Computational Modeling on Language Emergence:
A Coevolution Model of Lexicon, Syntax and Social Structure[*]

Tao Gong[1] and William S-Y. Wang[12]

[1]*Chinese University of Hong Kong*
[2]*Academia Sinica*

In this paper, after a brief review of current computational models on language emergence, a multi-agent model is introduced to simulate the emergence of a compositional language from a holistic signaling system, through iterative interactions among heterogeneous agents. A coevolution of lexicon and syntax (in the form of simple word order) is tracked during communications with indirect meaning transference, in which the listener's comprehension is based on interactions of linguistic and nonlinguistic information, and the feedback is not a direct meaning check. In this model, homonymous and synonymous rules emerge inevitably, and a sufficiently developed communication system is available only when a homonym-avoidance mechanism is adopted. In addition, certain degrees of heterogeneity regarding agent's natural characteristics and linguistic behaviors do not significantly affect language emergence. Finally, based on theories of complex networks, a preliminary study of social structure's influence on language emergence is given, and a coevolution of the emergence of language and that of simple social structure is implemented.

Key words: language emergence (phylogenetic), computational modeling, coevolution (lexicon & syntax), coevolution (language & social structure)

1. Introduction

Emergence, used in linguistics, has two distinct senses (Wang 1991a, Holland 1998). Ontogenetically, emergence refers to the process whereby an infant acquires language from its environment; phylogenetically, emergence refers to the process whereby our species, *Homo sapiens*, made the gradual transition from prelinguistic communication, perhaps not unlike those of our ape contemporaries, to communication with languages of the sort we use today. Our concern in this paper is exclusively with

the phylogenetic emergence of language at the **macrohistory level** (Wang 1991b). This area is also referred to as **Glossogenetics** (Grolier 1981), or **Language Origins** (Ruhlen 1994).

A major stimulus for research in this area was the paper published by Hockett (1960), where the idea of **design features** was introduced, and language was compared with various forms of animal communication. Another significant stimulus was provided by a major conference held in New York City, sponsored by the New York Academy of Sciences (Harnad et al. 1976). That this area has rapidly grown into a powerful magnet for interdisciplinary research can be seen in the numerous anthologies published in the past dozen years, from Hawkins and Gell-Mann (1992) to Minett and Wang (in press).

Many disciplines have contributed significantly to our knowledge on this topic. For example, anthropologists have uncovered more and more fossils of our ancestors, particularly in northern Africa. From these discoveries, we may conjecture that our species evolved into its modern form about 160,000 years ago, an important landmark for dating language emergence. This date matches the time range as some language-related developments in our genes. Another landmark must be placed at around 50,000 years ago, when cultural achievements blossomed in the form of stone tools, art forms in sculpture and cave paintings, and burial sites. Many studies by population geneticists have given us some baselines for the earliest major human migrations across water. All these indirectly indicate the gradually enriched and refined communicative abilities of our ancestors. In addition, the beginning of the twenty-first century is witnessing the coming together of molecular genetics and neuroscience (Marcus 2004). The integration of the new knowledge here will have important bearing on such age-old controversies as whether there is a 'language organ' (Wang 1984, Anderson & Lightfoot 2002), and whether language emerged **monogenetically** or **polygenetically** (Freedman & Wang 1996).

In this paper, instead of theoretical argumentation over or empirical investigation into primate communication and child-language acquisition (e.g., Hutchins & Hazlehurst 1995, Wagner 2000), we introduce another rapidly growing approach to study language emergence—**computational modeling**, as exemplified by several anthologies and reviews (e.g., Wang et al. 2004, Cangelosi & Parisi 2001, Christiansen & Kirby 2003, Wagner et al. 2003). Based on evolutionary and/or artificial life theories, such as **coevolution** or **self-organization** (e.g., Steels 1999, De Boer 2000), many models have been reported, covering various aspects of language emergence.

According to whether **agents** (language users) are situated in an "artificial world" and whether the communication acts use single or several unstructured tokens versus structured utterances composed of multiple tokens, most current models can be divided

into four types (according to Wagner et al. 2003):

1) **Nonsituated, unstructured** models (e.g., Ke et al. 2002);
2) **Nonsituated, structured** models (e.g., Smith et al. 2003);
3) **Situated, unstructured** models (e.g., Caine 1995);
4) **Situated, structured models** (e.g., Munroe & Cangleosi 2002).

Situated models place agents in an "artificial world". Besides linguistic communication, non-communicative interactions between agents and environmental entities, such as food or predators, can affect the environment and/or modify agents' internal state. However, in nonsituated models, agents only send and receive signals. The dynamics of the emergence of a communication system is the main focus of these models. Utterances used in structured models consist of smaller units for hearers to interpret. However, utterances in unstructured models have single units or consist of independent units obtained from multiple channels.

In the remainder of this section, we briefly review three computational models: a **vocabulary coherence model** (nonsituated, unstructured model), a **neural network model** (situated, structured model) and the **iterative learning model** (ILM) (nonsituated, structured model). Situated, unstructured models, basically considering animal communication systems, are not discussed. After that, the limitations of these models are discussed.

1.1 The vocabulary coherence model (Ke et al. 2002)

This model is developed from our group's first work on modeling language emergence (Wang & Ke 2001). Several simulation models are reported to show the **emergence** (convergence) of a coherent vocabulary through self-organization in a population. Human language is assumed to start from a consistent set of mappings between inseparable meanings and utterances (**M-U mappings**). These mappings, referred to as early vocabularies, are considered to be the result of conventions established among a population of agents. Each agent has his own way of naming a set of objects, and may concentrate only on his own communication performance with other agents. Without any explicit or implicit design, a consistent common vocabulary can be conventionalized as an emergent property of the population.

In these models, each agent has his own speaking and listening matrices for M-U mappings (see Figure 1(a)), containing numbers that indicate the probability of correlative M-U mappings. In production, according to his speaking matrix, the speaker encodes the meaning into the utterance of the M-U mapping, the probability of which is

the highest in that meaning's row. In comprehension, according to his listening matrix, the listener decodes the signal into the meaning of the M-U mapping, the probability of which is the highest in that utterance's column. After a direct check of whether the speaker's encoded meaning matches the listener's decoded one, a self-organization mechanism adjusts probabilities of the chosen M-U mappings in their speaking/listening matrices. If they match, the probability of the chosen M-U mapping in the speaker's speaking matrix is increased and other probabilities in the same row are decreased. In addition, the probability of the chosen M-U mapping in the listener's listening matrix is increased and other probabilities in the same column are decreased. Otherwise, an inverse adjustment is executed. All probabilities are randomly initialized at the beginning of the simulation. After recursive interactions, a coherent vocabulary "emerges" (see Figure 1(b)(c)).

Probabilistic Speaking and Listening Mapping Matrices of Agent S

	u_1	u_2	u_3
p_{ij}			
m_1	0.3	0.4	0.3
m_2	0.4	0.55	0.05
m_3	0.7	0.2	0.1
q_{ij}			
m_1	0.1	0.3	0.6
m_2	0.5	0.3	0.3
m_3	0.4	0.4	0.1

(a)

An Example of Convergent Vocabulary with Incompatible Speaking and Listening Matrices

	u_1	u_2	u_3
p_{ij}			
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1
q_{ij}			
m_1	0	1	0
m_2	1	0	0
m_3	0	0	1

(b)

An Example of a Stable State: All Agents Have the Same Speaking and Listening Matrices

	u_1	u_2	u_3
p_{ij}			
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1
q_{ij}			
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1

(c)

Figure 1: (a) Agent S's Speaking/listening matrices; (b) (c) Coherent vocabulary examples. (From Ke et al. 2002)

The convergence process is tracked by four measures:

1) **Similarity of the mapping matrices** (SI), measuring the similarity between the vocabularies of two agents, is defined as the sum of the differences between correlative elements in both their speaking and listening matrices;

2) **Individual convergence rate** (IC), measuring the degree of consistency of an individual's speaking and listening mappings, is defined as the proportion of correlative elements in each of the two matrices that are smaller than a certain threshold;

3) **Population convergence rate** (PC), an index of the population convergence, is the sum of communicative consistency between all possible pairs in the population;

4) **Convergence time** (CT), the number of interactions taken for PC to reach a certain threshold above which the population is considered to have converged.

Figure 2(a) shows that the convergence process follows a **phase transition**, a sudden increase of SI and IC. Phrase transition has been discussed as a common emergence pattern in physical, biological, and social systems. In addition, without **self-talk** (interaction with oneself), there seems to be an optimal population size for the smallest CT (see Figure 2(b)).
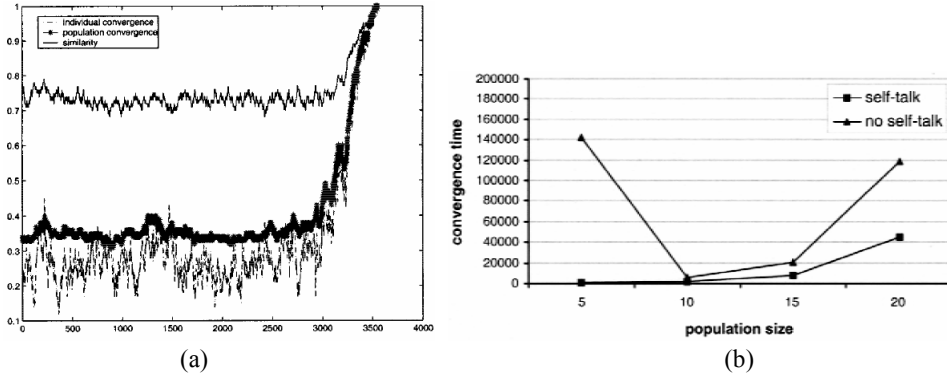
(a)   (b)

Figure 2: (a) Convergence process under conditions: P_s (population size) = 10, M (number of meaning) = U (number of utterance) = 3, Δ (adjustment magnitude in self-organization) = 0.2. An abrupt phrase transition is observed around 3,000 interactions; (b) Relationship between P_s and CT. Without self-talk, an optimum population size is observed; with it, a nonlinear increase of CT is observed. (From Ke et al. 2002)

1.2 The neural network model (Munroe & Cangelosi 2002)

Munroe and Cangelosi, using a neural network, implement a mushroom-foraging model to demonstrate how learning and natural selection interact under different conditions. It is a "situated" model; all agents are foraging in an artificial world with poisonous and edible "mushrooms" and different actions should be executed before eating different edible mushrooms. Correctly eating edible mushrooms can increase the **fitness** of that agent.

The structure of the neural network, simulating the internal state of each agent, is shown in Figure 3(a). The input layer includes the mushroom's visual property and linguistic utterance to describe the mushroom; language is assumed as a sensor to collect information. The output layer includes an action part, relating language communications with effective actions, and linguistic output, which can be used to train other agents. Two **clusters** (combined nodes, *winner-take-all* is executed inside the cluster) in language input and output are set up beforehand, which separate the mushroom type and appropriate actions.

The input and output of the neural network are signals of fixed length. The linguistic competence is stored as **connection weights** in connections among different layers. After a generation's foraging, twenty expert foragers are chosen as parents based on their fitness. Then, through asexual reproduction, each of them produces five offspring, who copy their parent's initial connection weights. The mutation operation in *Genetic Algorithm* (GA) (Holland 1995), a random change of some offspring's connection weights, introduces the variance. The performance of this new agent's **genome** is

chosen based on its fitness in future foraging.

The connection weights are adjusted during **lifetime learning**—a series of training tasks. In Stage 1 (foraging process), without linguistic communication, agents judge the edibility of mushrooms only by the visual information of the mushroom encountered in the artificial world. After several generations, Stage 2 begins, in which linguistic communication is allowed. The twenty expert foragers are carried over into the next generation as "teachers". Each new agent continues to forage as before, but lacking access to visual features most of the time. However, an additional linguistic input is always given by the parent. Lifetime learning is executed through three tasks during each cycle of parent-child interaction (Figure 3(b)): In task 1, the parent sends out a signal containing its own verbal description of the food closest to the child. The child uses this linguistic information to decide on an action vis-à-vis the mushroom. In task 2, the child performs a naming task in which the perceptual property of the food is available and a linguistic description of it is required. **Back-propagation** (adjustment of connection weights according to the difference between the child's linguistic input from its parent and the child's linguistic output) is executed to correct its linguistic output. In task 3, the child performs a linguistic imitation task in which it uses only its parent's description and reproduces its own linguistic description as output; once again, back-propagation is applied.
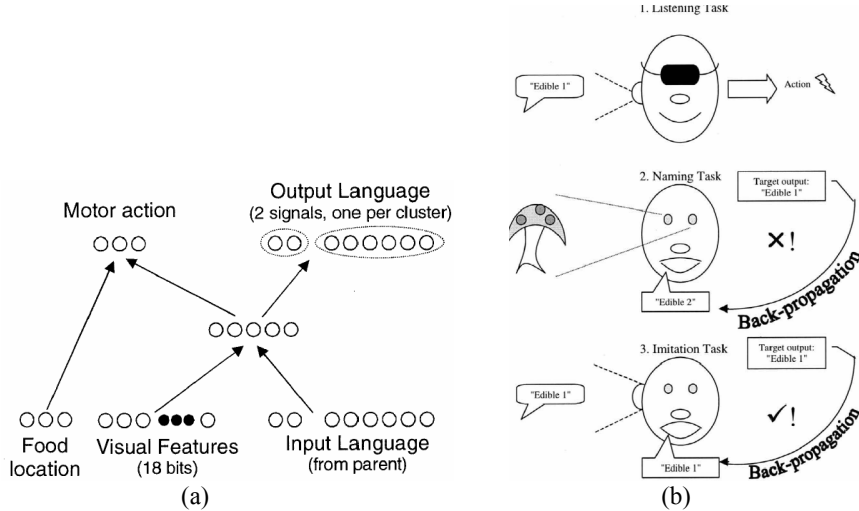

Figure 3: (a) Neural network structure; (b) The learning scenario. (From Cangelosi & Parisi 2002)

Based on visual information only, after foraging in Stage 1, agents can achieve the ability to distinguish different types of mushroom. In addition, with the allowance of language communication, this ability can be achieved within a shorter time. Meanwhile,

a simple syntactic structure, "object action", "emerges" (the preset two clusters already assume the compositional structure), with one node cluster for distinguishing mushroom type, and the other for different actions towards different types of edible mushroom. (See Figure 4).
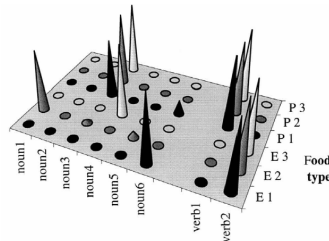

Figure 4: Compositional structure. Averaging 100 agents within one generation, cone height corresponds to the probability of using that node: greatest height corresponds to 100% use of that node among the whole population. (From Cangelosi & Parisi 2002)

Furthermore, this model also tests the **Baldwin Effects** (Baldwin 1896) from two aspects:

1) **Learning Cost**, whether a fitness penalty is given for eating poisonous mushrooms;

2) **Cultural Variation** (Language Distortions), when recording each generation's twenty expert foragers' language to train offspring, whether a random change of this language is added.

Through testing the difference between the acquired language and that of the parents and analyzing the efficiency of the language acquisition, this model shows that: learning cost can make agents gradually assimilate in their genomes some explicit features (e.g., lexical properties) of the specific language exposed to them; under cultural variation, Baldwinian processes cause the assimilation of a predisposition to learn any language exposed to them.

1.3 The Iterative Learning Model (ILM) (Smith et al. 2003)

Kirby (2002b) presents the ILM (Figure 5) to study the emergence of compositional languages from holistic signaling systems through vertical transmission of successive language learners. In a **holistic signaling system**, a signal stands for the meaning as a whole, with no subpart of the signal conveying any part of the meaning in and of itself. In a **compositional language**, the meaning of a signal is a function (combination) of the meaning of its parts (Krifka 2001).
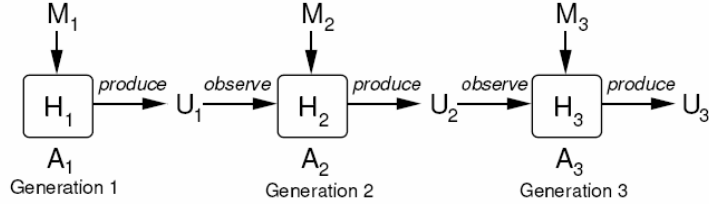
Figure 5: Iterative Learning Model (ILM). Linguistic competence (H_i) regulates an individual (A_i) of generation i to express certain meanings (M_i) with certain utterance (U_i). These utterances are also the primary linguistic data exposed to individuals in generation $i+1$. In this model, each generation has only one individual and learning happens only between individuals in successive generations. (From Smith et al. 2003.)

Smith et al. (2003) extend this ILM. In their model, the <u>meanings</u> are represented as points in an F-dimensional space where each dimension has V discrete values. The <u>utterances</u> (signals) are represented as strings of characters of length 1 to l_{max} and the characters w_i are drawn from an alphabet set Σ.

(1) $\quad M = \left\{ (f_1 f_2 ... f_F) : 1 \leq f_i \leq V \text{ and } 1 \leq i \leq F \right\}$

(2) $\quad U = \left\{ w_1 w_2 ... w_m : w_i \in \sum \text{ and } 1 \leq l \leq l_{\max} \right\}$

Components of meanings/utterances are vectors, each feature of which either has the same value as the meanings/utterances, or a wildcard (*). **Holistic components** in meanings/utterances have no wildcards (e.g., meaning (123) or utterance (ac)). **Compositional components** are vectors with a wildcard in a certain dimensional space or location (e.g., meaning (1*3) or utterance (a*)).

An **associative network** is used to store M-U mappings. Every crossing point represents a lexical mapping of a meaning component and an utterance component, and an associative weight is stored to indicate the possibility of this mapping. This associative network covers exhaustively all mappings between meaning and utterance components, and all associative weights are initialized at zero.

Adjustment of the associative weights happens in learning, based on the acquired M-U mapping from the agent in the previous generation. For example, in Figure 6(a), the agent hears M-U mapping: <(21), ab>. Then, related connection weights are either incremented (+) or decremented (−), and unrelated ones left unchanged. Competition of the associative weights happens in production through comparing the average weight of all applicable holistic or combinable compositional mappings to decide how to encode some meanings. For example, an agent wants to express the meaning (21) (see Figure 6(b)). There are 3 ways to express this meaning:

1) Using holistic utterance (gray circle with i);
2) Using compositional utterance (gray circles with ii) with one order;
3) Using compositional utterance (gray circles with iii) with another order.

Then the average strength of associative weights in all these conditions decides which utterance to use to encode the meaning.
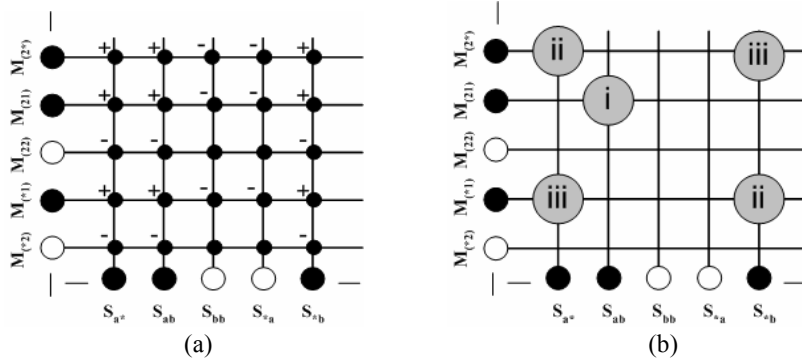

(a)   (b)

Figure 6: Learning and Production. (a) Learning: Large filled circles represent activated, related nodes (labeled with the component they represent) and small filled circles represent associative weights. (b) Production: The relevant connection weights are highlighted in gray. (From Smith et al. 2003.)

This model assumes that the **bottleneck in cultural transmission** is the stimulus for the emergence of a compositional language, since in vertical transmission only part of the previous generation's language is exposed to the next generation's learner. By simulating cultural transmission with and without this bottleneck, this model shows that a compositional language "emerges" (the associative network already assuming compositional structures) when there is a bottleneck in cultural transmission. In other words, compositionality is an adaptation by language allowing it to slip through the transmission bottleneck.

This model also discusses the relationships among meanings in the environment. Based on the number of these meanings in the environment (**density**) and whether they are closely related (**structured/unstructured**), four types of environment are introduced (see Figure 7). This model shows that maximum compositionality occurs when a language learner perceives his world as structured—when the objects in the environment relate to one another in structured ways—since a generalizable, compositional language is highly adaptive.
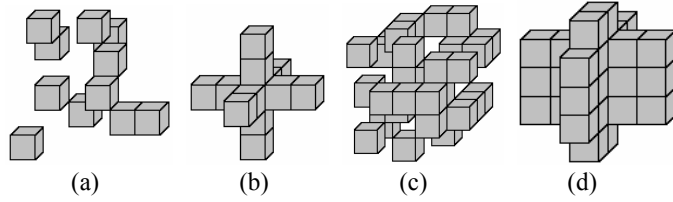
Figure 7: Four types of environments when $F = 3$ and $V = 5$. The selected meanings contained in each environment are highlighted in gray. (a) "Low-density, unstructured" environment; (b) "Low-density, structured" environment; (c) "High-density, unstructured" environment; (d) "High-density, structured" environment. (From Smith et al. 2003)

1.4 Discussions of current models

All these "emergent" models (according to Schoenemann 1999) discussed above view language evolution as **Complex Adaptive Systems** (CAS) (Holland 1986), and share several assumptions shedding light on real language development (e.g., self-organization strategies drive language evolution; language-specific syntactic predispositions are unlikely, etc.). However, there are still several limitations to these models.

First, most of these models (excluding Cangelosi & Parisi 2002) assume **direct meaning transference** in interactions among agents, i.e., the intended meanings, encoded in linguistic utterances produced by speakers, are always accurately available to listeners. It implies that accurate meaning transference through other channels is possible. However, if this were true, language as a communication medium would have been unnecessary, since intended meanings would always be available to listeners without linguistic communication. Moreover, it is obvious that there is no direct connection between speakers' production and listeners' comprehension—speakers always use utterances that they believe to represent intended meanings, and listeners always interpret utterances into the meanings that they believe these utterances express (Kirby 2002a). Other channels, such as pointing while talking or gestural/facial feedback, can only provide a certain degree of confirmation. Regarding pointing while talking, Quine's question (1960) is a good counterexample: If a child hears his mother use a word like *gavagai* as she points at a rabbit, what meaning should the child assign to it? The rabbit, some part of it, or any of a host of properties or details of that condition? Therefore, there is no telepathic access to other agents' minds. Furthermore, comprehension is not based only on linguistic information; nonlinguistic information provided by the environment, such as visual information, is important when linguistic information is inadequate for comprehension. It is worth studying the interactions between linguistic and nonlinguistic information in comprehension, especially in the early stage of language evolution where linguistic information is poor.

Second, these models either fail to model syntax (e.g., Ke et al. 2002), or implicitly assume that certain syntactic features are built in (e.g., Cangelosi & Parisi 2002), or do not adopt a coevolutionary view of the emergence of syntax and lexicon (e.g., Smith et al. 2003). From the evolutionary point of view, syntax also has its evolutionary origin. What the emerging process of syntax is and what its relation to the emergence of lexicon is should be considered.

Third, random interactions, adopted in these models, disregard social structure, which, as an intrinsic feature in human society, might influence language development (Romaine 1994) or may have coevolved with language use (Knight 1998). First, some social structures are formed based on biological and/or socio-economic factors irrespective of language. Such structures as kinship and social classes may place constraints on interactions between agents, and then have consequences in language acquisition/change. Second, language can be used to enhance social relationship, sense of solidarity/identity, etc., as reported in many sociolinguistic studies (e.g., Labov 1972). Mutual understanding of certain languages can be a factor to trigger changes in the social structure. Although sociological research has studied emergent structures based on stable or global factors, little research has touched upon the emergence of structure based on an evolving language. Therefore, it is worth seeing whether a global structure can be triggered during language emergence under certain social strategies, and what characteristics such structure might have. Besides, it is also worth studying whether and how different social strategies affect the emergence of language and that of social structure. Recently, the rapidly developing complex networks theory (Wang 2002, Barabási 2002, Newman 2003a) serves as an effective tool to explore these two aspects.

Fourth, many models are built with a homogeneous population, each member of which has identical natural characteristics and consistent linguistic behavior. However, sociolinguists have observed dramatic variations in speech communities (Romaine 1994), and studies on language acquisition have revealed basic differences in children's learning styles (Shore 1995). It is more realistic for computational models to take into account heterogeneity.

Addressing these limitations, a multi-agent model is presented to study language emergence. Two aspects of language emergence are considered: emergence of lexicon and emergence of syntax. There are two notable scenarios regarding syntactic origin:

1) A "**bootstrapping**" **scenario** (Bickerton 1998), which theorizes that a full language, originating from words, developed from word combination regulated by an innate syntax;

2) An "**emergent**" **scenario** (Wray 1998, 2002a), suggesting a holistic signal origin of language (cf. §1.3). Sporadic recurrent components in meanings and utterances are assumed to have triggered the **segmentation** of holistic signals, the break-down of

composite meanings and longer utterances into subpart meaning-constituents and subpart syllables; and this has led to the convergence of shared syntactic structures. With segmentation, holistic signaling systems are gradually changed into compositional languages. (See §1.3.)

From an evolutionary point of view, an "**emergent**" **scenario** appears to be more attractive and plausible. First, the protolanguage may have actually consisted of a number of holistic signals, similar to those found in primates and other animals such as birds (Hauser 1996), though their nature may be very different. According to the emergent theory, there may have been a stage of development in which early hominids began to detect recurrent patterns that appeared by chance in these holistic signals, which they then segmented into words. Second, in the "**emergent**" **scenario**, grammar is acquired through segmenting, detecting regularities in the meaning structures, and using sequences to combine words (Wray 2002a). This scenario presumes the existence of the sequencing ability, which makes it possible for certain sequences to get conventionalized and become dominant. The emergent process matches the general evolutionary principle that the syntactic feature also has its evolutionary origin. Grammatical rules in language are more likely to have emerged as a result of **conventionalization** due to language use, and not resulting from an innate, grammar-specific module (Schoenemann 1999). Syntax is assumed to have emerged from a pre-adapted cognitive capacity, also utilized in other cognitive processes (e.g., **sequencing ability** (Christiansen 2000)). Such a sequencing ability as a cognitive predisposition has been attested in other primates, as well as in pre-language infants. Finally, from holistic to analytic (Wray 2002b), language can emerge through iterative interactions among agents without any external guidance or innate language specific prerequisite. This developmental process has been attested in both first (Wray 1998) and second (Fillmore 1979) language acquisition in children.

Based on the emergent scenario, our model tracks a process from a non-syntactic, holistic signaling system to a syntactic, compositional language, along with which is the emergence of lexicon and syntax. Syntax, in this model, is in the form of dominant word order, which evolves from a set of optional word orders.

In addition, by introducing one type of nonlinguistic information, our model provides a solution to the unrealistic mind-reading adopted in many models, and implements an indirect meaning transference, in which the comprehension is decided by both linguistic and nonlinguistic information, and the feedback is not a direct meaning check. This scenario simulates realistic, multi-channel information processing in linguistic communication, and traces a process whereby linguistic information gradually achieves its advantage and reliability over other nonlinguistic information.

This model also assumes a heterogeneous population, considering heterogeneity to be a natural part of the agents' characteristics, as evidenced in memory capacity or

linguistic behavior. This makes it possible to study the effects of heterogeneity on language emergence.

Finally, the influence of social structure on language emergence is studied from two perspectives. One approach simulates language emergence under initialized, unchanged social structures, such as structures with popular agent(s) or those consisting of two-groups. The other approach is simulates a coevolution of the emergence of language and that of simple social structure based on certain social strategies, such as friendship or popularity.

The remainder of this paper discusses the model (§2), results of language emergence research (§3), the influence of social structure (§4), and several promising trends for the future (§5).

2. Description of the coevolution model

The current model simulates a language communication game, in which agents produce and comprehend utterances encoded with structured meanings. Language is indicated as M-U mappings, and stored as linguistic rules in agents. Through iterative communication, a common set of rules is shared among all agents, indicating the emergence of a common language. In what follows, the model's main components are briefly described, such as linguistic rules, agent ability, and the communication game.

2.1 Meaning, utterance and rule-based system

Meanings include single concepts such as objects or actions, e.g., 'dog'/'meat', 'bark'/'eat', and integrations of these concept constituents, such as "who do what (to whom)", e.g., "dog bark" or "dog eat meat". Two types of **integrated meaning** are considered, "predicate<agent>" (e.g., "run<dog>") and "predicate<agent, patient>" (e.g., "eat<dog, meat>" or "chase<fox, cat>"). **Predicate** usually is an action, **Agent** is the instigator of an action or a sentient causer, and **Patient** is an entity undergoing an action (Fromkin et al. 2003). Some integrated meanings are **transparent**, such as "bark<dog>" or "eat<dog, meat>", as these meanings are inferable from their constituents. It is obvious which is the agent and which is the patient, as long as the meaning constituents "eat<#, #>", 'dog' and 'meat', are identified. The whole meaning must be inferred as "eat<dog, meat>", instead of "eat<meat, dog>" in a normal situation. Other integrated meanings are **opaque**, such as "chase<fox, cat>", which cannot be inferred from their constituents. Although the meaning constituents of "chase<#, #>", 'fox' and 'cat', can be distinguished, it is not clear "who is chasing whom" without

further information, such as syntactic knowledge (e.g., word order) or environmental information (discussed later). Opaque meanings are often subject to misinterpretation.

In this model, the semantic space consists of forty-eight integrated meanings built upon twelve meaning constituents, each of which describes a certain environment event. Half of the meanings are transparent, half opaque. Agents only produce and comprehend integrated meanings in this semantic space.

Utterances consist of a string of syllables, and are combinable under the regulation of simple word order to map either constituent or integrated meanings.

Language is represented by a set of **linguistic rules**, comprising both **lexical rules** (M-U mappings + strength) and **word order rules** (sequencing orders + strength). Rule strength numerically indicates the frequency of successful uses of that rule. The agent's self-organization strategies include **rule competition** (decision-making during production and comprehension) and **rule adjustment** among available rules, both based on rule strength.

Lexical rules comprise **holistic** and **compositional** rules. Holistic rules are mappings between integrated meaning and holistic utterance. For example:

(3) "run<dog>"↔/a b c/ (0.4)

where the integrated meaning "run<dog>" and the utterance /a b c/ are associated with strength 0.4. Note that the mappings in all lexical rules are bidirectional, directing encoding in production and decoding in comprehension. Compositional rules include both **word** and **phrase** rules. Word rules are mappings between single meaning constituent and utterance. For example:

(4) "run<#>"↔/d e/ (0.3)   or
(5) "dog"↔/c/ (0.5)

where the pound sign (#) can be replaced by other compositional rule(s)'s meaning constituent(s) to form an integrated meaning together. Phrase rules are mappings between two constituents that do not form an integrated meaning and utterance. For example,

(6) "eat<dog, #>"↔/c * f/ (0.4).

where an integrated meaning can be formed with a word rule replacing the asterisk (*) by its utterance and the pound sign (#) by its meaning.

Word order rules cover all possible sequences to regulate utterances in

expressing integrated meanings. For example, to express "predicate<agent>" meanings, two orders are considered:

(7) "utterance for predicate before that for agent"(VS for sort), or
(8) "utterance for predicate after that for agent"(SV)

To express "predicate<agent, patient>" meanings, six orders are considered. For example:

(9) "utterance for agent first; that for predicate second; that for patient last" (SVO) or SOV, OSV, VSO, VOS, OVS

The word orders for "predicate<agent>" and "predicate<agent, patient>" meanings evolve independently. When producing and comprehending utterances with compositional rules, agents will choose word order rules with a higher strength among some potential orders rules.

In this model, following the <u>emergent</u> scenario, protolanguage is assumed as a holistic signaling system without dominant word order. Therefore, initially, all agents only share six holistic rules for expressing six integrated meanings, which contain all twelve meaning constituents. (We do not simulate the acquisition of semantic items.) All word order rules are initialized with the same strength for agents from which to choose randomly. A compositional language emerges if all agents share a set of common compositional rules as well as some dominant word order rules with high strengths.

2.2 Agent

Each agent uses a two-level memory system to store its lexical rules (Figure 8), which is inspired by the model of the **Learning Classifier System** (LCS; Holland 2001). The **buffer** stores "previous experiences" (M-U mappings obtained from previous communication(s), i.e., Com(i), Com(i+1), etc). The **rule list** stores "linguistic knowledge" (lexical rules) learned from "previous experiences". "Learning" takes place when the buffer is full, new lexical rules are generalized from those M-U mappings in the buffer, and updated into the rule list (discussed below). The buffer is then emptied to store new M-U mappings to be obtained in future communication(s). Lexical rules in the rule list, together with nonlinguistic information, are used in production or comprehension in future communications.
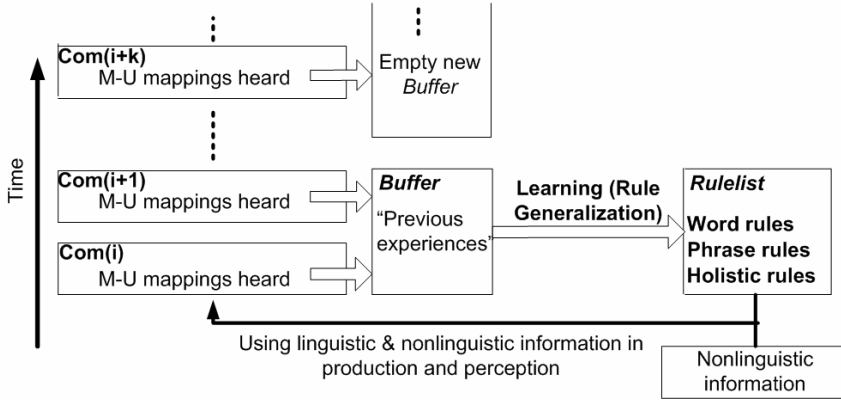
Figure 8: Two-level memory system. Com(i) is the index of communication.

There are two mechanisms for agents to acquire lexical rules:

1) **Random creation** in production. When encountering inexpressible integrated meanings, with certain possibility, the speaker may randomly select syllables to map either the whole integrated meaning, thus creating a holistic rule, or only those inexpressible constituent(s) in the encountered meaning, thus creating a compositional rule. The probability of random creation has an inverse proportion to the number of inexpressible constituents in the encountered meaning.

2) **Rule generalization** through detecting **recurrent patterns** in "learning". Figure 9 shows some examples of rule generalization. In ex. 2 of Figure 9, recurrent patterns are the identical meaning constituent(s) in the meaning parts (i.e., "fight<dog, #>") and the identical syllable(s) in the utterance parts (i.e., /c d * g/) contained in the two M-U mappings in the buffer. Recurrent utterance syllables do not require identical locations in M-U mappings. Agents focus on identical parts, and "don't care" about other parts indicated by "#" and "*". Once detecting the existence of these recurrent patterns, with certain possibility, the agent will create a compositional rule (a phrase rule, an M-U mapping, i.e., "fight<dog, #>"↔/c d * g/, plus an assigned initial strength) and update it into the rule list. Recurrence of some patterns triggers the segmentation of integrated meanings into meaning constituents and holistic utterances into substrings. If holistic mappings can be fully segmented into combinations of substrings, they are **fully decomposable**.
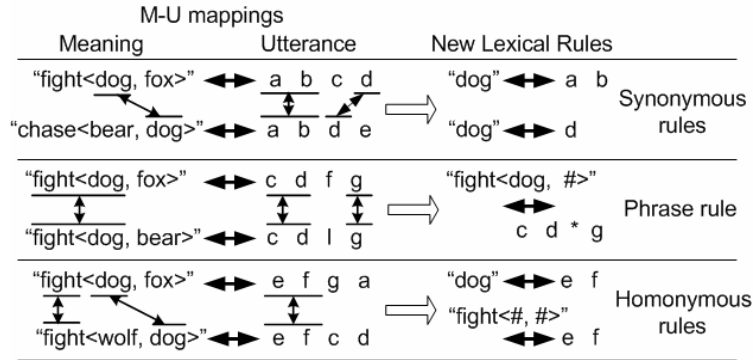
Figure 9: Rule generalization examples.

Synonymous and homonymous rules emerge inevitably during the rule acquisition. For example, the presence of multiple sets of recurrent utterance syllable(s) but only one set of recurrent meaning constituent(s) may cause one agent to learn many synonymous rules. (See ex. 1 in Figure 2.) Similarly, the presence of multiple sets of recurrent meaning constituent(s) but only one set of recurrent utterance syllable(s) may cause one agent to learn many homonymous rules. (See ex. 3 in Figure 2.) Synonymous rules increase the speaker's load for searching rules in production and take more space in the rule list. Homonymous rules may cause ambiguity in the listener's comprehension. Linguistic context can avoid such ambiguity. Besides, there are some internal avoidance mechanisms for the ambiguity caused by homonyms, suggested by some empirical research (e.g., **the principle of contrast** (Clark 1987)). Considering limited rule list size in our model, and lack of context since every meaning expressed by the speaker is independent, we adopt avoidance mechanisms in our model: after increasing the strength of a successfully used rule, decrease strengths of its synonymous and homonymous rules. A discussion of the necessity for avoiding homonymous rules is given in §3.3.

2.3 Communication

2.3.1 Nonlinguistic information—Cues

In order to show a comprehension process using linguistic as well as nonlinguistic information, we introduce cues as nonlinguistic information available to the listener during the communication. **Cues** describe the ongoing environmental events during the communication. In this model, cues are modeled as integrated meanings with some strength. For example:

(10)　　"chase<fox, dog>" (0.5)

It is obvious that cues, as semantic hints for comprehension, are not always reliable because the speaker may not always describe the ongoing events in the immediate environment and the listener may totally ignore certain events or pay attention to the wrong events. However, in the early stage of language development, there is a high probability that the speaker does describe the ongoing events and sometimes the listener can infer the speaker's intended meaning based on **shared attention** (Tomasello 2003). To simulate this, the **reliability of cues** (RC) is used to manipulate when the listener selects cues before comprehending the heard utterance; with what possibility the listener will acquire the cue containing the speaker's intended meaning; and in other situations, the listener simply selects a cue containing an integrated meaning in the semantic space. Multiple cues can be acquired by the listener simultaneously in one communication. Only referring to cues with the same strengths, the listener cannot tell which cue contains the speaker's intended meaning.

2.3.2 Communication game

Self-organization strategies in production and comprehension, together with interactions of linguistic and nonlinguistic information in comprehension, implement a communication game with indirect meaning transference.
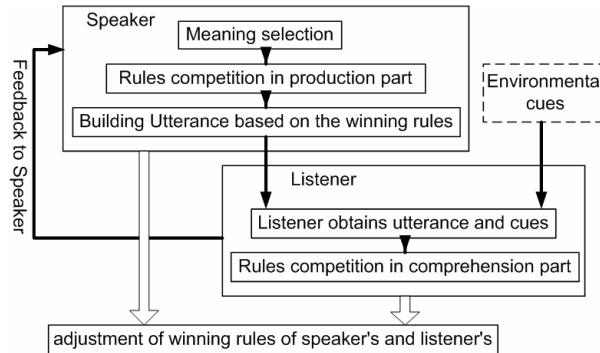

Figure 10: Communication game through indirect meaning transference.

Communication proceeds as follows. In the production, an integrated meaning is randomly chosen for the speaker to express. Several related or newly created lexical rules, together with appropriate word order rules, are activated in the speaker's mind. Then, among these activated rules, the speaker executes a rule competition strategy to select its winning rules which have the highest **combined strength for production**

(CS$_{production}$), defined by the following:

$$(11) \quad CS_{production} = average \left\{ \begin{array}{l} Strength(available \, lexical \, rules) \\ + \quad Strength(applicable \, word \, order \, rules) \end{array} \right\}$$

After that, the utterance, built up according to its winning rules, is sent to the listener. No production is allowed if the speaker has no lexical rules to encode the meaning with and the random creation fails. An example of production is stated in appendix A.

In comprehension, the listener receives an utterance, and occasionally, some **cues**. Then, the lexical rules in the listener's rule list, in which the utterances partially or fully match the received utterance, are activated. Similarly, the winning set of linguistic rules is chosen through rule competition in comprehension. Here in the comprehension, not only available linguistic rules, but also available cues contribute to the decision of the winning rules. The **combined strength for comprehension** (CS$_{comprehension}$) is defined by:

(12)

$$CS_{comprehension} = \begin{array}{l} LanguageWeight \left\{ average \left\{ \begin{array}{l} Strength(available \, lexical \, rules) \\ + \quad Strength(applicable \, word \, order \, rules) \end{array} \right\} \right\} \\ + \quad CueWeight \left\{ Strength(related \, Cues) \right\} \end{array}$$

where *LanguageWeight* and *CueWeight* are the proportions of linguistic and nonlinguistic information's contribution in the final decision. In order to simulate the transition from relying on nonlinguistic information to relying on linguistic information, we set both proportions at 0.5. In the early stage of language development, the listener tends to comprehend meanings supported by both linguistic rules and cues, since linguistic rules alone may not be sufficient to comprehend the utterance into an integrated meaning except for some holistic rules. Sometimes, due to the paucity of linguistic rules, unreliable cues would be the sole source for comprehension. No comprehension is allowed if no linguistic rules are available for inferring integrated meaning and either are cues available. An example of comprehension is stated in appendix B.

The meaning inferred from the listener's winning rules with the highest *CS$_{comprehension}$* is the listener's comprehended meaning. If this *CS$_{comprehension}$* exceeds a threshold, the listener sends a positive feedback to the speaker, indicating a strong confidence in the listener's comprehension. Otherwise, a negative feedback is sent, indicating that the listener cannot comprehend, or is not confident of his comprehension. Then under strong confidence feedback, both speaker and listener adjust rule strengths

of their activated rules to strengthen their winning rules; otherwise, an inverse adjustment, weakening their winning rules, is executed.

During the whole communication game, there is no direct connection between the speaker's production and the listener's comprehension, and the feedback only provides a certain degree of confirmation, since it is not a direct meaning check. The comprehension is determined by both linguistic and nonlinguistic information. The self-organization processes drive the emergence of a common language through iterative communication. Nonlinguistic information assists the comprehension and its unreliability may trigger a transition from relying on nonlinguistic information towards relying on linguistic information.

In this section, major components of our model have been described. In general, there are some differences in handling rules and communication scenario between this model and those discussed in §1. First, in this model, all M-U mappings (holistic or compositional) are acquired in communications, not set up beforehand as adopted by Smith et al.'s ILM model (2003). Adjustment of mappings is executed only among available mappings, not among all possible mappings as in the ILM model. Second, utterances in this model are flexible in length, instead of fixed in length as in Munroe and Cangelosi's mushroom foraging model (2002). And utterance in this model is connected with semantics through linguistic rules, while, in the mushroom foraging model, they are connected by weighted connections in a neural network; in the ILM model, they are connected with a weighted associated network. Third, this model focuses on the **horizontal transmission** (communication among agents in the same generation), which can supplement some language that is not acquired through vertical transmission, thus weakening the bottleneck effect. Finally, in our model, language is divided into lexical mappings and regulating sequences (syntax). This separation provides an opportunity to study the relation of syntax evolution and lexicon evolution.

3. Coevolution of lexicon and syntax

In the following section, we discuss the results of the model described above. The population size is ten, and a concurrent, iterative communication system based on the communication game described in §2.3 is adopted. In one round of communication, many (five in the population of ten) communications among different pairs of agents happen simultaneously, and in one communication, there are many instances of the communication game between speaker and listener. Reliability of cues (*RC*) is 0.7. For storage capacity, buffer size is 30 and rule list size is 40; for linguistic abilities, the possibility in random creation is 0.5 and the possibility for rule generalization is 0.5. In fact, results in populations more or fewer than ten agents are similar, but requiring more

or fewer rounds of communication to get a shared language.

The structure of this section is organized as follows. First, in order to test whether a common language can emerge in the population, determine what its features are, and trace the emerging process, we define several indices. Then, using these indices, we trace the emergence of a compositional language from a holistic signaling system and the coevolution of lexicon and syntax. After that, we discuss some factors that determine how sufficient the emergent compositional language is. Finally, by introducing a heterogeneous population, we demonstrate that a certain degree of heterogeneity does not significantly influence the emerging process.

3.1 Indices to test the performance

1) **Rule expressivity** (RE)—the average number of meanings that all agents can express:

$$(13) \quad RE = \frac{\sum_i \text{number of meanings that agent i can express}}{\text{number of agents}}$$

One holistic rule can only express one integrated meaning contained in its meaning part. Although one compositional rule can only express meaning constituent(s) instead of the whole integrated meaning, through combination, a limited number of compositional rules can express many integrated meanings. This **compositionality** endows human languages with the ability to use limited material to express infinite meanings, and *RE* can trace the transition from holistic rules to compositional rules in the agents' rule list.

2) **Understanding rate** (UR)—the average number of meanings understandable to every pair of agents in the group based on linguistic information only:

$$(14) \quad UR = \frac{\sum_{i,j} \text{number of understandable meanings between agent i and j}}{\text{number of all possible pairs of i, j}}$$

Models using direct meaning transference only trace an increase of *RE* of the emergent language, but do not test whether these linguistic expressivities are reliable and can be accurately comprehended. Considering the important linguistic characteristic of **displacement** (speech signals can refer to objects or events that are removed from the present in both space and time, but still can be accurately understood (Hockett 1960)), we use *UR* to evaluate such characteristics of the emergent language. When testing *UR*, there is no other assistance in comprehending the utterance but the linguistic

rules, whether the speech signals of the emergent language can be accurately understood based on linguistic rules only shows whether this language is reliable for agents to interchange information describing events not in the immediate time/space. A mature language should be a language with high *UR* (over 80%, say), instead of simply a high *RE*.

3) **Convergence time** (CT)—in certain times of simulations under the same conditions, the average number of rounds of communication required to achieve a mature language.

3.2 Coevolution process of lexicon and syntax

The coevolution of lexicon and syntax is summarized in Figure 11. Figure 11(a) shows the **Rule Expressivity** (RE) of both holistic and compositional rules; the decrease of the former and the increase of the latter show the transition from an initially holistic signaling system, to a compositional language. The **Understanding Rate** (UR), shown in Figure 11(a), undergoes an S-shape **phase transition** (Monasson et al. 1999), which is similar to the result of Ke et al.'s model (2002). Figures 11(b-c) show the emergence of dominant word orders from all possible sequential orders; each curve tracks the average strength of each of the eight word order rules across all agents. Two dominant word orders emerge, one for each of the two integrated meaning types. The chances for any order to become the dominant one are *a priori* equally likely; the random seed initialized in every communication can cause any word order to be the dominant one.
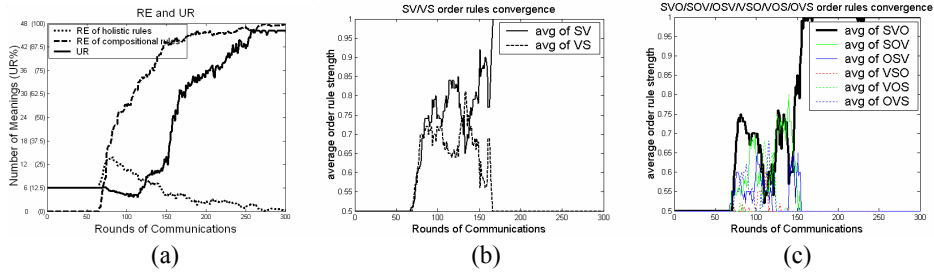

(a)   (b)   (c)

Figure 11: Coevolution of the lexicon and the syntax (a typical run). (a) Rule Expressivity (RE) and Understanding Rate (UR): Vertical axis is the number of meanings and understanding rate, horizontal axis is the number of rounds of communication. (b) Emergence of dominant word order for "predicate<agent>" meanings. (c) Emergence of dominant word order for "predicate<agent, patient>" meanings. In (b) & (c), the vertical axis is the rule strength; the horizontal axis is the number of rounds of communication.

The coevolution process typically proceeds as follows: In the beginning, agents only understand meanings produced by the six shared holistic rules. Then, more holistic rules emerge through random creations, which gradually increase the holistic rules' *RE*. Later on, the presence of recurrent patterns emerging by chance and acquisition of them greatly increases the compositional rules' *RE*, which initiates the transition from a holistic signaling system to a compositional language. Using compositional rules requires consistent word orders, so the convergence of the dominant word order begins, too. In this stage, due to random selection of word order rules, almost every order rules' strengths are increasing. The *UR* will gradually rely on compositional rules, although the use of compositional rules may cause some meanings that were initially understandable when produced by holistic rules to be misunderstood, causing a slight drop of *UR*. A similar drop of understandability is traced in the children's language acquisition (Fillmore 1979). Then, self-organizing mechanisms, such as rule competition and rule adjustment will get certain compositional and word order rules to win the competition and to be shared among almost all agents. Both the *UR* and those potential dominant word order rules' strengths increase sharply. The sharing of a set of common lexicons is finalized after the dominant word order emerges. In certain simulations, *UR* can reach very high, indicating a full share of lexicons sufficient to produce and comprehend all integrated meanings in the semantic space.

There are two driving forces for this emergence process:

1) **Mutual understanding**. Depending on whether the meanings produced are "accurately" comprehended (under positive feedback), a self-organizing process can adjust linguistic rules to drive the emergence or convergence of linguistic rules.

2) **Nonlinguistic information** (Cues). When linguistic rules are inadequate, cues are the only source for comprehension. When the listener acquires the cue containing the speaker's intended meaning, holistic/compositional rules used by the speaker to create the utterance may be accurately learned by the listener. In addition, cues can indirectly boost the strengths of compositional rules to gain advantage over holistic ones. During the calculation of combined strengths in comprehension, one holistic rule can only be assisted by one cue containing the holistic rule's integrated meaning. However, one compositional rule can be assisted by many cues containing that rule's meaning constituent. In some communications, although intended meanings are misunderstood, positive feedback is possible and compositional rules aided by related cues have greater chance of being boosted. Besides a self-organizing process, this is another way to cause some compositional rules with low strengths when created, to get their strength increased, and finally, shared among agents. A certain degree of misunderstanding, from this point of view, is not bad, but even necessary!

In all, Figures 11(a-c) show the coevolution of lexicon and syntax. Mutual

understanding requires not only common lexical rules but also a shared syntax to regulate utterances. The sharp increase of *UR* and strengths of the dominant order rules are almost **synchronized**: the use of compositional rules triggers syntax convergence, which in turn boosts lexicon convergence.

3.3 Homonym avoidance and reliability of cues (*RC*)

To acquire a mature language in this model, several internal and external constraints are necessary.

From an internal perspective, something has to be done to constrain the ambiguity caused by homonymous rules resulting from unreliable cues and the lack of context or any direct meaning check. Empirical research has detected homonym-avoidance mechanisms during child language acquisition (e.g., Clark 1987, Fillmore 1979). Children avoid mapping utterances already mapped to an extant meaning to novel salient meanings, especially those within the same semantic category ("agent", "patient", or "predicate"). For example, if a child has learned the word *apple*, when given an apple and a banana and told {"Bring me the banana."}, he always brings you the banana. Since the child already maps 'apple' with one object, he will not map the same word form with other objects so as to avoid possible ambiguity in the future, especially those in the same category, such as 'fruit' in this example.

According to this model, similar homonym avoidance is implemented in rule adjustment. After adjusting the strength of the winning form, any homonymous rules with their meaning parts within the same semantic category ("agent/patient" or "predicate") as the winning form, get their strengths adjusted inversely. This homonym avoidance allows the existence of homonymous rules in different semantic categories distinguishable by dominant word orders. Statistical results show that without homonym avoidance, the peak *UR*, under certain *RC* (say, 0.7), is much lower (12/48, 25%) than that for homonym avoidance (42/48, 87.5%).

Viewed externally, it is obvious that a high *RC* can accelerate language emergence by increasing the chances for accurate comprehension. However, even with highly reliable nonlinguistic information, without internal homonym avoidance, a mature language cannot be acquired. Figure 12 shows the average *UR* under different *RC* with and without homonym avoidance. The *UR* increases with the *RC* in both cases, but the *UR* with homonym avoidance is higher than that without such an avoidance mechanism. Besides, even if the cue were always reliable (i.e., *RC*=1.0), without homonym avoidance the peak *UR* is not that high. Checking the rule list, the existence of homonymous rules shared among agents still causes ambiguity when nonlinguistic information is absent. Therefore, with highly reliable cues alone, agents may not acquire a mature language.
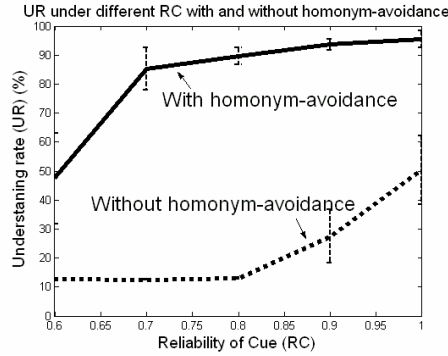
Figure 12: Average *UR* (Understanding Rate) under different *RC* (Reliability of Cue) with and without homonym avoidance. Horizontal axis is degree of *RC*; vertical axis is degree of *UR*. Results here are average *UR* in 10 simulations under the same conditions.

The preceding discussion suggests that for this model certain internal and external constraints are necessary for language emergence. Homonym avoidance is one of the internal constraints that can get around ambiguity, and certain avoidance mechanisms have been detected empirically (Clark 1987, Fillmore 1979). Reliable nonlinguistic information is one external requirement for sufficient language acquisition. However, reliable nonlinguistic information alone may not be enough for the development of a mature language. Similar results are found in language acquisition studies (Clark 2003). Ensuring that children always get accurate nonlinguistic information can usually increase learning efficiency. However, there are cases where such careful management of external information alone does not work.

3.4 Influence of the heterogeneous population

Here we shall analyze convergence time (CT) in a heterogeneous population to study the influence of heterogeneity on language emergence.

1) **Storage capacity of the buffer and rule list**. Heterogeneous capacity refers to the fact that various agents can have differently sized buffers and rule lists. A Gaussian distribution of different capacities with certain mean and variance is adopted to simulate this heterogeneity. Figure 13(a) shows the Convergence Time (CT) of heterogeneous buffer capacities with different mean values and a fixed rule list capacity of 40 (indicated by the dotted line). Figure 13(b) shows the *CT* of heterogeneous rule list capacities with different mean values and a fixed buffer capacity of 40 (the dotted line). The variance is 5 in either case. The *CT* of homogeneous buffer and rule list capacities under values equal to those mean values in the Gaussian distribution are also shown in the two figures (solid lines).
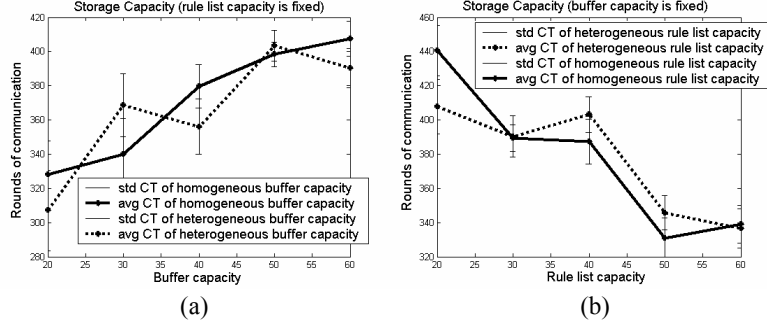
Figure 13: Heterogeneity: memory capacity. (a) *CT* (Convergence Time) of heterogeneous buffer capacity under a fixed rule list capacity; horizontal axis is different mean value of buffer capacity; vertical axis is the number of rounds of communication. (b) *CT* of heterogeneous rule list capacity under a fixed buffer capacity; horizontal axis is different mean value of rule list capacity; vertical axis is number of rounds of communication.

Buffer capacity affects rule generalization. Numerous slots in the buffer can store many M-U mappings, increasing the probability of recurrent patterns among them, as well as the chances of simultaneous generalization of many new rules. However, a bigger buffer needs more communications to fill, which may delay the rule updating rate. In Figure 13(a), the slow increases of the *CT* with the increase of the buffer capacity indicates that the second effect is a little bit stronger than the first, but not significantly so. For most reasonable buffer capacities (similar in magnitude to the size of semantic space), the emergent language has high *UR*s, even under heterogeneous conditions.

Rule list capacity determines rule storage. Increasing rule list capacity for more easily storing new rules may accelerate the convergence process. However, in our model, twelve-word rules are the minimum requirements to produce and comprehend all integrated meanings. The redundancy introduced by the large rule list can delay the convergence because those redundant rules may distract comprehension. In Figure 13(b), the slow decrease of the *CT* with an increase in rule list capacity indicates that the redundancy does slow down the convergence process, but this tendency is insignificant. For most reasonable rule list capacities (in this model, not lower than 12), the emergent language has high *UR*s, even under heterogeneous conditions.

2) **Linguistic behavior**. This includes the potential for random creation (**Creation Rate** or CR), which controls the rate for creating salient linguistic mappings, and the ability to detect recurrent patterns (**Detection Rate** or DR), which controls the rate for acquiring new rules from available M-U mappings. This heterogeneity is also simulated using Gaussian distribution. Figure 14(a)(b) shows the *CT* under heterogeneous *CR* with different mean values and a fixed *DR* of 0.5 and the *CT* under heterogeneous *DR* with different mean values and a fixed *CR* of 0.5. The *CT*s of homogeneous conditions

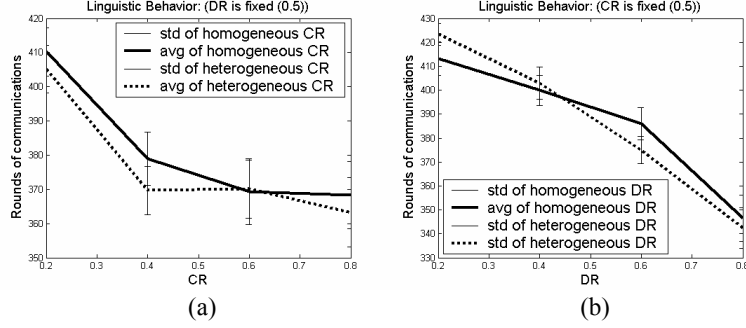are also indicated by dotted lines in the two following figures.

Figure 14: Heterogeneity: linguistic behavior. (a) *CT* under fixed *DR* and different *CR*; horizontal axis is different mean value of *CR*; vertical axis is the number of rounds of communication. (b) *CT* under different *DR* and fixed *CR*; horizontal axis is different mean value of *DR*; vertical axis is the number of rounds of communication.

The emergence of a common set of lexicons is impossible without random creation, since no salient linguistic materials for segmentation will be acquired. Similarly, lexicon emergence is impossible without detection of recurrent patterns, since there is no segmentation process at all. Besides, under a certain value of *DR*, if *CR* is too small, acquisition of new rules will be delayed by the insufficient number of randomly created salient linguistic materials. Similarly, under a certain value of *CR*, if *DR* is too small, few recurrent patterns are extracted and few holistic rules are decomposed. Both will delay the emergence of common compositional rules. Except for these extreme cases, high *UR*s are usually achieved even under heterogeneous conditions.

In this section, we prove that certain heterogeneities (e.g., storage capacity and linguistic behavior) cannot greatly affect the emergence of a mature language. This shows the self-organization processes in this model are **robust**; i.e., conditions with interference caused by either external noises or internal heterogeneous properties cannot significantly influence the emerging process of the compositional language.

4. Social structure's influence on language emergence

Social structures found in both human and primate societies may play certain roles in language evolution. Some models (e.g., Nettle 1999a, 1999b, Livingstone 2001) have touched upon this topic. In this section, we import **complex networks** (Albert & Barabási 2001, Newman 2003a) to study social structures, which view agents as vertices and communications among them as edges. Two aspects are discussed:

1) Language emergence under initialized, unchanged social structure;

2) Coevolution of language emergence and social structure based on the mutual understanding when using an evolving language.

4.1 Language emergence in initialized, unchanged social structure

One of the possible social structures in early hominids might be a structure with a **popular agent** (leader) (Labov 2001), which is the vertex with the highest connectivity. In this social structure, there are two types of communication.

1) Communications between the popular agent and other normal ones. **Popularity rate** (PR) is used to indicate the percentage of this type of communication in all communications.

2) Communications between two normal agents. The percentage of this type of communication is 1-*PR*.

PR indicates degree of global centralization. With an increase of *PR*, the popular agent will participate in more communications, and normal ones will gradually surround the popular agent. Figure 15 shows the Convergence Time (CT) and the Understanding Rate (UR) under different *PR*. There seems to be an optimal *PR* (0.7 in these simulations) where the highest *UR* is achieved.
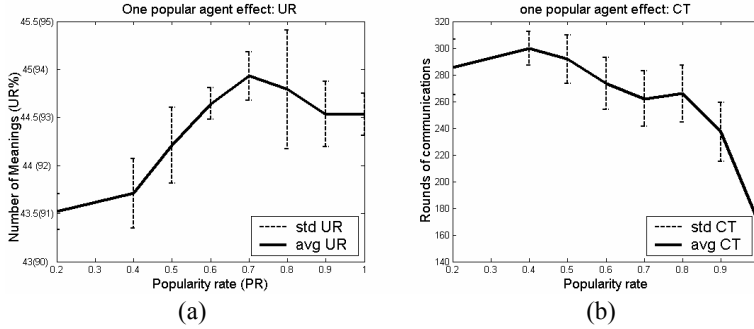

Figure 15: Popular agent effect. (a) *UR*. (b) *CT*. Horizontal axis is the *PR*; vertical axis is the number of rounds of communication. Simulation condition: 10 agents, 500*5 communications, *RC* = 0.7, 10 simulations.

The global centralization around some agent(s) has two effects:

1) **Acceleration effect**. Popular agents connect many normal ones, like a network **hub**. Centralization around it can increase the chances for information exchange among normal agents through the popular agent, which may accelerate the convergence of common rules in the whole group and reduce the *CT*.

2) **Deceleration effect**. Effective information transference between two agents (say, A_1 and A_2) requires a direct connection or a connection through a **stable intermediary**. (I.e., if the popular agent's internal rules do not change much, the information received by A_2 via the popular agent will not change much compared with the original information sent by A_1). However, with an increase of global centralization, every unpopular agent has a better chance of contacting the popular agent and influencing its rules. This makes the popular agent **unstable**; i.e., although the input information is the same, the output information may greatly differ from time to time, which may greatly affect the information transference and the convergence of common rules between A_1 and A_2 via the popular agent.

Balancing these two contradictory factors, the optimum UR occurs at an intermediate level of global centralization; absolute "democracy" (PR=1/number of agents) and absolute "dictatorship" (PR=1.0) will not achieve the best performance as measured by UR.

Another social structure concerns communication between two groups. Two types of communication are considered.

1) Communication among agents inside one group. **Intra-rate** represents the percentage of this type of communication.

2) Communication among agents in different groups. The percentage of this type of communication is 1-*Intra-rate*.

Intra-rate indicates the closeness among agents in one group. **Cross-group Understanding Rate** ($UR_{cross-group}$) tests the understandability among agents in different groups.

$$(15) \quad UR_{cross-group} = \frac{\sum_{i,j} \text{number of understable meanings between agent i and j}}{\text{number of all possible pairs of agent i (in group1) and agent j (in group2)}}$$

Figure 16 shows the UR and the $UR_{cross-group}$ within two groups under different **Inter-rates**, which gives a graphic demonstration that there is a transition from a high similarity of the languages emerged in two groups to a low similarity of them with the increase of **Intra-rate**, the chances for inside group communication.
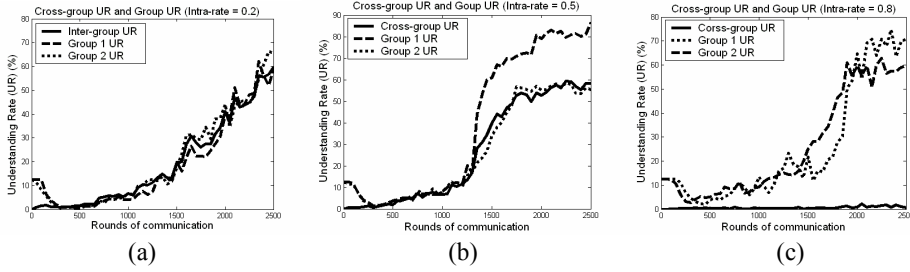
Figure 16: Communication between two groups (a typical run). (a) *Intra-rate* = 0.2; (b) *Intra-rate* = 0.5; (c) *Intra-rate*=0.8. Horizontal axis is the number of rounds of communication; vertical axis is *UR* (Understanding Rate). Simulation condition: 2 groups, 10 agents in each group, 500*5 communications, *RC* = 0.7.

In real human histories, communications between two groups usually happen between their leaders or representatives. If through few representatives, these limited communications provide a tendency towards **language divergence** (Nettle 1999c), while, if through leaders (popular agent), considering their accelerating effect stated above, there may actually be a tendency towards **language convergence**. In addition, at group boundaries, inevitable communications between members of two groups may introduce innovations into each other's languages, even developing an intermediate language. Research on communication among groups may furthermore shed light on **language change** (Fisiak 1995). However, the driving force of language change may not be simply restricted to contact alone, and other **socio-economic factors** should be considered as well (Mufwene 2004).

4.2 Coevolution of language emergence and social structure

Mutual understanding of language in use may be a factor in the adjustment of relationships among agents (Labov 2001), and the structure triggered accordingly may adversely influence language evolution. Our group began to move in this direction in Gong et al. (2004).
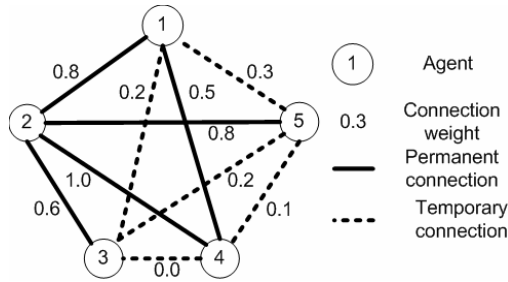

Figure 17: Fully-connected, undirected, weighted network.

A fully-connected, undirected, weighted network is introduced to indicate the social relationships among agents (see Figure 17). The connection weight, initially zero, is increased in **successful communications** where most of the feedback (say, over 80%) are positive; otherwise, it is decreased. Once the connection weight exceeds a threshold, a permanent edge is built up. Agents connected to each other through permanent edges have **linguistic friendship**, i.e., the propensity of mutual understanding between them. However, these permanent edges can still be broken if their connection weights fall below the threshold in future, similar to the breaking of friendship due to dissension. The number of permanent edges of one agent indicates its **linguistic popularity**, i.e., its propensity to communicate successfully with others. We also introduce a **local-view** assumption (drawn from Li & Chen 2003); i.e., in one communication an agent only views several members (local-view), instead of all in the whole group, and only communicates with someone in a local-view. These two local features (local-view and linguistic popularity) will separately adjust the relationship among agents during language evolution.

Two types of simulation are implemented to study the influence of these local features on the emergent social structure.

*Sim.*1 considers local features. In one generation, each agent selects agents into its local-view, those to which it is permanently connected having a higher chance of being chosen. Then agents communicate with a subset of agents in their local-view, preferring those having higher popularity. A new generation begins after all agents have executed this process.

*Sim.*2 does not consider local features. In one generation, agents randomly select agents into their local-view, and randomly communicate with some of them; i.e., agents randomly select other agents to communicate with during the entire time.

Several indices commonly used in analyzing complex networks are adopted to track the global structure.

1) **Average Degree** (AD) and **degree distribution** (P_k). Average degree is the average number of permanent edges per agent. The degree distribution is a histogram of the number of nodes with a given degree k.

2) **Average shortest path length** (L), the average shortest number of connections between every two different vertices.

3) **Clustering Coefficient** (C), the average fraction of pairs of neighbors (permanently connected) of one node that are also neighbors of each other.

$$(16) \quad C = avg(\sum_i \frac{E_i}{K_i(K_i - 1)/2})$$

Here K_i is the number of neighbors of node i; E_i is the number of edges that exist among node i's K_i neighbors. This index indicates the closeness of the network.
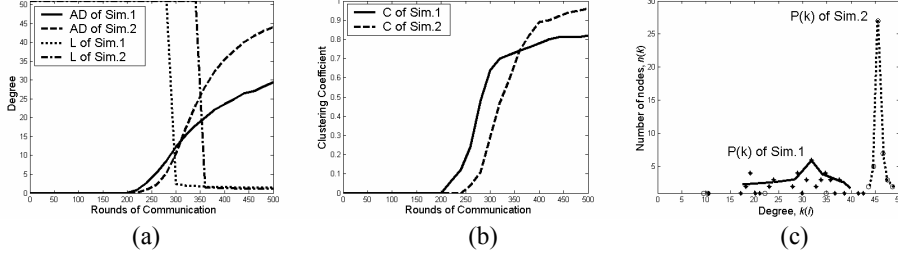

Figure 18: The emergence of social structure (two typical runs using two types of simulation). (a) *AD* and *L*. Horizontal axis is the number of rounds of communication; vertical axis is the number of degree. (b) *C*. Horizontal axis is the number of rounds of communication; vertical axis is the clustering coefficient; (c) P_k. Horizontal axis is the number of degrees; vertical axis is the number of agents that have a given degree. Simulation condition: 50 agents, Local-view = 10, Communication per agent = 5, 500 generations, 500*50*5 communications, *RC* = 0.7, 10 simulations.

During the emergence of language, a global social structure is triggered in both simulations based on the mutual understanding of an evolving language (see Figure 18) (the emergence of language follows a similar process shown in Figure 11). First, the Average Degree (*AD*) and the *C* in Figure 18(a), also following an *S*-curve, track the emergence of social structure. Due to the restriction of the local features, the *AD* and the *C* of Sim.1 are smaller, but having an earlier increase than those in Sim.2. Second, the high *C* and the low *L* in Figure 18(b) indicate that the emergent social structure in Sim.1 has **Small-world** (Watts 1999, 2003) characteristics, i.e., the average distance between nodes of the network increases slowly compared with the increase of the number of nodes. Third, seen from the degree distribution in Figure 18(c) in Sim.2, a network that is almost fully-connected emerges, with most agents having the same high degree. However, in Sim.1 the degree distribution is more uniform. This is because that local-view and the linguistic popularity only trigger a local centralization inside the local-view. Although members in agents' local-views are different at different times, the preference for permanently connected members will gradually restrict members in their local-views, and the preference for linguistic popularity inside the local view will trigger a local centralization among some members in their local views, agents within these local-views might have intensive connections with one another, but they don't frequently connect to outsiders except to those with higher connectivity. Therefore, the restriction of local information and local centralization prevents some agents' degrees from increasing greatly.
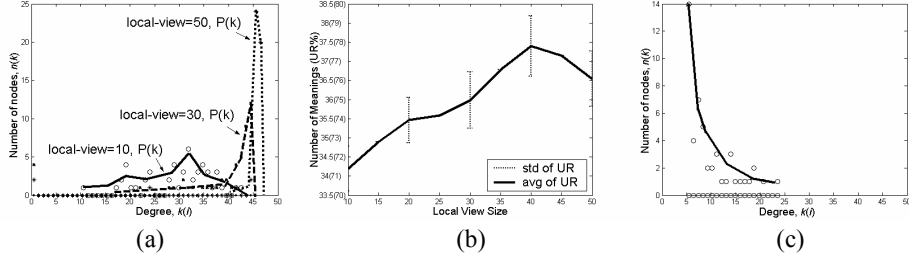
Figure 19: Local-view size effect: (a) P_k; (b) UR; (c) P_k of local world structure (based on the Li & Chen (2003) model).

In addition, different local features, e.g., different sizes of local-view, can affect the emergent social structure. With an increase in the local-view size, the influence of linguistic friendship is gradually reduced, and the local centralization is broken down. Seen from Figure 19(a), with the increase of local-view size, the degrees of most agents gradually increase. As for the emergent language, with the increase of the local-view size, the centralization is more global. An optimal local-view size for the peak UR is seen in Figure 19(b). This matches well the result in the structure with popular agent(s) as shown in Figure 15(a), presumably for similar reasons, though the structure here is gradually formed during the emergence of language, instead of initialized and unchanged.

Finally, the same social strategies, if based on a non-evolving (stable) language, can trigger a **Local-world** (Li & Chen 2003, see Figure 19(c)) or a **Scale-free** (Barabási 1999) structure (if the local-view is the whole group); i.e., a network lacks an intrinsic scale in some of its properties. However, no such structures emerge in our model. This shows a significant influence of the evolutionary characteristics of languages on the emergent social structure.

5. Conclusions and future directions

Computational simulation offers an efficient tool for the study of language evolution. It can evaluate linguistic theory, demonstrate evolutionary processes, and even raise questions that traditional empirical study may not explicitly answer. Several models have already touched upon many aspects of language evolution. In this paper, a coevolution model is presented, modifying some limitations of current models. It tracks a coevolution of lexicon and syntax in a concurrent communication system among a group of agents with a certain degree of heterogeneity. It simulates an interaction of linguistic and nonlinguistic information in comprehension and avoids direct meaning transference. Several internal/external strategies, required to acquire a mature language, are discussed. Based on this model, a preliminary study of the influence of social structure

on language evolution and vice versa is given, which connects to another important application of computation modeling: the study of complex networks. However, there are many important aspects of language and social structure still awaiting exploration.

On the linguistic side, the current model assumes a built-in semantic space. However, semantics, similar to syntax, is also an evolutionary feature gradually acquired by agents. Some computational models have already developed several scenarios on semantics. For example, in Steels et al. (2002), a **specification** mechanism is simulated; i.e., whenever ambiguity between two objects occurs, a new property is adopted by agents to distinguish them. Through gradual specification, agents acquire the properties necessary to distinguish different objects, and then build up some semantic concepts. In our model, a **generalization** process that extracts similarities among available M-U mappings is simulated. This generalization process can also be found during the acquisition of semantics. For example, through generalizing similarities between 'horse' and 'dog', agents can acquire the concept of 'animal'. The acquisition of semantics may be an integrative process of specification and generalization. The simulation of the emergence of semantics, along with the emergence of syntax and lexicon, can shed light on the relationship among semantics, lexicon, and syntax during language evolution. In addition, the semantic space in the current model is fixed. It is interesting to see whether the emergent process still follows a phase transition under an open-ended semantic space. Finally, this model does not touch upon more complex syntax, such as embedding or recursion. The building-up of complex syntax based on a simple sequencing ability and a built-in semantic space can shed light on the relationship between syntax and semantics.

From a social structure angle, several immediate extensions, based on the current framework, are required. For example, parameters like **betweenness** (Newman, in press) or **assortative degree** (Newman 2003b) can track whether the emergence of language in a community "grows and blossoms" from a sole source ("backbone" nodes having the highest betweenness) or "outbreaks' simultaneously from multiple sources. Meanwhile, besides linguistic communication, other factors like common interest or geographical limitation, can adjust relationships among agents and give agents **multiple social identities** (Watts 2003). Influence of these social identities on language evolution as well as other related phenomena, such as linguistic innovation and diffusion, is also worth studying.

Appendix A: An example of production

Suppose a speaker wants to express an integrated meaning: "fight<dog, fox>". According to his own rule list, he has three ways of expressing this meaning, as shown in Figure A.1. 1) By using a holistic rule, no word order is considered. 2) By using three word rules, all six possible word orders are applicable, so the strongest order rule (VSO(0.6)) is chosen. 3) By using a word rule and a phrase rule, the phrase rule's utterance part requires that only VSO and OSV are applicable orders, so that the strongest word order rule VSO(0.6) is chosen. In each condition $CS_{production}$ is calculated and $CS3$ (0.7) is the highest. Therefore, rules (boldface in Figure A.1) in that case are chosen as the speaker's winning rules. The utterance is built up accordingly, and sent to the listener.

Production Part		Meaning to express: "fight<dog, fox>"		
	Activated rules	Applicable word order rules	Combined Strength (CS)	Utterance created
1 holistic rule	"fight<dog, fox>"←→/a b/ (0.6)		CS1 = 0.6	/a b/
3 word rules	"dog"←→/b/ (0.8) "fight<#, #>"←→/c e/ (0.5) "fox"←→/g/ (0.2)	VSO (0.6)	CS2 = 1/2(1/3(0.8+0.5+0.2)+0.6) = 0.55	/c e b g/
1 word rule 1 phrase rule	**"dog"←→/b/ (0.8)** **"fight<#, fox>"←→/e * f/ (0.8)**	**VSO (0.6)** OSV (0.5)	**CS3 = 1/2(1/2(0.8+0.8)+0.6) = 0.7**	/e b f/
	Utterance built up: /e b f/			

Figure A.1: Example of rule competition in production.

Appendix B: An example of comprehension

Suppose a listener, in the comprehension (see Figure B.1), hears the speaker's utterance, /e b f/, and selects some cues: Cue1: "eat<dog, meat>" (0.5) and Cue2: "run<cat>" (0.5) (neither of which contains the speaker's intended meaning: "fight<dog, fox>"). Then, in the listener's rule list, lexical rules whose utterance parts partially or fully match the heard utterance are activated. He has three ways to decode the heard utterance and $CS_{comprehension}$ is calculated in each case (both *LanguageWeight* and *CueWeight* are 0.5). 1) By using a holistic rule, since Cue1's meaning matches that of the holistic rule, both the strength of the holistic rule and that of Cue1 are used to calculate $CS_{comprehension}$. 2) By using three word rules, no cue is related. An order (SVO) is detected by using these word rules to match the heard utterance. So, the strengths of the three word rules and that of SVO are used to calculate $CS_{comprehension}$. 3) By using a word rule, and since Cue2's meaning is related to this word rule's meaning, and no clear order rule can be detected, both the strength of the word rule and that of Cue2 are used to calculate $CS_{comprehension}$. $CS2$ (0.65) is the highest and rules (boldface in Figure

B.1) in that case are chosen as the listener's winning rules. The comprehended meaning is built up accordingly. Since *CS2* exceeds the confidence threshold (0.5), a positive feedback is sent back to the speaker, though the interpreted meaning does not match the intended meaning.

Figure B.1: Example of rule competition in comprehension.

References

Albert, Réka, and Albert-László Barabási. 2001. Statistical Mechanics in Complex Networks. arXiv:cond-mat/0106096.

Anderson, Stephen R., and David W. Lightfoot. 2002. *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge: Cambridge University Press.

Baldwin, James M. 1896. A new factor in evolution. *American Naturalist* 30:441-451, 536-553.

Barabási, Albert-László. 1999. Mean-field theory for scale-free. *Physica A* 272:172-182.

Barabási, Albert-László. 2002. *Linked: The New Science of Networks*. Cambridge: Perseus.

Bickerton, David. 1998. Catastrophic evolution: the case for a single step from protolanguage to full human language. *Approaches to the Evolution of Language: Social and Cognitive Bases*, ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight, 341-358. Cambridge: Cambridge University Press.

Briscoe, Ted. (ed.) 2002. *Linguistic Evolution through Language Acquisition*. Cambridge: Cambridge University Press.

Caine, Nancy G. 1995. Factors affecting the rates of food calls given by red-bellied tamarins. *Animal Behavior* 50:63-60.

Cangelosi, Angelo, and Domenico Parisi. (eds.) 2001. *Simulating the Evolution of Language*. London: Springer Verlag.

Christiansen, Morten H. 2000. Using artificial language learning to study language evolution: exploring the emergence of word order universals. *The Proceedings of Evolution of Language: The 3rd International Conference*, ed. by J. L. Dessalles, and L. Ghadakpour, 45-48. Paris, France.

Christiansen, Morten H., and Simon Kirby. (eds.) 2003. *Language Evolution.* New York: Oxford University Press.

Clark, Eve V. 1987. The principle of contrast: a constraint on language acquisition. *Mechanisms of Language Acquisition*, ed. by Brian M. MacWhinney, 1-33. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, Eve V. 2003. *First Language Acquisition*. Cambridge: Cambridge University Press.

De Boer, Bart G. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28: 441-465.

Fillmore, L. W. 1979. Individual differences in second language acquisition. *Individual Differences in Language Ability and Language Behavior*, ed. by Charles J. Fillmore, Daniel Kempler, and William S-Y. Wang, 203-228. New York: Academic Press.

Fisiak, Jacek. (ed.) 1995. *Linguistic Change under Contact Conditions*. Berlin, New York: M. de Gruyter.

Freedman, D. A., and William S-Y. Wang. 1996. Language polygenesis: a probabilistic model. *Anthropological Science* 104:131-37.

Fromkin, Victoria, Robert Rodman, and Nina Hymas. 2003. *An Introduction to Language*, Seventh Edition. Boston: Thomas, Heinle.

Grolier, Eric. (ed.) 1981. *Glossogenetics: The Origin and Evolution of Language*. Paris: Hardwood Academic Publishers.

Gong, Tao, Jinyun Ke, James W. Minett, and William S-Y. Wang. 2004. A computational framework to simulate the coevolution of language and social structure. *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, ed. by Jordan Pollack, Mark Bedau, Phil Husbands, Takashi Ikegami, and Richard A. Watson, 158-163, Cambridge: MIT Press.

Harnad, Stevan R., Horst D. Steklis, and Jane Lancaster. (eds.) 1976. *Origins and Evolution of Language and Speech*. New York: The New York Academy of Sciences.

Hauser, Marc D. 1996. *The Evolution of Communication*. Cambridge: MIT Press.

Hawkins, John A., and Murray Gell-Mann. (eds.) 1992. *The Evolution of Human Languages*. Redwood City: Addison-Wesley.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203:88-96.

Holland, John H. 1986. Escaping brittleness: the possibilities of general-propose learning algorithms applied to parallel rule-based systems. *Machine Learning: An Artificial Intelligence Approach*, ed. by Ryszard S. Michalski, Jamie G. Carbonell, and Tom M. Mitchell, 593-623. Los Altos: Morgan Kaufmann.

Holland, John H. 1995. *Hidden Order*. Cambridge: Perseus. 約翰‧霍蘭著，周曉牧、韓輝譯《隱秩序》(2000)。上海：上海科技教育出版社。

Holland, John H. 1998. *Emergence*. Cambridge: Perseus. 約翰・霍蘭著，陳禹等譯《湧現》(2001)。上海：上海科學技術出版社。

Holland, John H. 2001. Exploring the evolution of complexity in signaling networks. *Complexity* 7.2:34-45.

Hutchins, Edwin, and Brian Hazlehurst. 1995. How to invent a lexicon: the development of shared symbols in interaction. *Artificial Societies: The Computer Simulation of Social Life*, ed. by Nigel Gilbert and Rosaria Conte, 157-189. London: UCL Press.

Ke, Jinyun, James W. Minett, Ching-Pong Au, and William S-Y. Wang. 2002. Self-organization and selection in the emergence of vocabulary. *Complexity* 7.3:41-54.

Kirby, Simon. 2002a. Natural language from artificial life. *Artificial Life* 8.2:185-215.

Kirby, Simon. 2002b. Learning, bottlenecks and the evolution of recursive syntax. *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, ed. by Ted Briscoe, 173-205. Cambridge: Cambridge University Press.

Knight, Chris. 1998. Introduction: grounding language function in social cognition. *Approaches to the Evolution of Language—Social and Cognitive Bases*, ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight, 9-17. Cambridge: Cambridge University Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Li, Xiang, and Guan-Rong Chen. 2003. A local-world evolving network model. *Physica A* 328.1-2:274-286.

Livingstone, Daniel. 2001. The evolution of dialect diversity. *Simulating the Evolution of Language*, ed. by Angelo Cangelosi and Domenico Parisi, 99-117. London: Springer Verlag.

Marcus, Gary. 2004. *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought*. New York: Basic Books.

Minett, James W., and William S-Y. Wang. (eds.) (in press). *Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics*. Hong Kong: City University of Hong Kong Press.

Monasson, Rémi, Riccardo Zecchina, Scott Kirkpatrick, Bart Selman, and Lidror Troyansky. 1999. Determining computational complexity from characteristics "phrase transitions". *Nature* 400.8:133-137.

Mufwene, Salikoko S. 2004. Language birth and death. *Annual Review of Anthropology* 33:201-222.

Munroe, Steve, and Angelo Cangelosi. 2002. Learning and the evolution of language:

the role of culture variation and learning cost in the Baldwin Effect. *Artificial Life* 8.4:311-340.

Nettle, Daniel. 1999a. Using social impact theory to simulate language change. *Lingua* 108:95-117.

Nettle, Daniel. 1999b. Is the rate of linguistic change constant? *Lingua* 108:119-136.

Nettle, Daniel. 1999c. *Linguistic Diversity*. Oxford: Oxford University Press.

Newman, Mark E. J. 2003a. The structure and function of complex networks. *SIAM Review* 45.

Newman, Mark E. J. 2003b. Mixing patterns in networks. *Phys Rev. E67*.

Newman, Mark E. J. (in press). Who is the best connected scientist?—a study of scientific coauthorship networks. *Complex Networks*, ed. by Eli Ben-Naim, Hans Frauenfelder, and Zoltan Toroczkai. Berlin: Spring.

Quine, Willard V. 1960. *Word and Object*. Cambridge: MIT Press.

Romaine, Suzanne. 1994. *Language in Society, an Introduction to Sociolinguistics*. Oxford: Oxford University Press.

Ruhlen, Merritt. 1994. *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.

Schoenemann, P. Thomas. 1999. Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines* 9:309-346.

Shore, Cecilia M. 1995. *Individual Differences in Language Development*. Thousand Oaks: Sage.

Smith, K., H. Brighton, and Simon Kirby. 2003. Complex systems in language evolution: the cultural emergence of compositional structure. *Advances in Complex Systems* 6.4:537-558.

Steels, Luc. 1999. The spontaneous self-organization of an adaptive language. *Machine Intelligence* 15, ed. by K. Furukawa, Donald Michie, and S. Muggleton, 205-224. Oxford: Oxford University Press.

Steels, Luc, F. Kaplan, A. McIntyre, and J. V. Looveren. 2002. Crucial factors in the origins of word-meaning. *The Transition to Language*, ed. by Alison Wray, 252-271. Oxford: Oxford University Press.

Strogatz, Steven H. 2003. *Sync: The Emerging Science of Spontaneous Order.* New York: Theia.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition.* Cambridge: Harvard University Press.

Wagner, Kyle. 2000. Cooperative strategies and the evolution of communication. *Artificial Life* 6.2:149-179.

Wagner, Kyle, James A. Reggia, Juan Uriagereka, and Gerald S. Wilkinson. 2003. Progress in the simulation of emergent communication and language. *Adaptive Behavior* 11.1:37-69.

Wang, William S-Y. 1984. Organum ex machina. *Behavioral and Brain Sciences* 7.2:210-211.

Wang, William S-Y. (ed.) 1991a. *The Emergence of Language: Development and Evolution*. New York: W. H. Freeman.

Wang, William S-Y. 1991b. The three scales of diachrony. *Explorations in Language*, ed. by William S-Y. Wang, 60-71. Taipei: Pyramid.

Wang, William S-Y. (王士元), and Jinyun Ke (柯津雲). 2001.〈語言的起源及建模仿真初探〉[A preliminary study on language emergence by modeliing and simulation],《中國語文》2001.3:195-200。

Wang, William S-Y., Jinyun Ke, and James W. Minett. 2004. Computational studies of language evolution. *Computational Linguistics and Beyond: Perspectives at the Beginning of the 21st Century*, ed. by Chu-Ren Huang and Winfried Lenders. Taipei: Institute of Linguistics, Academia Sinica.

Wang, Xiaofan. 2002. Complex networks: topology, dynamics and synchronization. *International Journal of Bifurcation and Chaos* 12.5:885-916.

Watts, Duncan J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton: Princeton University Press.

Watts, Duncan J. 2003. *Six Degrees: The Science of a Connected Age*. New York: Norton.

Wray, Alison. 1998. Protolanguage as a holistic system for social interaction. *Language & Communication* 18:47-67.

Wray, Alison. 2002a. *Formulaic Language and the Lexicon*. New York: Cambridge University Press.

Wray, Alison. 2002b. Dual processing in protolanguage: performance without competence. *The Transition to Language*, ed. by Alison Wray, 113-137. Oxford: Oxford University Press.

Tao Gong
Department of Electronic Engineering
Chinese University of Hong Kong
Hong Kong
tgong@ee.cuhk.edu.hk

William S-Y. Wang
Department of Electronic Engineering
Chinese University of Hong Kong
Hong Kong
wsywang@ee.cuhk.edu.hk

語言產生建模仿眞：
一個詞匯、語法和社會結構交互演化模型

龔　濤 [1]　　王士元 [12]

[1] 香港中文大學

[2] 中央研究院

　　本文在簡略回顧了當前語言產生建模仿真的發展後，提出了一個多個體模型來模擬通過異質個體間的反覆交流，以詞匯爲主的語言如何從原始不可分信號中產生。該模型模擬了詞匯和語法（簡單的詞序）的交互演化，並模擬了間接語意傳輸：聽者在交流中通過處理語言和非語言方面的信息來理解語意，所採用的反饋毋需直接核對語意。獲得語言規則的過程中不可避免的產生同音和同義詞。避免同音詞機制的採用使有效語言交流系統的產生成爲可能。另外，個體間不同的自然屬性和語言行爲並不會顯著影響語言的產生。最後，從複雜網絡理論出發，本文初步研究了社會結構對語言產生的影響，並模擬了語言與簡單社會結構的產生同交互發展。

關鍵詞：語言產生，交互演化，異質，社會結構