

Coevolution of Lexicon and Syntax from a Simulation Perspective

TAO GONG,¹ JAMES W. MINETT,¹ JINYUN KE,² JOHN H. HOLLAND,³ AND WILLIAM S.-Y. WANG¹

¹Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong; and ²English Language Institute, and ³Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109

Received March 8, 2004; revised July 2, 2004, January 27, 2005; accepted April 25, 2005

Whether simple syntax (in the form of simple word order) can emerge during the emergence of lexicon is studied from a simulation perspective; a multiagent computational model is adopted to trace a lexicon-syntax coevolution through iterative communications. Several factors that may affect this self-organizing process are discussed. An indirect meaning transference is simulated to study the effect of nonlinguistic information in listener's comprehension. Besides the theoretical and empirical argumentations, this computational model, following the Emergentism, demonstrates an adaptation of syntax from some domain-general abilities, which provides an argumentation against the Innatism. © 2005 Wiley Periodicals, Inc. Complexity 10: 50–62, 2005

Key Words: language emergence; multiagent; coevolution; indirect meaning transference

1. INTRODUCTION

Language emergence has long been of interest to scholars from multiple backgrounds, such as linguistics, biology, psychology, and anthropology. Emergence in linguistics, has two distinct senses [1, 2]: ontogenetic emergence refers to the process whereby an infant acquires language from its environment, and phylogenetic emergence refers to the process whereby our species, *Homo sapiens*, made the gradual transition from prelinguistic communication to communication with languages of the sort we use today. There have been two well-known hypotheses regarding the phylogenetic emer-

gence of language: Innatism (e.g., [3–5]) believes that the human language faculty is determined by certain innate, language-specific features. This hypothesis has been the dogma in linguistics. Emergentism (e.g., [6, 7]) assumes that evolutionary forces drive the adoption of certain domain-general features into language and the development of language from primitive forms into modern complex, human-specific forms.

Although both hypotheses share an evolutionary perspective on human language, the crucial distinction between them is whether syntax, an important feature in human languages, results from innate, domain-specific features or gradually developed from some domain-general abilities. Regarding the syntax origin, there are two notable scenarios:

Correspondence to: Tao Gong, E-mail: tgong@ee.cuhk.edu.hk

1. A “bootstrapping” scenario with innate syntax [8], which conceives that a full language, originating from words, developed through word combination regulated by an innate syntax. This scenario follows the Innatism hypothesis;
2. A “formulaic” scenario without innate syntax [9], which suggests that language may have started from a *holistic signaling system*, where 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 [10]. Then, sporadic recurrent components in utterances and in meanings are assumed to have triggered the segmentation of holistic signals (i.e., break down longer utterances with integrated meanings into combinations of subpart syllables with meaning constituents) and then to have led to the convergence of shared syntactic structures. Through the segmentation, compositional languages gradually emerge, where the meaning of a signal is a function (combination) of its meaning constituents.

The “formulaic” scenario without innate syntax is supported by the following arguments. First, the primitive human cognitive system may not allow human to develop a clear concept of “the instigator of an action” or “the entity that undergoes an action” when they experience environmental events. Therefore, in that stage, signals in the primitive human communication system must have consisted of a number of holistic items, similar to those found in the communication systems of primates and other animals such as birds [11], although their nature may be very different. Second, detecting similar patterns and sequencing items are two domain-general abilities shared by human and other animals, though the actual level of these abilities is different. For example, the sequencing ability as a cognitive predisposition has been attested in prelanguage infants [12] as well as other primates. According to this, after human grasp compositional semantic concepts, there may have been a stage of development in which early hominids began to detect recurrent patterns in the holistic signals that encode compositional meanings and appeared by chance and segment these patterns into words. After that, the sequencing ability was adopted to regulate these words to express complex meanings, making it possible for certain sequences to become conventionalized into dominant word orders, i.e., the emergence of simple syntax. Finally, without any external guidance or innate language specific prerequisite, this developmental process, from holistic to analytic [13], has been attested in both first [14] and second [15] language acquisition in children. To summarize, this scenario provides an emergent origin of syntax from some domain-general abilities, and more linguists gradually believe that grammatical rules in language are more likely to have emerged as the result of conventionalization due to

language use, rather than as the result of an innate, grammar-specific module [16].

Besides the above theoretical and empirical arguments, computational modeling has recently joined the endeavor to serve as an effective methodology to study language evolution, and many models have been reported covering various areas of language evolution (e.g., [17–23], a brief review of some of which is given in Wagner et al. [19] and Gong and Wang [24]).

All these “emergent” models (according to [19] view language evolution as a Complex Adaptive System (CAS) [25]. They adopt many assumptions supported by linguistics or theories in other subjects; they take objective mechanisms and follow traceable procedures to obtain replicable, convincing results, some of which can be verified by empirical data (e.g., the results of human subject test on perception of the artificial language match the simulation results in [12]; they shed light on the development of natural languages. For example, interactions between agents and learning through generations drive the emergence of language [22], and language-specific syntactic predispositions may also be gradually adapted from other domain’s abilities [21]. In all, computational simulation is gradually accepted as an alternative way to explore problems, for which, theoretical argumentations alone are not sufficiently reliable and incomplete empirical findings may not allow the original developmental history to be established. For example, both the scenarios discussed above only focus on the general process and the main cause, without clear explanations on the detail or the exception and on the influences of other related forces along this process. Meanwhile, most empirical findings or empirical research only cover limited time slots in history or touch on some short-term phenomena.

In this article, we adopt a multiagent computational model that operates the evolution of the lexicon and syntax through interactions among individual agents. It allows us to make inferences about the language evolution at the mesohistory level [26] by simulating the diachronic change of language from one primitive stage to another at the community level. This model traces a lexicon-syntax coevolution, which instantiates the “formulaic” scenario without innate syntax. It also modifies several shortcomings that exist in current computational models, such as direct meaning transference in [20,22,23], which disregard the interaction of linguistic and nonlinguistic information in the comprehension process. Besides, a primary study on the influence of individual differences is given based on this model. The rest of the article discusses the model (§2), summarizes the results (§3) and presents conclusions and some future directions (§4).

2. DESCRIPTION OF THE MODEL

In this model, language is treated as a set of mappings between meanings and utterances (M-U mappings). Com-

munication is a language game, in which, agents, based on linguistic and nonlinguistic information, produce and comprehend utterances that encode with integrated meanings. The emergence of a common language is indicated by sharing a common set of rules among agents through iterative communications. The main components of this model are introduced in this section, including linguistic rules, agents' abilities and communication.

2.1. MEANING, UTTERANCE AND RULE-BASED SYSTEM

In this section, we define the meanings that agents can express and comprehend, the utterances that encode the meanings and the linguistic rules that agents refer to in production and comprehension.

Meanings are either single concepts such as objects or actions, e.g. "dog"/"meat," "run"/"eat," or semantic integrations of single concepts, such as "dog run"/"dog eat meat." There are two types of integrated meaning, "predicate{agent}" (e.g., "run{dog}") and "predicate{agent, patient}" (e.g., "eat{dog, meat}" or "chase{fox, cat}"), where Predicate usually is an action, Agent is an instigator of an action or a sentient causer, and Patient is an entity undergoing an action [27]. Semantically, some integrated meanings are transparent, after whose meaning constituents are identified, it is obvious that what is the agent and what is the patient. For example, if all the meaning constituents of an integrated meaning are identified as "eat," "dog," and "meat," the whole meaning of it, in normal situations, must be inferred as "eat{dog, meat}," instead of "eat{meat, dog}." Other integrated meanings are opaque, even whose meaning constituents are identified, the whole meanings are still not clear. For example, if the meaning constituents of an integrated meaning are identified as "chase," "fox," and "cat," without further information, it is not clear "who is chasing whom." The further information required in comprehension of opaque meanings includes syntactic knowledge (e.g., word order) and nonlinguistic information (e.g., environmental information). In this model, through introducing one type of nonlinguistic information (environmental cues, discussed in Section 2.3), the role of this information in the comprehension is studied, and how the comprehension gradually shifts from relying on nonlinguistic information to relying on gradually formed syntactic knowledge, as in modern languages, is traced.

In this model, the semantic space contains 48 integrated meanings (half transparent and half opaque) that are built upon 12 meaning constituents belonging to 3 semantic categories: "Agent," "Predicate," and "Patient" (some unreasonable combinations of these meaning constituents are omitted in the semantic space). Each integrated meaning describes a realistic environmental event. Agents only produce and comprehend these integrated meanings.

The material for agents to encode these integrated meanings and transfer information among other agents is

the utterance. Utterances consist of a string of syllables, such as /a/, /b/, /c/, etc. Utterances are combinable under the regulation of simple word order to map either meaning constituents or the whole integrated meanings.

Linguistic information is stored as a set of linguistic rules, including both lexical (M-U mappings + strength) and word order rules (sequencing orders + strength). The strength of each rule numerically indicates the probability of successful use of that rule. The self-organization strategies of agents, including rule competition and rule adjustment among available rules (discussed in §2.3.2), which simulate agents' decision making process in communications, are both based upon the rule strength.

Lexical rules include holistic and compositional rules. Holistic rules are mappings between integrated meanings and holistic utterances, e.g.,

$$\text{"run{dog}"} \leftrightarrow /a\ b\ c/(0.4),$$

where the integrated meaning "run{dog}" and the utterance /a b c/ are associated with a strength of 0.4. Compositional rules include both *word* and *phrase rules*. Word rules are mappings between one meaning constituent and an utterance, e.g.,

$$\text{"run{\#}"} \leftrightarrow /d\ e/(0.3) \text{ or } \text{"dog"} \leftrightarrow /c/(0.5),$$

where "#" can be replaced by other compositional rule(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 an utterance, e.g.,

$$\text{"eat{dog, \#}"} \leftrightarrow /c*f/(0.4).$$

where a partial integrated meaning is encoded with an utterance. To form an integrated meaning, it still needs another word rule which can fill its meaning gap "#," and insert an utterance into its "*." These bidirectional mappings in lexical rules guide meaning encoding in production and utterance decoding in comprehension.

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:

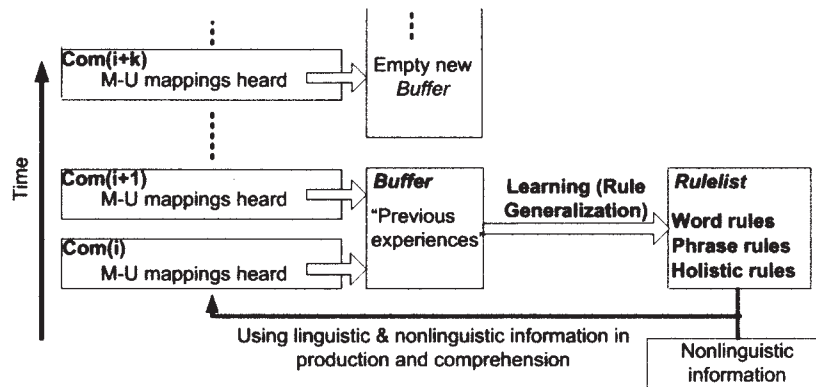
"utterance for predicate before that for agent"

(in short, VS) or "utterance

for predicate after that for agent"(SV).

To express "predicate{agent, patient}" meanings, six orders are considered:

FIGURE 1



Two-level memory system. Com(i) is the index of the *i*th communication.

“utterance for predicate first; that for agent second; that for patient last” (VSO) and VOS, OVS, SVO, SOV, OSV,

where, *S* represents the utterance for agent, *V* for predicate, and *O* for patient, not the usual terms (Subject, Verb, or Object) in linguistic sentence analysis.

The orders for “predicate{agent}” and “predicate{agent, patient}” meanings evolve independently. When many order rules are optional for regulating utterances, agents prefer the one with a higher strength.

This model assumes that the primitive communication system is a holistic signaling system without dominant word order. To simulate this, initially, all agents only share a limited number of holistic rules (6) containing all 12 meaning constituents (we do not simulate how agents acquire meaning constituents, so all constituents appear in the pre-shared holistic rules), and all orders are treated equally (initialized with same strength 0.5). Under the guidance of these semantic constituents and other self-organizing mechanisms during iterative communications, when all agents use common compositional rules and some dominant word order rules to express integrated meanings, a compositional language emerges.

2.2. AGENTS, RULE STORAGE AND RULE ACQUISITION

The agent in this model is an autonomous entity with an internal memory system for storing linguistic rules and some particular abilities which determine his behavior in interactions with other agents.

The agent’s memory system has two layers; a buffer and a rule list (see Figure 1). This structure is inspired by the Learning Classifier System model [28]. The buffer stores “previous experiences,” which are M-U mappings obtained from previous communication(s). The rule list stores “linguistic knowledge” (lexical rules), which are learned from

“previous experiences.” Learning happens when the buffer is full; new lexical rules are generalized among M-U mappings in the buffer and updated into the rule list. Lexical rules in the rule list, together with nonlinguistic information, are used in production and comprehension in future communications, and the buffer is emptied to store new M-U mappings to be obtained in future communications.

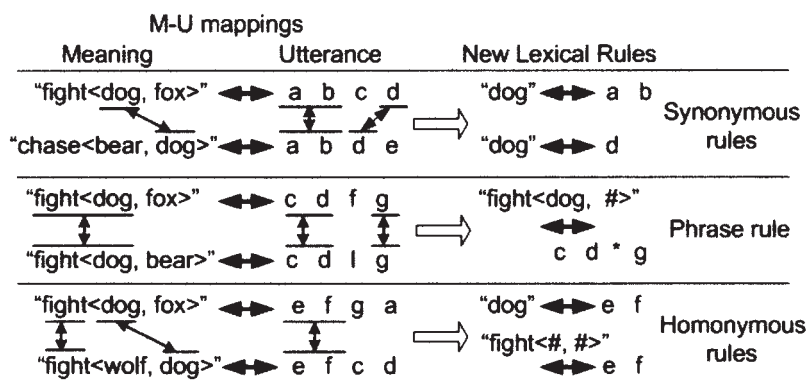
The agent’s particular abilities include storing linguistic information (rules), acquiring rules and communicating with other agents. Agents have two mechanisms to acquire lexical rules.

1. *Random creation* in production. With certain probability, for an integrated meaning without sufficient rules to express, the speaker may randomly select a set of syllables to map either a) the whole integrated meaning, or b) only the inexpressible constituent(s) in the integrated meaning (similar to [22]). And these newly created holistic or compositional rules are updated into the speaker’s rule list.

To express meanings with more constituents usually requires more efforts, so the probability of random creation is inversely proportional to the number of inexpressible constituents in the chosen meaning, i.e., the more inexpressible constituents contained in it, the less the probability for agents to create utterance to map these inexpressible items.

2. *Rule generalization* through detecting recurrent patterns among M-U mappings comprehended in previous communications. Some examples are given in Figure 2. Recurrent patterns are similar features repeatedly appearing in many M-U mappings in the buffer, e.g., identical meaning constituent(s) in meaning parts (e.g., “dog” or “fight{dog, #}”) or identical syllable(s) in utter-

FIGURE 2



Rule generalization examples.

ance parts (e.g., /a b/ or /c d * g/). Agents only focus on identical parts, and do not care about the other meaning constituents (indicated by "#"), the other syllables (indicated by "*") or the locations of them. These identical meaning constituents and identical syllables are mapped as compositional rules with certain initial strength 0.5 (e.g., a word rule "dog" ↔ /a b/ or a phrase rule, "fight{dog, #}" ↔ /c d * g/), and these rules are inserted into the listener's rule list. Detecting recurrent patterns that emerge inevitably triggers the segmentation of integrated meanings into meaning constituents, and holistic utterances into combinations of separated, compositional utterances. This mechanism is the internal mechanism by which compositional language is transmitted horizontally among agents. Other external mechanisms, such as the bottleneck effect during the vertical transmission (communication among agents of different generations), are discussed in [22, 23].

Natural languages contain many synonyms and homonyms. The former are words with the same or similar meaning, but pronounced differently, such as "liberate/release," "farmer/peasant," etc.; the latter are words pronounced the same, but having different meanings, such as "aye/eye," "dear/deer," etc. [27]. In our model, rule generalization results in many synonymous and homonymous rules. For example, the presence of multiple sets of recurrent utterance syllables but one set of recurrent meaning constituents causes the emergence of synonymous rules in an agent's rule list (see Ex. 1 in Figure 2). Similarly, the presence of multiple sets of recurrent meaning constituents but one set of recurrent utterance syllables causes an agent to generalize many homonymous rules (see Ex. 3 in Figure 2). Synonymous rules may increase the speaker's load for searching rules in production and occupy more space in the rule list. Homonymous rules may cause ambiguity in the

listener's comprehension. Linguistic context may avoid such ambiguity, which is widely used in modern languages. However, some internal mechanisms for avoiding ambiguities caused by homonyms have been suggested based on some empirical research. For example, according to the principle of contrast [29], children tend to avoid mapping utterances that are already mapped to an extant meaning to novel, salient meanings, especially those meanings in the same semantic category. For instance, if the child has already learned the word "apple," when given an apple and a banana, and is asked "bring me the banana," usually, he brings the banana. This is because, after the child has mapped "apple" to one type of or one genus of objects with certain features, in order to avoid potential ambiguity in future, he will not map the same word form to other objects or other genus with different features, especially those in the same semantic category, in this case "fruits." Such homonym-avoidance mechanism can be traced in some research (e.g., [15]) on language acquisition by children.

Considering the lack of linguistic context and inevitable emergence of homonymous rules in this model and based on the empirical findings, we adopt a similar homonym-avoidance mechanism in agents. After the normal rule adjustment in a communication, an additional penalty on rule strength is executed on some homonymous rules to the rules chosen by agents to produce and comprehend utterances in that communication. These homonymous rules have the identical utterance syllables to those of the chosen rules, and their meaning parts belong to the same semantic category as those of the chosen ones. The penalty on rule strength could be either a strength increase of those homonymous rules if the chosen rules failed to help the listener to get a confident comprehension or a strength decrease of those homonymous rules if the chosen rules succeed in getting a confident comprehension to the listener. How to judge the confident comprehension is introduced in §2.3.

Here, we give an example of this homonym-avoidance mechanism. Suppose in a communication, one chosen rule of the listener is “dog” \leftrightarrow /a c/ (0.5) and it succeeds to give a confident comprehension to him, then, another rule in the listener’s rule list, “cat” \leftrightarrow /a c/ (0.4), is penalized since it is homonymous to the chosen rule and its meaning part, “cat,” belongs to the same semantic category of the chosen one’s, “dog.” However, another homonymous rule, “run{#}” \leftrightarrow /a c/ (0.7), is not penalized because “run{#}” and “dog” belong to different semantic categories. The necessity of this avoidance mechanism is discussed in §3.2, based on the simulation result.

In this section, the agent’s memory system and the rule acquisition mechanisms are discussed. Sociolinguists have observed dramatic variations in speech communities, including various linguistic abilities [30], and studies on language acquisition have revealed various dichotomies in children’s learning styles [31]. Therefore, these natural characteristics (memory size) and linguistic abilities (rule acquisition mechanisms) should be individually different. However, many computational models which adopt multiagent systems are built with homogeneous agents, all of which have identical characteristics and consistent strategies. These models can not study the influence of these heterogeneities on language emergence. In our model, these heterogeneities are considered, and their influences are discussed in §3.4.

2.3. Communication

Many computational models simplify the communication process by adopting a direct meaning transference, i.e., the intended meanings, encoded in linguistic utterances produced by speakers, are always accurately available to listeners. This assumption is similar to that of mind-reading. However, there is no direct connection between speakers’ production and listeners’ comprehension [32]. Other channels, such as pointing while talking or feedback by nodding, can only provide certain degrees of confirmation. Quine presented a good counterexample regarding pointing while talking [33]: if someone points to a dog and says: “look at the dog!”, how do listeners know that the word “dog” refers to the animal, instead of the grass on which it sits or even the pointing finger itself? Meanwhile, feedback through facial expressions or gestures (e.g., nodding) may not allow speakers to know for sure whether listeners correctly infer speakers’ intended meanings. Furthermore, such simplification disregards the comprehension process, in which, both the agent’s linguistic information and nonlinguistic information provided by the environment (e.g., visual information) can assist the comprehension. In the early stage of language development, based on the *shared attention* [34], a listener can sometimes infer the speaker’s intended meaning when describing an ongoing event. Therefore, the nonlinguistic information can provide sources for comprehension, espe-

cially when linguistic information is inadequate, i.e., the available linguistic rules can not provide an integrated meaning for the listener to comprehend. In order to study how linguistic and nonlinguistic information aid each other and how the linguistic information gradually gains its advantage over the nonlinguistic information, the simplification of the direct meaning transference is abandoned, and a comprehension process with indirect meaning transference is adopted.

2.3.1. Nonlinguistic Information—Cues

In this model, one type of nonlinguistic information, cues, is introduced. Cues describe environmental events and are available to the listener during communications. Cues are modeled as integrated meanings having a certain strength (here, all strengths are equal to 0.5), e.g.,

“chase(fox, dog)”(0.5).

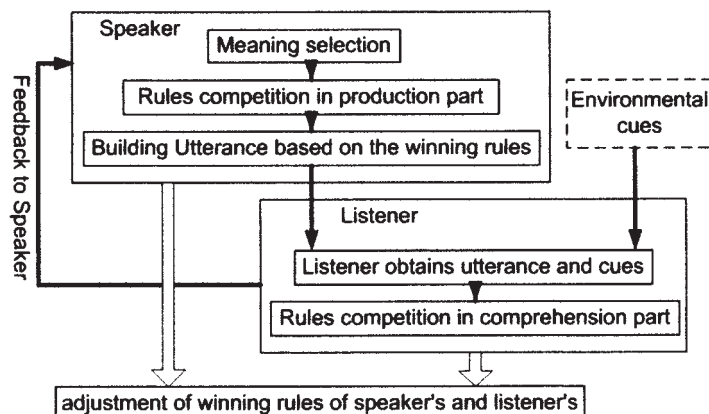
Cues, as semantic hints for comprehension, are sometimes ambiguous, like in Quine’s example. Besides, cues may be unreliable because the speaker might not always be describing an event that is ongoing in the immediate environment. Together, such sources of unreliability may cause the listener to totally ignore certain events or pay attention to the wrong events. However, as stated before, considering the shared attention, sometimes, the listener can infer the speaker’s intended meaning from cues. To simulate this ambiguity, reliability of cues (RC) is used to manipulate the probability that a cue containing the speaker’s intended meaning is available to the listener. During each information exchange in a communication, multiple cues comprising different integrated meanings are simultaneously made available to the listener. The effect of RC on language emergence is discussed in §3.2, based on the simulation result.

2.3.2. Communication Game

In the communication game of this model, there is no direct connection between the speaker’s production and the listener’s comprehension. The listener’s comprehension is based on interactions of linguistic and nonlinguistic information, together with a confidence feedback without direct access to the speaker’s interpreted meaning.

In each communication game, there are many times of information exchange, one such information exchange proceeds as follows (see Figure 3). In the production, the speaker randomly chooses an integrated meaning to express. The speaker decides how to encode the meaning into utterance according to his linguistic rules. If no rules are available and random creation fails, the speaker makes no production at all. Otherwise, a set of extant or newly created lexical rules that can be combined to express the integrated meaning is activated, together with the appropriate word order rules that can combine the activated compositional

FIGURE 3



Information exchange in a communication game through indirect meaning transference.

rules. The speaker selects the set of rules with the greatest combined strength for speaking, CS_{speak} . CS_{speak} is defined by

$$CS_{speak} = Average \left\{ \begin{array}{l} Strength(\text{available lexical rules}) \\ + Strength(\text{applicable word order rules}) \end{array} \right\}. \quad (1)$$

It shows that the speaker's production is decided by his available lexical and syntactical information. The utterance, built up according to the winning rules, is sent to the listener for comprehension. An example of rule competition in the production process is given in the Appendix.

In the comprehension, the listener receives the utterance and, sometimes, some cues. Then, the lexical rules in the listener's rule list whose utterance parts partially or fully match the received utterance are activated. The rule competition here not only takes account of the strength of the listener's available lexical and word order rules, but also considers the cues whose semantic hints support any of the activated lexical rules. For example, the rules available to the listener might only allow the utterance to be interpreted as "eat{dog, #}." But if an available cue contains a meaning "eat{dog, meat}," the strength of this cue will be included in the calculation of the combined strength for listening, CS_{listen} defined by

$$CS_{listen} = Average \left(\begin{array}{l} Average \left\{ \begin{array}{l} Strength(\text{activated rules}) \\ + Strength(\text{applicable word order rules}) \end{array} \right\} \\ + \{Strength(\text{related Cues})\} \end{array} \right). \quad (2)$$

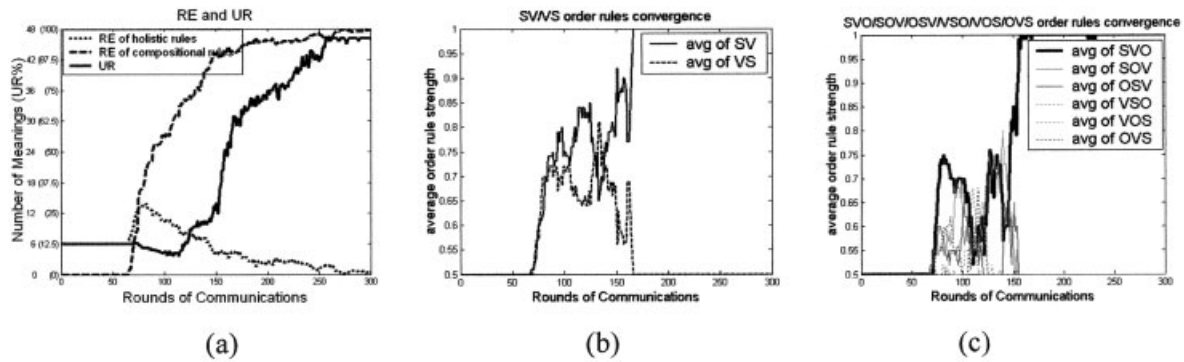
This definition indicates that the listener's comprehension is determined by both linguistic and nonlinguistic information. The listener tends to comprehend meanings that are supported by both linguistic rules and cues, especially when there is a lack of linguistic rules that allow the utterance to be interpreted as an integrated meaning. CS_{listen} is used to compare the optional integrated meanings that the listener can comprehend based on the activation of different rules, supported by cues or not. The listener can make no comprehension if there are no rules for comprehending the heard utterance and no cues available to him.

The listener interprets the utterance as the meaning inferred from the set of rules that has the highest CS_{listen} . If this CS_{listen} exceeds a threshold (say, 0.5), a positive feedback, indicating a strong confidence on the listener's comprehension, is sent back to the speaker. Otherwise, a negative feedback is sent, indicating that the listener is either unable to comprehend the utterance or is not confident of his comprehension. Note that, even a positive feedback is sent, the listener's comprehended meaning is not necessary to be identical to the speaker's intended meaning, the confidence of the listener is decided by the combined strength of the rules he uses to comprehend the heard utterance. Then, based on this confidence feedback, instead of a direct meaning check, both the speaker and the listener perform rule adjustment to their activated rules. Under a positive feedback, they increase the strengths of their winning rules and decrease those of the losing ones; under a negative feedback, an inverse adjustment is executed.

3. RESULTS AND DISCUSSIONS

The simulation results under one set of parameters are discussed here. The population size is 10, the reliability of cues (RC) is 0.7, the average buffer and rule list sizes are 40, and the

FIGURE 4



Lexicon-syntax coevolution (a typical run): (a) Rule expressivity (*RE*) and understanding rate (*UR*); (b) Emergence of dominant word order for “predicate(agent)” meanings; (c) Emergence of dominant word order for “predicate(agent, patient)” meanings. The abscissa is the number of communication rounds, the ordinate is the number of meanings (and understanding rate in brackets) in (a) and the rule strength in (b) and (c).

average probability for random creation and for rule generalization are 0.5. We adopt a concurrent simulation system, i.e., each agent communicates with each other agent during each communication round, in all, $C_{PopulationSize}^2$ number of communications happen simultaneously. In one communication, there are 20 times of information exchange between the speaker and the listener. Results under other sets of reasonable parameters are similar.

In order to analyze the results, 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 the lexicon and syntax. After that, we discuss some factors that may affect this self-organizing process, some of which may shed light on the evolution of natural languages.

3.1. Indices to Test the Behavior

1. Rule expressivity (*RE*): the average number of integrated meanings that each agent can express:

$$RE = \frac{\sum_i \text{number of meanings that agent } i \text{ can express}}{\text{number of agents}} \quad (3)$$

One holistic rule expresses one integrated meaning. One compositional rule can express one or two meaning constituents, but not an entire integrated meaning. However, through combination, a limited number of compositional rules can express many integrated meanings. For example, in this model, 12 word rules are enough to express all 48 integrated meanings in the semantic space. Compositionality endows human languages with the ability to express infinite meanings with limited materials. *RE* is used to trace

the transition from holistic rules to compositional rules in agents’ rule lists.

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

$$UR = \frac{\sum_{i,j} \text{number of understandable meanings between agent } i \text{ and } j}{\text{number of all possible pairs of } i, j} \quad (4)$$

The *UR* can evaluate another feature of real languages, the displacement [35], i.e., whether expressions with these linguistic rules can be accurately comprehended even though they describe events that are not happening in the present space and time or there are no related cues. In models adopting the direct meaning transference (e.g., [20, 23]), each speaker’s intended meaning is always accurately transferred to the listener, so these models may not be suitable for evaluating the displacement.

3. Convergence time (*CT*): the average number of communication rounds required to achieve a mature language ($UR > 80\%$).

3.2. Lexicon-Syntax Coevolution

The lexicon-syntax coevolution is summarized in Figure 4. Figure 4(a) shows the *RE* of both holistic and compositional rules; the decrease of the former and the increase of the latter show a transition from an initially holistic signaling system to a compositional language. The *UR*, shown in Figure 4(a), undergoes an S-shaped curve, which indicates that the size of the common lexicons undergoes a phase transition [36] (a sharp increase of *UR*), similar to the result of Ke et al.’s model [20]. Figure 4(b,c) shows the convergence of the syntax from all possible sequential order rules

to the dominant word orders; the curves trace the average strengths of each of the eight order rules. Two dominant word orders emerge, one for each meaning type. No prior bias is conferred to any particular word order; each is *a priori* equally likely and, initially, with the same strength.

Without considering other restrictions on the dominance of certain orders (e.g., *SOV* or *SVO*) as in natural languages, any order can be the dominant one. This indicates, from a usage point of view, that any order can fulfill the regulating role. Language usage provides an arena for multiple factors to take effect on driving the emergence of syntax, from simple dominant word order as in this model to the complex syntactic features shown in modern languages. The multiple factors include linguistic factors, such as frequency-based construction, phonology change [37], and factors from other aspects, such as individual perspective [38].

The lexicon-syntax coevolution typically proceeds as follows: 1) Holistic period (from 0 to 60 rounds in the run summarized in Figure 4). At the beginning, agents only understand meanings produced by the shared holistic rules. Then, more holistic rules emerge due to occasional random creation, which gradually increases the *RE* of holistic rules. The *UR*, in this period, relies on holistic rules. 2) Random order period (from 60 to 150 rounds). Later on, the presence and acquisition of recurrent patterns that appear by chance greatly increase the *RE* of the compositional rules. A consistent word order becomes necessary when agents come to use more compositional rules to encode meanings, so the convergence of dominant word order begins too. But in this period, the strength of almost every order rule is increased, because agents have no prior bias toward any particular word order. The comprehension gradually relies on compositional rules, though their use causes some meanings, initially understandable when produced by holistic rules, to be misunderstood. This may cause a slight, temporary drop of the *UR*, which matches a similar drop of understandability during the children's language acquisition [15]. 3) Self-organizing period (from 150 to 170 rounds). Self-organizing mechanisms gradually make certain compositional and word order rules win the competition and get shared among agents. In this period, there is a sharp increase of both the *UR* and the strengths of those will-be dominant word order rules. 4) Conventionalization period (from 170 rounds till the end). The sharing of a common lexicon is finalized in this period. In certain simulations, *UR* can approach 100%, indicating that the shared lexicons are sufficient to produce and comprehend almost all integrated meanings in the semantic space.

Mutual understanding requires not only common lexical rules but also a shared syntax to regulate utterances. Unreliable cues drive the transition from relying on nonlinguistic information to relying on linguistic information. This transition is a lexicon-syntax coevolution process: from Figure 4(a–c), it is seen that the sharp increase of *UR* and that of

TABLE 1

Shared Rules among Agents in a Typical Run with Homonym-Avoidance

Round = 492

Common Rules: 11

"cry(#)" ↔ /a/	"run(#)" ↔ /d/
"chew(#)" ↔ /b f/	"fight(#, #)" ↔ /e/
"chase(#, #)" ↔ /c/	"dog" ↔ /b f/
"bear" ↔ /h/	"fox" ↔ /c/
"wolf" ↔ /i/	"water" ↔ /k/
"meat" ↔ /a/	

Average UR = 40.80 (85.00%)

the strengths of the dominant order rules are almost synchronized, i.e., the use of compositional rules triggers the syntax to converge, which in turn boosts the lexicon convergence.

3.3. Factors Affecting the Coevolution Process

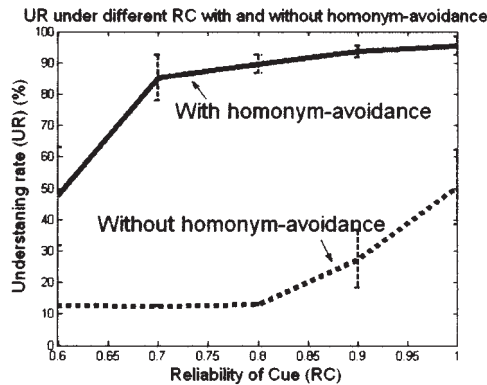
In this section, some internal (homonym-avoidance, heterogeneous individual characteristics, and linguistic abilities) and external (cue reliability) factors that affect the lexicon-syntax coevolution in this model are discussed.

3.3.1. Homonym-avoidance and Cue Reliability (RC)

The empirical basis and the execution procedure of the homonym-avoidance mechanism are given in §2.2. In this section, the necessity of this mechanism is strengthened by the simulation result. Statistical results show that without homonym-avoidance, the peak *UR* is very low (12/48, 25%). Table 1 lists the shared rules in a typical run with homonym-avoidance. Although some homonymous rules with meaning constituents in different categories (in bold) exist, the high *UR* indicates less ambiguity.

From an external perspective, the higher the reliability of cues, the higher chances for listeners to get assistance from cues in their comprehensions, which may accelerate the conventionalization of common lexicons. But the performance of this external factor relies much on the internal homonym-avoidance mechanism. Figure 5 shows the average understanding rate for different *RC* values, with and without homonym-avoidance. In both cases, the understanding rate increases with the *RC*, and the understanding rate with homonym-avoidance is much higher than that without homonym-avoidance. However, even when cues are always reliable (i.e., *RC* = 1.0), without homonym-avoidance, the peak *UR* (50%) is much lower than the peak *UR*

FIGURE 5



Average *UR* (understanding rate) under different *RC* (reliability of cue) with and without homonym-avoidance. The abscissa is the degree of *RC* and the ordinate is the degree of *UR*. Results here are average *UR* in 10 simulations under the same conditions.

(94*) with homonym-avoidance. This is because, although accurate comprehension is possible with the aid of highly reliable nonlinguistic information, homonymous rules could still be shared among agents and cause ambiguity when nonlinguistic information is absent. Besides, with homonym-avoidance, there seems to be a threshold of *RC* for understanding rate to reach a high level (0.7 in Figure 5); *RC* exceeding that threshold may not greatly increase the *UR*, whereas, *RC* below that threshold, is not enough to allow a high *UR*. This indicates that certain reliable external assistance or feedback is necessary.

Both the empirical findings (e.g. [39]) and the simulation results suggest that homonym-avoidance, one internal constraints for avoiding ambiguity, and reliable nonlinguistic information, one external requirements, are necessary for a mature language ($UR > 80\%$) to emerge.

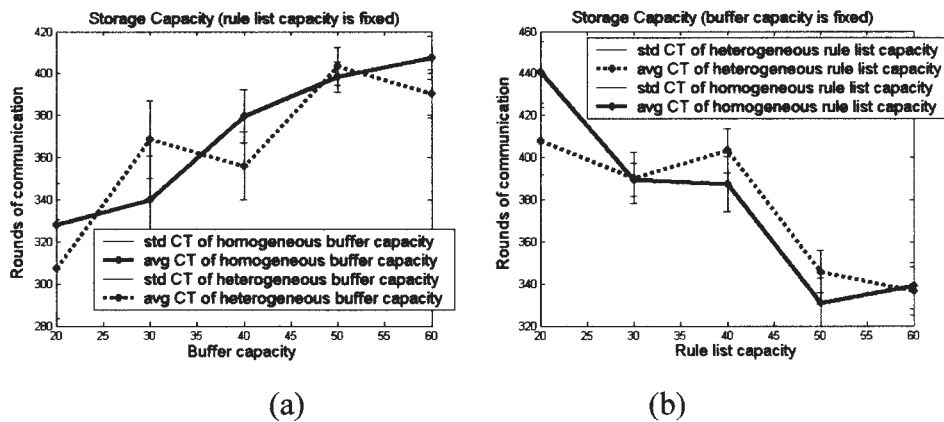
3.3.2. Heterogeneous Population

This section describes the influence of two types of agent heterogeneity on the emergence of language. One way to simulate the heterogeneity is to use a Gaussian distribution, i.e., each agent has a value of that heterogeneous feature, randomly drawn from a Gaussian distribution. The influence of the heterogeneity is shown through comparing the convergence time (*CT*) in a heterogeneous situation with that in a homogeneous situation. In the homogeneous situation, all agents share the same value of that feature which is equal to the mean value of the Gaussian distribution in the heterogeneous situation.

1. Heterogeneous storage capacities, i.e., different agents can have various sizes of buffer and rule list. Under a fixed rule list capacity (40), Figure 6(a) shows the *CT*s under a heterogeneous buffer capacity (dashed line); under a fixed buffer capacity (40), Figure 6(b) shows the *CT*s under a heterogeneous rule list capacity (dashed line). In both heterogeneous situations, the mean values of the Gaussian distribution are from 20 to 60, the variances are 10, and the corresponding *CT*s under homogeneous situations also are shown in Figure 6 (solid lines).

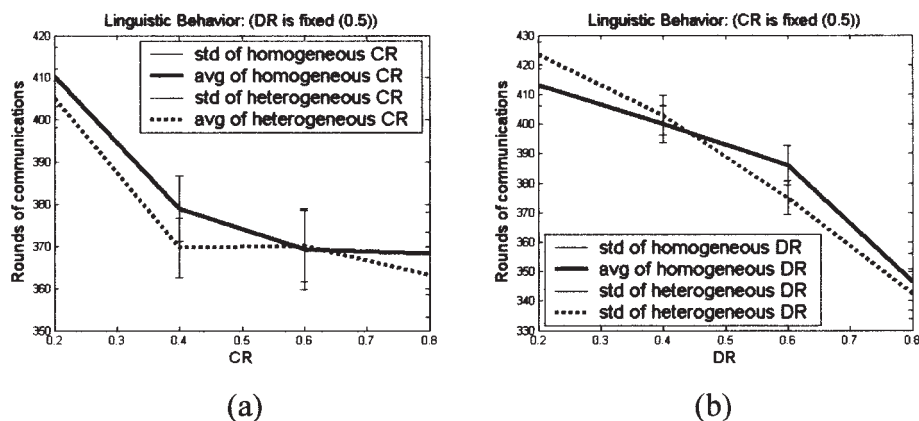
The buffer capacity affects the rule generalization. With abundant M-U mappings stored in the buffer, chances for recurrent patterns among them are high, and many new rules may be generalized simultaneously. However, a bigger buffer needs more communications to fill, which may delay

FIGURE 6



Heterogeneity: convergence time (*CT*) as a function of different memory capacity: (a) *CT* of heterogeneous buffer capacity under a fixed rule list capacity, the abscissa is the mean value of buffer capacity; (b) *CT* of heterogeneous rule list capacity under a fixed buffer capacity, the abscissa is the mean value of rule list capacity. The ordinate in (a) and (b) is the number of communication rounds.

FIGURE 7



Heterogeneity: convergence time as a function of different linguistic abilities: (a) *CT* under fixed detection rate and different creation rate, the abscissa is the mean value of creation rate; (b) *CT* under different detection rate and fixed creation rate, the abscissa is the mean value of detection rate. The ordinate in (a) and (b) is the number of communication rounds.

the rule updating rate. The *CT* curve in Figure 6(a) shows that with the increase of the buffer capacity, the *CT* increases slowly, which indicates that the delayed updating rate in bigger buffer capacities holds back the convergence time slightly. For most buffer capacities within the magnitude of the semantics space, most simulations converge to languages with high understanding rates, even in heterogeneous conditions.

The rule list capacity determines the rule storage. The bigger the capacity, the more easily agents store new rules, which can accelerate the convergence process. However, for a fixed semantic space, a large rule list capacity introduces redundancy. For example, in the current model, 12 word rules are sufficient for all integrated meanings in the semantics space. When the system converges, there are empty slots or slots storing non-shared rules in some agents' rule list. As shown in Figure 6(b), there is a slow decrease of the *CT* with the increase of the rule list in both homogeneous and heterogeneous situations. This shows that for most rule list capacities that are not lower than 12, most simulations converge to languages with high understanding rates, even in heterogeneous conditions.

2. Heterogeneous linguistic abilities, i.e., different abilities of random creation [creation rate (*CR*)], which controls the rate for creating salient linguistic mappings, and different abilities to detect recurrent patterns [detection rate (*DR*)], which controls the rate for acquiring new rules from available M-U mappings. As for the modeling of heterogeneous storage capacities, Gaussian distribution is used to simulate this heterogeneity. Under a fixed *DR* (0.5), Figure 7(a) shows the *CT*s under a heterogeneous creation rate; under a fixed *CR* (0.5), Figure 7(b) shows the *CT*s under a heterogeneous detection rate. In both heterogeneous situ-

ations, the mean values of the Gaussian distribution are from 0.2 to 0.8, the variance is 0.2, and the corresponding *CT*s under homogeneous situations also are shown in Figure 7 (solid lines).

It is obvious that without an intentional creation mechanism, no salient linguistic materials will emerge. Similarly, without detection of recurrent patterns, though many linguistic rules may emerge, similarities among them will only occur by chance and the emergence of a common set of lexicon will be virtually impossible. Furthermore, in cases with extremely low creation rate or detection rate, the acquisition of new rules can still be delayed. If the creation rate is low, although agents can detect recurrent patterns, acquisition of new rules is delayed because of the insufficient amount of salient linguistic materials. If the detection rate is low, although many recurrent patterns appear in newly created linguistic materials, few of them can be detected and learned as linguistic rules, acquisition of new rules is delayed, too. As shown in Figure 7, except for these extreme cases, most simulations converge to languages with high *UR*s, even in heterogeneous conditions.

To summarize, the limited heterogeneities on storage capacities and linguistic abilities can not significantly affect the emergence of language. The self-organization process in this model is robust to a certain degree, i.e., a mature language can still emerge in a population, each agent of which may have small differences among each other in his natural characteristics or linguistic abilities.

4. CONCLUSIONS AND FUTURE DIRECTIONS

In this article, the coevolution of the lexicon and syntax is traced in a computational model. Nonlinguistic information's influence and some individual differences' effects on

the emergence of language are discussed by adopting a communication system with indirect meaning transference among a heterogeneous population. This model also discusses some external and internal requirements for effectively acquiring a compositional language. However, there are several unexplored aspects of language emergence based on this model.

First, this model shows the emergence of language through horizontal transmission, which can supplement the conclusions of models considering the vertical transmission. In order to test whether the self-organizing mechanisms in this model are also applicable for vertical transmission, there is necessity to adopt the vertical transmission in this model. Second, human's semantic knowledge is gradually increasing, what is the influence of this increasing semantic space on language evolution can be explored by adopting an open-ended semantic space in this model. Third, other types of communication, such as more than two individuals communicate in a group with one as the speaker, the others as the listeners, can be taken into account. Finally, random communication adopted in this model, as in many other models, is unrealistic, because it disregards the effect of social factors, which may also influence language evolution, as discussed in [40]. The recently developing complex networks theory (e.g., [41, 42] provides an effective tool to explore language emergence at a social com-

munity level. Based on the current model, our group has already obtained some preliminary results which are introduced in [43].

ACKNOWLEDGMENTS

The authors thank Prof. G. R. Chen and T. Lee for helpful discussions and suggestions. This work is supported by City University of Hong Kong Grants 9010001, 9040781.

APPENDIX: AN EXAMPLE OF RULE COMPETITION IN THE PRODUCTION PROCESS

Suppose a speaker wants to express the integrated meaning "fight{dog, fox}." Based on his rule list, there are three ways to express it, as shown in Figure A1: 1) using a holistic rule, no word order is considered; 2) using 3 word rules, all 6 orders are applicable, so the strongest one [VSO(0.6)] is chosen; 3) using a word rule and a phrase rule, due to the restriction of the phrase rule's utterance, only VSO and OSV are applicable orders, so the strongest one VSO(0.6) is chosen. In each case, CS_{speak} is calculated, and CS_3 (0.7) is the highest. Therefore, rules in that case (boldface in Figure A1) are his winning rules and the utterance, built up accordingly (/e b f/), is sent to the listener. The comprehension is a reversed process and the calculation of CS_{listen} considers not only the activated linguistic rules but also the related cues.

FIGURE A1

Production Part	Meaning to express: "fight<dog, fox>"		Combined Strength (CS)	Utterance created
	Activated rules	Applicable word order rules		
1 holistic rule	"fight<dog, fox>" \leftrightarrow /a b/ (0.6)		$CS_1 = 0.6$	/a b/
3 word rules	"dog" \leftrightarrow /b/ (0.8) "fight<#, #>" \leftrightarrow /c e/ (0.5) "fox" \leftrightarrow /g/ (0.2)	VSO (0.6)	$CS_2 = 1/2(1/3(0.8+0.5+0.2)+0.6) = 0.55$	/c e b g/
1 word rule 1 phrase rule	"dog" \leftrightarrow /b/ (0.8) "fight<#, fox>" \leftrightarrow /e * f/ (0.8)	VSO (0.6) OSV (0.5)	$CS_3 = 1/2(1/2(0.8+0.8)+0.6) = 0.7$	/e b f/
Utterance built up: /e b f/				

Example of rule competition in the production process

REFERENCES

1. Wang, W.S-Y., Ed. *The Emergence of Language: Development and Evolution*. W. H. Freeman: New York, 1991.
2. Holland, J.H. *Emergence*; Perseus Publication: Cambridge, MA, 1998.
3. Chomsky, N. *The Minimalism Program*; MIT Press: Cambridge, MA, 1995.
4. Jackendoff, R. *The Architecture of Language Faculty*. MIT Press: Cambridge, MA, 1997.
5. Hauser, M.D.; Chomsky, N.; Fitch, W.T. *The Faculty of Language: What Is It, Who Has It, and How did It Evolve?* *Science* 2002, 298, 1569–1579.
6. Knight, C.; Hurford, J.R.; Studdert-Kennedy, M., Eds. *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Forms*; Cambridge University Press: Cambridge, MA, 2000.
7. Oller, D.K.; Griebel, G., Eds. *Evolution of Communication System: A Comparative Approach*; MIT Press: Cambridge, MA, 2004.
8. Calvin, W.H.; Bickerton, D. *Lingua Ex Machine: Reconciling Darwin and Chomsky with the Human Brain*; MIT Press: Cambridge, MA, 2000.
9. Wray, A. *Formulaic Language and the Lexicon*; Cambridge University Press: New York, NY, 2002.
10. Krifka, M. *Compositionality*. In: *The MIT Encyclopedia of the Cognitive Sciences*; Wilson, R.A., Keil, F., Eds.; MIT Press: Cambridge, MA, 2001.
11. Hauser, M.D. *The Evolution of Communication*; MIT Press: Cambridge, MA, 1996.
12. Christiansen, M.H. Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In: *The Proceedings of Evolution of Language: the 3rd International Conference*; Dessalles, J.L., Ghadakpour, L., Eds.; Paris, France, 2000; pp 45–48.
13. Wray, A. *Dual Processing in Protolanguage: Performance without Competence*. In: *The Transition to Language*; Wray, A. Ed.; Oxford University Press: Oxford, NY, 2002; pp 113–137.
14. Wray, A. *Protolanguage as a holistic system for social interaction*. *Language Commun* 1998, 18, 47–67.
15. Fillmore, L.W. *Individual differences in second language acquisition*. In: *Individual Differences in Language Ability and Language Behavior*; Fillmore, C.J., Kempler, D., Wang, W.S-Y, Eds.; Academic Press: New York, NY, 1979; pp 203–228.
16. Schoenemann, P.T. *Syntax as an emergent characteristic of the evolution of semantic complexity*. *Minds Machines* 1999, 9, 309–346.
17. Christiansen, M.; Kirby, S., Eds. *Language Evolution*; Oxford University Press: New York, 2003.
18. Cangelosi, A.; Parisi, D., Eds. *Simulating the Evolution of Language*; Springer Verlag: London, 2001.
19. Wagner, K.; Reggia, J.A.; Uriagereka, J.; Wilkinson, G.S. *Progress in the simulation of emergent communication and language*. *Adaptive Behavior* 2003, 11, 1, 37–69.
20. Ke, J.; Minett, J.W.; Au, C. P.; Wang, W.S-Y. *Self-organization and selection in the emergence of vocabulary*. *Complexity* 2002, 7, 3, 41–54.
21. Munroe, S.; Cangelosi, A. *Learning and the evolution of language: The role of culture variation and learning cost in the Baldwin effect*. *Artif Life* 2002, 8, 4, 311–340.
22. Kirby, S. *Learning, bottlenecks and the evolution of recursive syntax*. In: *Linguistic Evolution through Language Acquisition: Formal and Computational Models*; Briscoe, T. Ed.; Cambridge University Press: Cambridge, 2002; pp 173–205.
23. Smith, K.; Kirby, S.; Brighton, S. *Iterative learning: A framework for the emergence of language*. *Artif Life* 2003, 9, 371–386.
24. Gong, T.; Wang, W. S-Y. *Computational modeling on language emergence: A Coevolution model of lexicon, syntax and social structure*. *Language and Linguistics* 2005, 6(1), 1–41.
25. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology: Control and Artificial Intelligence*; MIT Press: Cambridge, MA, 1992.
26. Wang W.S-Y. *The three scales of diachrony*. In: *Explorations in Language*; Wang, W.S-Y. Ed.; Pyramid Press: Taiwan, 1991; pp 60–71.
27. Fromkin, V.; Rodman, R.; Hyams, N. *An Introduction to Language*, 7th Ed.; Thomson, Heinle: Boston, MA, 2003.
28. Holland, J.H. *Escaping brittleness: The possibilities of general-propose learning algorithms applied to parallel rule-based systems*. In: *Machine Learning, an Artificial Intelligence Approach*, Vol. 2; Michalski, R.S., Carbonell, J.G., Mitchell, T.M., Eds.; Morgan Kaufmann: Los Altos, CA, 1986; pp 593–623.
29. Clark, E. *The principle of contrast: A constraint on language acquisition*. In: *Mechanisms of Language Acquisition*; MacWhinney, B. Ed., Lawrence Erlbaum: Hillsdale, NJ, 1987; pp 1–33.
30. Romaine, S. *Language in Society, an Introduction to Sociolinguistics*; Oxford University Press: Oxford, 1994.
31. Shore, C.M. *Individual Differences in Language Development*; Sage Pub: Thousand Oaks, CA, 1995.
32. Kirby, S. *Natural language from artificial life*. *Artif Life* 2002, 8(2), 185–215.
33. Quine, W.V. *Word and Object*; MIT Press: Cambridge, MA, 1960.
34. Tomasello, M. *Constructing a Language: A Usage-based Theory of Language Acquisition*; Harvard University Press: Cambridge, MA, 2003.
35. Hockett, C.F. *The origin of speech*. *Sci Am* 1960, 203, 88–96.
36. Monasson, R.; Zecchina, R.; Kirkpatrick, S.; Selman, B.; Troyansky, L. *Determining computational complexity from characteristics “phrase transitions.”* *Nature* 1999, 400, 8, 133–137.
37. Bybee, J. *Mechanisms of change in grammaticization: The role of frequency*. In: *The Handbook of Historical Linguistics*; Joseph, B.D., Janda, J., Eds.; Blackwell: Oxford, 2003; pp 602–623.
38. MacWhinney, B. *The emergence of grammar from perspective*. In: *Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics*; Minett, J.W., Wang, W. S-Y., Eds.; City University of Hong Kong Press: Hong Kong, 2005; 129–186.
39. Clark, E.V. *First Language Acquisition*; Cambridge University Press: Cambridge, UK, 2003.
40. Nettle, D. *Linguistic Diversity*; Oxford University Press: Oxford, 1999.
41. Barabási, A. *Linked: The New Science of Networks*; Perseus Publication: Cambridge, MA, 2002.
42. Newman, M.E.J. *The structure and function of complex networks*. *SIAM Rev* 2003, 45, 167–256.
43. Gong, T.; Ke, J.; Minett, J.W.; Wang, W.S-Y. *A Computational Framework to Simulate the Coevolution of Language and Social Structure*. In: *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*; Pollack, J., Bedau, M., Husbands, P., Ikegami, T., Watson, R.A., Eds.; MIT Press: Cambridge, MA, 2004; pp 158–163.