterances ➔ social cohesion;
  - Geographical constraints on communications ➔ social clusters;

- Available studies:
  - Predefined social structures:
    - Single popular agent;
    - Power-law distributed popularities;
    - Scale-free, small-world networks;
    - Intra-community vs. inter-community communications;
  - Language models: the lexicon-syntax coevolution model, the naming game/the category game;

- Some discussions with Justin:
  - Whether agents can acquire sufficient social knowledge via gossiping;
  - Whether changes in network structure can be spread by linguistic communications;
  - Whether different ways of gossiping can affect social structure;

Gong, T., Loreto, V., and Wang, W. S-Y. 2008. Conventionalization of linguistic knowledge under communication constraints. *Biological Theory*, 3(2): 154-163.

Work to do: 3) design empirical experiments for simulated behaviors

- Several assumptions on learning behaviors built in the model:
  - Pattern extraction without location constraint;
  - Sensitivity to local order relations;
  - Generalization based on similar usage in terms of order relations;
- Cooperation with Anna's and Monica, Barbora & Andrew's experiments;

Work to do: 4) other simulation work to do

- Polya urn and Price Equation simulation (with Monica);
- 2-D space category game based on empirical experiments;
- The word order game:
  - Individuals develop only local orders between 2 lexical items;
  - Whether re-use of local orders leads to consistent global order(s) to regulate multiple lexical items;
  - What are the necessary constraints avoiding for agents to build up global orders → syntactic hierarchy or constructions;