

Working Group and Workshop Exploratory Models of Language Acquisition  
October 17-22, 2003:

---

Language Emergence: a Self-Organized Model  
using Indirect Meaning Transference based on  
interaction of multiply factors

---

*Tao Gong*

Language Engineering Laboratory (LEL),  
City University of Hong Kong

*Oct. 18th, 2003 on SFI, Santa Fe, NM, USA*

# *Contents*

- Background
- Description of the model
- Main results & Discussions
- Future directions

## Background

Mainstream scenarios on language emergence: (Schoenemann, 1999)

1) **Bickerton's "bootstrapping, innate" scenario:**

Words → "innate" syntax for combination to express integrated meanings

2) **Wray's "emergent" scenario:**

Holistic units for simple meanings → phonetic similarity (recurrent patterns) triggers decomposition → syntax emerges during decomposition.

**Language is the result of many complex interacting dynamical systems and acquired in communications. (Kirby, 2003, Oliphant, 1997)**

Many "emergent" models simulate the emergence of lexicon or syntax in a communication system:

1) **Ke et al, 2002:** vocabulary emergence, no syntax.

2) **Kirby, 1998, 2002:** learning, direct meaning transference;

3) **Batali, 1998:** communication, direct meaning transference;

4) **Munroe & Cangelosi 2002:** learning, simple language, feedback from environment;

## Limitations of these models:

- 1) **Direct meaning transference (Kirby's model and Batali's model)**  
In direct meaning transference, both speaker and listener can get access to the real meanings, so they clearly know the inferred meaning is right or not, the adjustment of rules is always right. So there is no need to communicate.
- 2) **Main functions of language: representation and communication, the latter is not fully studied in these models.**

Factors in communication, like group size, social structure can influence the language acquisition.

*Our model is also an “emergent” model, simulating the emergence of lexicon as well as simple syntax (word order) at protolanguage level.*

# Description of the model

## ***Features of this model***

- Compositional rules emergence from initial holistic units is simulated.
- **An overlapping convergences of lexicon and syntax.**
- **Indirect meaning transference**, an understanding process based on multiple factors, linguistic rules as well as environment cues, is simulated
- **SOM, rule based competition and adjustment.**
- Inspirations from **Classifier Systems**, a machine learning framework.

## ***Basic Assumptions:***

- Semantic space is set up beforehand; **only integrated meanings (either “agent action” or “agent predicate patient”)** are expressed and interpreted.
- **Word Order** are general features not restricted in language. Agents have the ability to use word order to regulate linguistic utterances.
- Cognitive ability to **acquire environmental information (cues).**
- **Ability to detect recurrent patterns and store them as compositional rules.**
- **No explicit access to other agent’s language**

# Major components

## 1) Transparent and Opaque meanings.

**Transparent meanings:** meanings directly inferable from their constituents:

**Opaque meanings:** meanings not directly inferable from their constituents:  
*drink dog water, chase bear wolf* .

**Transparency and opacity of meaning types show that interpretation is not only based on language information itself. Other information like syntax or environmental information is necessary.**

## 2) Rules of this model: [ mappings/methods + strength ]

(1) **Lexicon Rules:** mappings [ meanings  $\leftrightarrow$  utterances ]

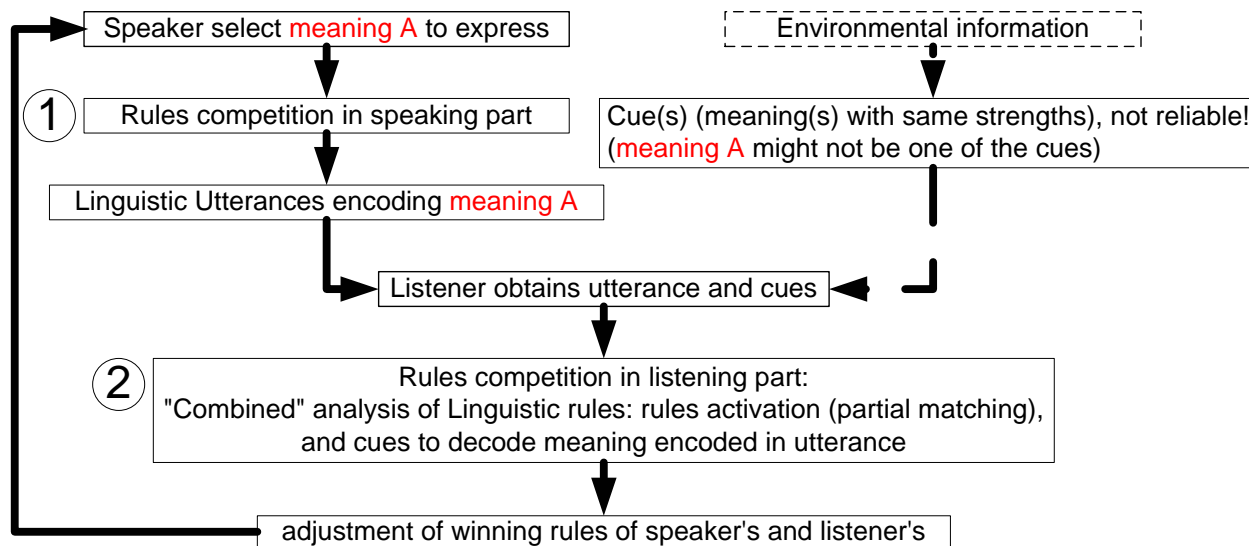
(2) **Word order rules:** methods to construct compositional rules' utterance.

(3) **Strength:** a numeric parameter indicating the frequency of successful use of the rules. Competition and adjustment are based on rule's strength.

**A set of lexicon and order rules with high strengths shared by all agents is an indication of the convergence of language.**

### 3) Indirect meaning transference and rules competition

Normal communication between Speaker and Listener



①  $LexiconWeight \times Avg(LexiconRules \cdot Str) + OrderWeight \times (OrderRule \cdot Str)$

②  $LanguageWeight \times (LexiconWeight \times Avg(LexiconRules \cdot Str) + OrderWeight \times (OrderRule \cdot Str)) + CueWeight \times (Cues \cdot Str)$

1) Environmental information is necessary in early stage of language evolution, it is the main source of meanings.

2) Independent processes happen in Speaking and Listening parts. Expression and Interpretation are the results of many complex interacting dynamical factors.

## 4) Listener's Feedback to Speaker:

**feedback 1: listener believes he understands, if not, feedback 0**

It is an unreliable criterion: the meaning that listener believe to understand might not the one speaker said.

## 5) Inspirations from Classifier Systems:

**Common features with Classifier Systems:**

- a) Similar acquisition process.
- b) Rule activation and rule-based decision.

**Modifications:**

- a) Rules are not only mappings but also combination methods.
- b) Rule activation can be decided by both sides of mappings (meaning, utterance).
- c) Activation and interplay of different rules (lexicon and word order rule).

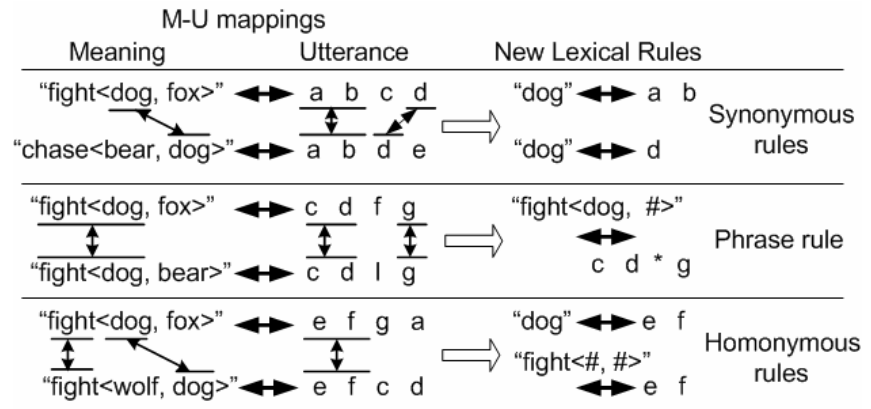
## 6) rule acquisition

### 1) Random invention

- Randomize syllables to map inexpressible integrated meanings or inexpressible constituents. the more the number of the inexpressible items, the less successful chance the speaker express it.

### 2) Detecting recurrent patterns

Flexible search of recurrent patterns to decompose holistic/phrase units without location restriction



a) Synonyms and homophones emerge in detecting recurrent pattern.

b) The residual utterance can be mapped into new rules (cf. Kirby's).

**Later slides will discuss these two!**

# Results and Discussions

## a) Linguistic results

### 1) Simple syntax (Lexicon Order) convergence

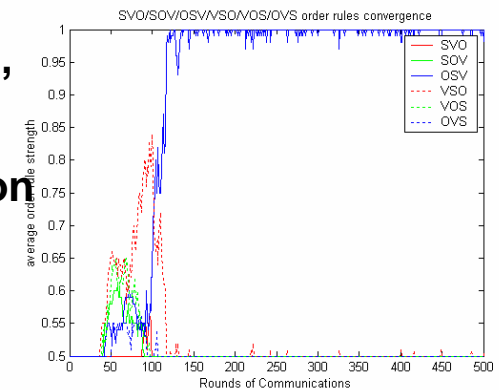
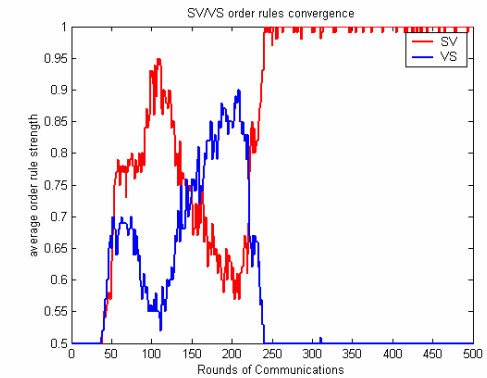
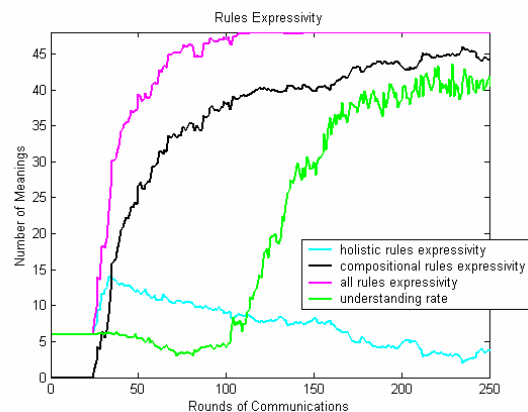
No order → random order → dominant order

a) Using of compositional rules (Word/Phrase) triggers word order convergence.

b) Converge to 2 orders, one for “agent action” meanings, the other for “agent predicate patient” meanings.

c) Order distribution is random, no prerequisite convention is necessary.

### 2) Lexicon Convergence



1) **Understanding rate** (averagely, how many meanings that all random pairs of agents can understand each other using linguistic rules only) indicates the robustness of language: effective representation and understanding **without other information's help.**

2) Lexicon Convergence follows a **“S-shape” Phase transition** (Ke et al, 2002).

*In this model,*

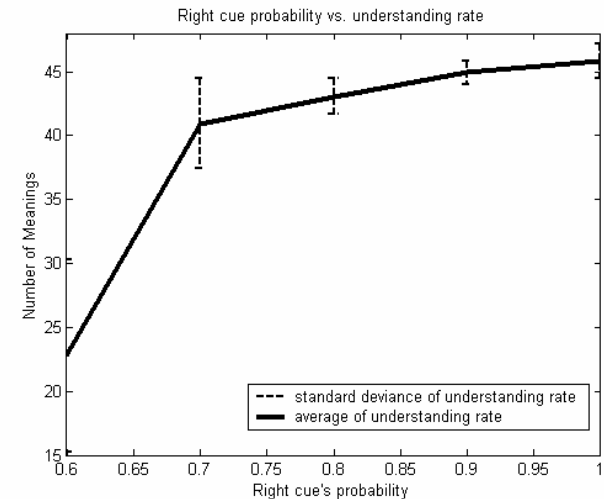
- a) A transition from **holistic units without internal structure to combination of compositional utterances with dominant order** can be traced.
- b) Two important processes: **Decomposition** related to holistic units and **Combination** related to compositional utterances.
- c) Both lexicon and syntax emerge during this process without prerequisite conventions; Converging processes are overlapping.
- d) **Mutual understanding** is the driving force for convergence of both lexicon and syntax

### 3) Robustness of acquired language

#### a) Cues are vague information

**Reliability (Probability for real meaning contained in one of the cues) test:**

- 1) Other information outside language is necessary in language evolution.
- 2) Even unreliable cues can still help language evolution.



#### b) Displacement test

**Without cues or under wrong cues, the acquired language is always reliable to express meanings not presented by environment and robust enough to withstand the wrong cues.**

*The acquired language based on this vague information is independent of this vague information and effective in representation and communication.*

## 4) Synonymy and Homophony

### *Easy emergence of synonym and homophone*

1) No explicit access to other agent's language

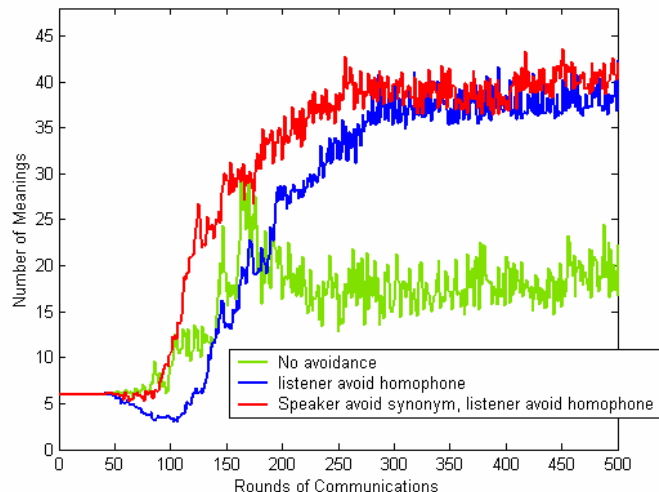
e.g., randomizing inexpressible meanings of an individual causes synonyms and homophones in language of the whole group)

2) Flexible storing recurrent patterns as new rules

### *Avoidance strategies are necessary (Homophone and synonym avoidance)*

Linguistic communication exchanging utterances and exclusive mappings of utterance and meaning, especially in meanings interpretation, are required when no non-language information's assistant is available.

Understanding rate comparison of with and without synonym and homophone avoidance



1) *Without synonym avoidance and homophone avoidance, understanding rate is very low, few meanings can be clearly understood.*

2) *Only certain degree of homophone avoidance can achieve a high understanding rate.*

## *Is homophone avoidance applicable in real language evolution?*

**Possibility:** Humans can distinguish different meaning items and utterances

**Necessity:**

- 1) Assuming both meaning items and phonemes are gradually increased, in the early stage of language evolution, selecting syllables from these phonemes to map meaning items would have easily caused homophone.
- 2) Linguistic communication without cues are common.
- 3) Some experiments show that recognition of spoken homophones depends on interaction of contextual and lexical information in sentence (Ping Li, 2001). Homophone avoidance might be a primitive strategy used in language emergence to avoid ambiguity of language.

## *Why synonyms/homophones still exist in modern language?*

- 1) **Syntax can allow homophones/synonyms in different categories.** e.g. train, well.
- 2) **Other Information (Context/Conditions/Written language) allows synonyms or homophones.** e.g. pitcher, contemporary.
- 3) Some synonyms or homophones can be frequently used in different location, and coexist very well

e.g. subway/tube, faucet/tap; /ai/, I/eye/aye

## 5) Full scale decomposition or not

Residual utterance after extraction of recurrent patterns can still be mapped as word/phrase rules for future confirmation or decomposition.

---

```
no use residue:
avgrate highrate lowrate stdrate avgcom highcom lowcom stdcom
45.8356 48      40.7556 2.7169 278.1000 98      439      125.2131
use residue:
avgrate highrate lowrate stdrate avgcom highcom lowcom stdcom
46.1856 48      41.9111 2.0061 246.8000 94      448      121.3451
```

**Using residue's result: no significant difference.**

### ***Considering:***

- Utterances in real language communication are not always fully understood, residual part of partial understanding is useless.
- Memory capacity limitation restricts storing rules extracted from residual parts. (cf Kirby, 1998, 2002)
- even residual utterances are allowed to be stored as new rules, due to lack of recurrence, these rules are less reliable than rules extracted from recurrent patterns.

---

***Using recurrent patterns only or mainly is more realistic!***

## b) Communication factor (**Social Structure**) effect

**Popular agent:** agent has many communications with other normal agents.

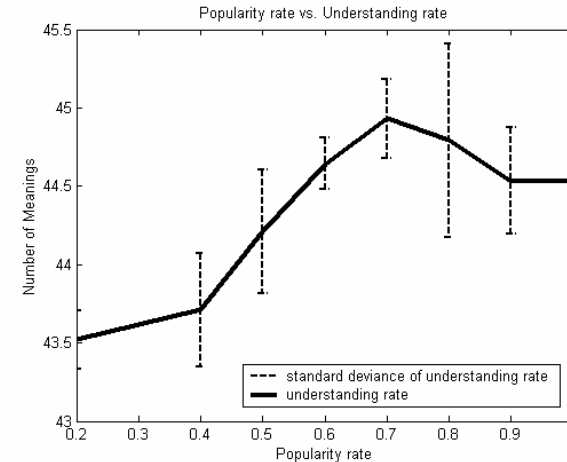
In view of network theory: popular agents are nodes with higher connectivity!

### 1) One popular agent

**Two types of communication:**

- 1) Popular agent  $\leftrightarrow$  normal agent (**Popularity Rate**)
- 2) Normal agent  $\leftrightarrow$  another normal agent

**Result:** *Partial “democracy” and partial “dictatorship” achieves a good convergence.*



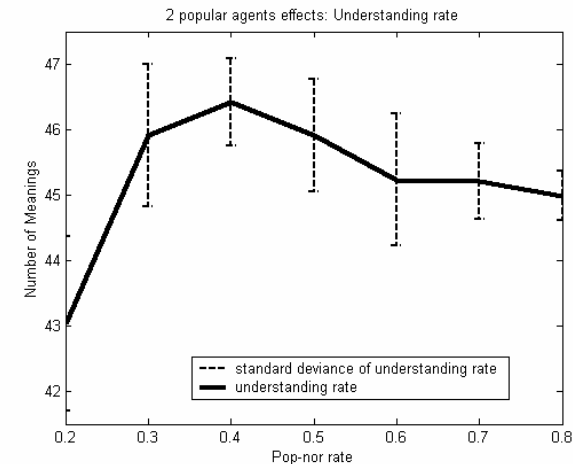
### 2) Two popular agents (2 nodes with high connectivity)

**Two types of communications:**

- 1) Popular agent 1  $\leftrightarrow$  Popular agent 2 (**Pop-pop rate**)
- 2) One Popular agent  $\leftrightarrow$  normal agent (**Pop-nor rate**)

**Pop-pop rate + Pop-nor rate = 1.0**

**Result:** *Best performance (highest understanding rate) happens at a certain Pop-nor rate.*



## Future work:

### 1) Theoretical discussion on Language emergence

Based on this model:

- Mainly a segmentation process.
- Decomposition and combination are overlapping in this model.
- Extracting meanings from environmental information can allow agents to acquire the concept of salient meanings, holistic or specific (events or objects).

*Language possibly emerged during an iterative process of cognizing, decomposing and combining. No clear distinction of word-originated or holistic unit-originated. both lexicon and simple syntax at protolanguage level emerged during this iterative process!* -- Extending Wray's "emergent" scenario  
*Models implementing all these processes are required!*

## 2) Interactions or constraints among speaking, listening and decomposition.

- To avoid further decomposition of existent word rules, rules generalization through detecting recurrent patterns should execute under the restriction of current rules in rule list.
- Interactions between speaking and listening.

## 3) Embedding, another important function of language besides representation and communication

## 4) In-depth study of social connection using network theory is a promising future.

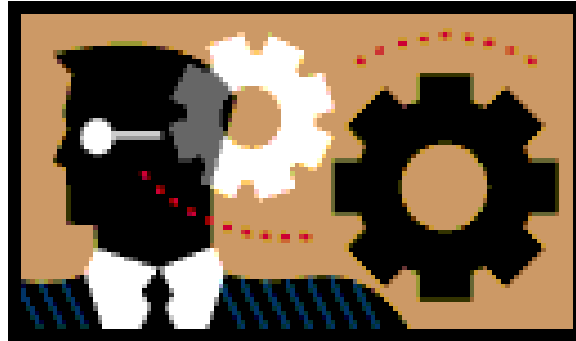
- a) The performance of language acquisition under different social structure.
- b) Can language acquisition trigger the forming of different social structure?

# Selected bibliography

- W. H. Calvin, Derek Bickerton. *Lingua ex machine: reconciling Darwin and Chomsky with the human brain*. Cambridge, Mass: MIT Press, 2000
- A. Wray. *Protolanguage as a Holistic System for Social Interaction*. *Language & Communication* 18: 47-67, 1998.
- J. H. Holland. *Exploring the Evolution of Complexity in Signaling Networks*. No. 01-10-062, Working Papers 2001 in Santa Fe Institute, 2001.
- S. Munroe, A. Cangelosi. *Learning and the Evolution of Language: The Role of Culture Variation and Learning Cost in The Baldwin Effect*. *Artificial Life* 8(4): 311-340, 2002.
- J. Batali. *Computational Simulations of the Emergence of Grammar*. *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press, Cambridge, MA, 1998.
- S. Kirby. *Learning, Bottlenecks and the Evolution of Recursive Syntax*. *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge University Press, Cambridge, MA, 2002.
- J. Ke, et al. *Self-Organization and Selection in the Emergence of Vocabulary*. *Complexity* 7(3): 41-54, 2002.
- S. Kirby. *Natural Language from Artificial Life*. *Artificial Life* 8: 185-215, 2002.
- M. Oliphant. *The Learning Barrier: Moving from Innate to Learned Systems of Communication*. *Adaptive Behavior* 7: 371-384, 1999.
- P. Li, et al. *Lexical ambiguity in sentence processing: Evidence from Chinese*. *Crosslinguistic Sentence Processing*: 111-129. Stanford, CA, 2002.

# Acknowledgements

- This work received many useful discussions and helpful suggestions from Professors *Guanrong Chen, Thomas Lee, John Holland, and William Wang*, as well as lab members, *Jinyun Ke* and *James W. Minett*. I hereby give them all my warmest thanks.



**Thank you !**