Title: A Computational Model of the Coevolution of Lexicon and Syntax

Correspondent author: Tao Gong

Address: Department of Electronic Engineering,

City University of Hong Kong,

Kowloon, Hong Kong

Telephone: (852)90398466

Fax: (852)27887791

Email: 50005488@student.cityu.edu.hk

Other authors:

Jinyun Ke, James W. Minett, William S-Y. Wang

Affiliation: Language Engineering Laboratory, Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong

John H. Holland

Affiliation: Department of Computer Sciences and Engineering, University of Michigan, USA

No. of pages main text: 19.

No. of figures: 8.

No. of tables: 2.

A Computational Model of the Coevolution of Lexicon and Syntax

Tao Gong*, Jinyun Ke, James W. Minett, John H. Holland, William S-Y. Wang

Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong

*50005488@student.cityu.edu.hk

Abstract

In this paper, a multi-agent computational model is used to simulate the emergence of a compositional language from a holistic signaling system through iterative interactions among heterogeneous agents. Syntax, in the form of simple word order, coevolves with the emergence of the lexicon through self-organization in individuals. We simulate an indirect meaning transference, in which the listener's comprehension is based on the interaction of linguistic and nonlinguistic information, together with a feedback without direct meaning checking. Homonyms and synonyms emerge inevitably during the rule acquisition. Homonym avoidance is assumed to be a necessary mechanism for developing an effective communication system.

Key words: Language emergence, multi-agent model, coevolution, indirect meaning transference, heterogeneous learning

1 Introduction

Recently, there has been a resurgence of interest in a multidisciplinary approach to language emergence [1]. Along with the theoretical argumentation, computational modeling has joined the endeavor to serve as an effective methodology. The computational approach has grown rapidly, as exemplified by several reviews (e.g., [2][3][4]). Different types of models, based on various evolutionary or artificial life theories, have been reported and various areas of language evolution have been discussed. For example, Ke et al. [5] report several simulation models to show that a coherent vocabulary can be reached through self-organization in a population. Munroe & Cangelosi [6] implement a neural network model to demonstrate how learning and natural selection interact under different conditions. A simple compositional structure to represent simple meanings, such as "object action", emerges in their model. Kirby [7][8] presents an iterated learning model (ILM) to simulate the emergence of compositional language with more complex syntactic structures, such as recursion, from a holistic signaling system through iterated learning by successive generations of learners.

All of these "emergent" models (according to [4]) view language evolution as a Complex Adaptive System (*CAS*) [9] and share several assumptions which shed light on the real language development. For example, interactions between agents and learning through generations drive the emergence of language; language-specific syntactic predispositions are unlikely. However, there are still several limitations in these models.

First, most of these models (excluding [6]) assume *direct meaning transference* in interactions among agents; i.e., the intended meanings, encoded in the linguistic utterances produced by speakers, are always accurately available to listeners. This approach is based on the assumption that accurate meaning transference through other channels is possible, especially in supervised learning. However, if this were true, language as a communication medium would have been unnecessary since the intended meaning would always be available without linguistic communication. Moreover, it is obvious that there is at least no direct connection between speakers' production and listeners' comprehension — speakers always use utterances that they believe represent the intended meanings and listeners always interpret utterances into the meanings that they believe such utterances express [10]. Other channels, such as pointing while talking or feedback, can only provide certain degrees of confirmation. Quine's question [11], regarding pointing, is a good counterexample: 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 countenances or gestures may not allow speakers to know for sure whether listeners have correctly inferred the speaker's intended meaning. Therefore, always assuming direct meaning transference between speakers and listeners in communication is unrealistic. Furthermore, comprehension is not based only on linguistic information: nonlinguistic information provided by environment should also be considered.

Second, these models either fail to model syntax (e.g., [5]), build in the syntactic features (e.g., [6]), or else do not adopt a coevolutionary view of the emergence of syntax and lexicon (e.g., [8]). From an evolutionary point of view, the emergence of lexicon and the convergence of syntax should be interwoven, i.e., they should coevolve.

Third, many models are built using homogeneous agents, who have identical characteristics and consistent strategies. However, sociolinguists have observed dramatic variations in speech communities [12], and studies on language acquisition have revealed various dichotomies in children's learning styles [13]. Therefore, it is more realistic for computational models to take account of heterogeneity.

Addressing these limitations, we present a computational model of language emergence at the *macrohistory* level [14] to show a coevolution of lexicon and syntax from a holistic signaling system. Regarding the origins of syntax, there are two notable scenarios: 1) a "bootstrapping" scenario [15], which conceives that a full language, originating from words, developed from the combination of words regulated by an innate syntax; 2) an "emergent" theory [16] which suggests that language may have started from a holistic, "formulaic", signaling system. In the latter theory, sporadic recurrent components in utterances and common meaning aspects are assumed to have triggered the emergence of words through segmentation, and then to have lead to the convergence of shared syntactic structures.

Considering the following arguments, the "emergent" scenario appears to be more attractive and plausible. First, protolanguage may have consisted of a number of holistic signals, similar to those found in primates and other animals such as bees and birds [18], though their nature may be very different. According to the "emergent" theory, there existed 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, from evolutionary principles, grammatical rules in language are more likely to have emerged as the result of conventionalization features due to language use, rather than as the result of an innate, grammar-specific module [17]. Syntax is assumed to have emerged from a pre-adapted cognitive capacity reflected in other cognitive processes (e.g., sequencing ability [19]). Such a sequencing ability as a cognitive predisposition has been attested in other primates as well as in pre-language infants. In the “emergent” theory, grammar is acquired through segmentation, incorporating systematic regularities involving the meaning structures and the sequences used to combine words [16] — word combination therefore requires the sequencing ability, which provides opportunities to conventionalize certain sequences such as simple, dominant word orders. Finally, an analytic language can emerge through iterative interactions among agents without any external guidance. Such a developmental process has been attested in both first [20] and second [21] language acquisition in children.

Based on the “emergent” scenario, we present here a model that shows a development from a non-syntactic, holistic proto-language to a syntactic, compositional proto-language. At first, the initial holistic signals have no hierarchical structure. Then, basic syntax, in the form of simple word order, is introduced to regulate compositional utterances after they are segmented from holistic signals. These word orders, freely chosen by individuals, converge to a dominant word order through iterative communications. This process occurs simultaneous to the emergence of the lexicon.

We model communication using a process of indirect meaning transference, in which the interaction of both linguistic and nonlinguistic information determines comprehension. This indirect meaning transference, involving feedback without direct meaning check, simulates a more realistic process for handling the multiple channels of information that are available in communications.

Finally, the model introduces heterogeneity into the natural characteristics and linguistic behaviors of agents, allowing us to study the influence of a population having heterogeneous cognitive capabilities on language emergence.

The rest of the paper is organized as follows: In Section 2, we provide a concise description of the model. Section 3 discusses some major results. Section 4 proposes several promising future directions.

2 Description of the model

The model for language emergence that we adopt is basically a linguistic communication game. Agents produce and comprehend language via communication. We treat language as a set of mappings between meanings and utterances; these mappings are acquired in communication, not set up beforehand, and are modified during communication, as in the model adopted by Kirby [8]. The adjustment of mappings is only executed among available mappings instead of all possible mappings, an approach that differs from Kirby's [8].

2.1 Meaning, utterance and rule-based system

Meanings are single constituents representing discrete features (e.g. "agent/patient", "predicate<#, #>") or integrations of constituents representing integrated features (e.g. "predicate<agent>") in a semantic space. We consider two types of integrated meaning, "predicate<agent>" (e.g. "run<dog>") and "predicate<agent, patient>" (e.g. "eat<dog, meat>" or "chase<fox, cat>"). Some integrated meanings are *transparent*, such as "run<dog>" or "eat<dog, meat>", since their meanings are inferable from their constituents; others are *opaque*, such as "chase<fox, cat>", whose meanings are not inferable from their constituents. Without further information, such as nonlinguistic information or regulation method in utterances, such meanings could be misinterpreted. The transparency and opacity of meanings shows that accurate interpretation requires both linguistic and nonlinguistic information. In this model, 12 constituent and 48 integrated meanings make up the semantic space. (The integrated meanings are built upon the 12 constituent meanings; half are transparent, half are opaque) Agents only produce and comprehend integrated meanings.

Utterances consist of a string of syllables, and can be combined under the regulation of simple syntax to map to either constituent or integrated meanings.

Language is represented by a set of rules, and comprises both *lexical rules* (mappings between meaning and utterance) and *word order rules* (regulating utterances). Each rule has a *strength*, which indicates numerically the frequency of successful use of the rule. The self-organizations strategies of the agents include *rule competition* (decision-making during both production and comprehension) and *rule adjustment* (adjustment among available rules), both based on rule strength. When all agents share a set of common lexical rules with high strengths, a common language has emerged.

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

“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, encoding both production and comprehension. Word rules are mappings between a single constituent and an utterance. For example:

“eat<#, #>” ↔ /d e/ (0.3)   or   “dog” ↔ /c/ (0.5), where “#” can be replaced by constituents to form an integrated meaning. Compositional rules include both *phrase rules* and *word rules*. Phrase rules are mappings between two constituents (that do not form an integrated meaning) and an utterance. For example:

“eat<dog, #>” ↔ /c * f/ (0.4),

where “#” represents an arbitrary meaning constituent (in this case, an agent) and “*” represents an arbitrary syllable(s) of a word rule, so forming an integrated meaning when combined with this phrase rule.

Word order rules cover all possible sequences of constituents to regulate utterances in expressing integrated meanings. For example, to express “predicate<agent>” meanings, two orders are considered:

“utterance for predicate precedes that for agent” (SV) (0.4),

“utterance for predicate follows that for agent” (VS) (0.3)

To express “predicate<agent, patient>” meanings, 6 orders are considered. For example:

“utterance for agent first; that for predicate second; that for patient last” (SVO)(0.2)

Word orders are randomly chosen at the beginning. The word orders for “predicate<agent>” and “predicate<agent, patient>” meanings evolve independently. A dominant word order emerges when all agents share a high strength for a particular word order rule.

2.2 Agent

Each agent has his own set of rules. Lexical rules are stored in a two-layer memory system (see Figure 1), inspired by *Learning Classifier System* (*LCS*) [22]. The *buffer* stores an array of “previous experiences” — each element of the buffer consists of a mapping between an utterance and an inferred meaning (M-U mapping) from some previous communication. The *rule list* stores the lexical rules that the agent uses for both production and comprehension — lexical rules are learned from previous experiences: when the buffer is full, new lexical rules are generalized from the M-U mappings, and are updated into the rule list. The buffer is then emptied to allow new M-U mappings to be recorded.

[Figure 1 is about here.]

Agents use two mechanisms to acquire lexical rules. a) *Random creation* in production. When the speaker attempts to convey an integrated meaning for which he has no set of lexical and word order rules with which to encode it, he may select a random set of syllables to map either 1) the whole integrated meaning, thus creating a holistic rule, or 2) only those inexpressible constituent(s) in the chosen meaning, thus creating a compositional rule(s) or phrase rule (similar to Kirby's model [7]). The probability of so doing is restricted by the number of inexpressible constituents in the chosen meaning — the greater the number of inexpressible constituents, the smaller the probability that the speaker creates new rules to encode them.

b) *Rule generalization* through detecting *recurrent patterns*. Recurrent patterns are identical meaning constituent(s) in the meaning parts and identical syllable(s) in the utterance parts contained in two or more M-U mappings. It is assumed that agents can detect recurrent patterns. Recurrence of some patterns triggers the segmentation of integrated meanings into meaning constituents and holistic utterances into substrings. If a holistic utterance can be fully segmented into compositional utterances, we say, it is fully decomposed. Figure 2 shows some examples of rule generalization. Regardless of its location in M-U mappings, once a recurrent pattern is detected, a new compositional rule, mapping the identical meaning constituent(s) to the identical syllable(s), is created. The residue part, after removing the recurrent patterns, sometimes, can also be mapped as new compositional rules. Agents can create new rules based on recurrent patterns in the M-U mappings in the buffer alone or both by recurrent patterns and residue parts into the rule.

[Figure 2 is about here.]

Synonym and homonymous rules (see Figure 2) emerge inevitably during the rule acquisition. For example, multiple recurrent patterns in the utterance parts but only one recurrent pattern in the meaning parts can cause many synonymous rules to be generalized in an agent's rule list (see Example 1 in Figure 2). In addition, neglecting extant rules, flexible detection of recurrent patterns can map some extant rule's utterance to salient constituent(s) (see Example 3 in Figure 2), thus introducing homonymous rules in an agent's rule list.

When synonyms arise, agents randomly learn one form in a set of synonymous rules, according to the principle of contrast [23]. Other forms can still be learned in future communications, but through iterative communication, only one form will win the competition and be preferred by agents. For homonyms, a detailed discussion will be given in a later section.

2.3 Communication

This model simulates communication through indirect meaning transference. Self-organizing strategies are implemented in production and comprehension. Comprehension models the interaction of both linguistic and nonlinguistic information.

[Figure 3 is about here.]

Communication (see Figure 3), using self-organizing strategies through indirect meaning transference, proceeds as follows. In production, the speaker randomly chooses an integrated meaning to express. If he has no lexical rules with which to fully encode it and random rule creation has failed, the speaker makes no production. Otherwise, a set of extant or newly created lexical rules that can express the meaning are activated together with an appropriate word order rule , used to combine compositional rules. The speaker decides which set of rules should be used to encode the chosen meaning using rule competition, selecting the set having the greatest *combined strength for speaking*, CS_{speak} , defined by

$$CS_{speak} = \begin{array}{c} \mathrm{Str}(\text{available rules}) \\ +\mathrm{Str}(\text{applicable word order rules}) \end{array} \qquad (1)$$

An example of rule competition is stated in Appendix A. The utterance, built up accordingly, is transferred to the listener, who then attempts to comprehend the utterance.

In comprehension, the listener receives the utterance and, sometimes, some cues, which represent the nonlinguistic information available to the listener during communication. *Cues* are modeled as integrated meanings having some strength (here, all set to the same value, 0.5), for example

“chase<fox, dog>” (0.5).

[1] In cue selection, if a randomized number is less than *RC*, the intended meaning is chosen as the cue, otherwise, a randomly chosen meaning from the semantic space is chosen as the cue. Identical cues are not allowed.

Cues provide semantic hints for comprehension. Based on the principle of shared attention, agents are often able to detect salient cues from the environment. But cues are not always reliable; that is, the integrated meaning that a cue encodes may not match the meaning intended by the speaker. To model this, we manipulate the *reliability of cues* (*RC*, the probability one of the available cues matches the speaker's intended meaning)[2].

The rules in the listener's rule list whose utterance parts partially or fully match the received utterance are activated. The rule competition process that we execute for listening is more complex than that for speaking — we consider not only the strength of the listener's available lexical and word order rules, but also the cues whose semantic hints support some of those lexical rules: the *combined strength for listening* is defined by

$$CS_{listen} = \begin{array}{l} LangWeight \begin{Bmatrix} \mathrm{Str}\left(\text{activated rules}\right) \\ + \ \mathrm{Str}\left(\text{applicable word order rules}\right) \end{Bmatrix} \\ \\ + \ EnvWeight \ \left\{\mathrm{Str}\left(\text{related Cues}\right)\right\} \end{array}, \tag{2}$$

where *LangWeight* and *EnvWeight* are the relative weights of the linguistic and nonlinguistic information. In this model, they are treated as equally important (both are set to 1.0). Note that CS_{listen} has the value zero when the listener has no set of rules with which to decode the received utterance and no cues are available. Rule competition in listening is executed analogously to the rule competition in production.

For example, the rules available to the listener might only allow the utterance to be interpreted as "eat<dog, #>". But if the cue "eat<dog, meat>" is available, then the strength of the cue

[2] In cue selection, if a randomized number is less than *RC*, the intended meaning is chosen as the cue, otherwise, a randomly chosen meaning from the semantic space is chosen as the cue. Identical cues are not allowed.

“eat<dog, meat>” (here, 0.5) is included in the calculated of the combined rule strength, CS_{listen}. Meanings that are supported by both the available rules and cues tend to be the inferred, particularly in the early stages of the emergence of the language.

The listener infers the meaning based on his winning rules having the greatest combined CS_{listen}. If this CS_{listen} exceeds a certain threshold, the listener sends a positive feedback to the speaker indicating the listener’s confidence in the comprehension. Otherwise, a negative feedback is sent, indicating that the listener was either unable to infer a meaning or else was not confident of correctly inferring the intended meaning. Based on this feedback (rather than on a direct meaning check), both the speaker and the listener perform rule adjustment to their available rules, increasing the strengths of the winning rules and decreasing those of the losing rules.

This process of indirect meaning transference and feedback makes an indirect connection between the production and comprehension of different speakers. The interaction of linguistic information and nonlinguistic information, rather direct meaning check, direct the emergence of mutual understanding.

3 Results and discussions

In this model, we simulate iterative *concurrent* communications among randomly chosen agents. At one time step, many communications among different pairs of agents happen simultaneously. In this section, we begin by introducing several indices that we use to test the behavior of this communication system. Then, we trace the emergence of a compositional language from a holistic signaling system, and the coevolution of the lexicon and syntax. We also discuss some factors that determine how effectively a compositional language is acquired. Finally, we test the emergence of language in a heterogeneous population, and demonstrate that certain types of heterogeneity do not significantly influence the behavior of this communication system.

3.1 Indices to test the behavior

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

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

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

$$UR = \frac{\sum_{i,j} \text{number of understable meanings between agent } i, j}{\text{number of all possible pairs of } i, j} \tag{4}$$

In models using direct meaning transference, only the *RE* of the emergent language is tested. But the ability of a language to represent meanings should not only consider the *RE*, but also the characteristic of *displacement* (speech signals can refer to objects and events that are removed from the present in both space and time, and can be accurately understood [24]). The *UR* can evaluate such characteristics of the emergent language. A *mature* language shared by agents should be a language with high *UR* (over 80%, say) in communications using this language.

c) *Convergence time* (*CT*) —how many iterations of communication are required to achieve a mature language.

3.2 Coevolution of lexicon and simple syntax (word order)

The coevolution[3] of the lexicon and syntax 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 the transition from an initially holistic signaling system to a compositional language. The *UR*, shown in Figure 4(a), undergoes an S-shaped evolution, indicating the emergence of a common lexicon — this is similar to the result of Ke et al.'s model [5]. Figures 4(b–c) show the convergence of the syntax from all possible sequential order rules to the dominant word orders; the curves trace the average strength of each of the eight order rules. Two dominant word orders emerge from the initial state of no syntax, one for each of the two meaning types. No prior bias is conferred to any particular word order; each is *a priori* equally likely.

[Figure 4 is about here.]

The coevolution process typically proceeds as follows: At first, agents only understand meanings that are expressed by 6 initial holistic rules. Then, more holistic rules emerge through random creation, gradually increasing the *RE* of holistic rules. The *UR* also relies on holistic rules. Later on, recurrent patterns emerge by chance, and their acquisition greatly increases the *RE* of the compositional rules, triggering the transition from a holistic signaling system to a compositional language. A consistent word order becomes necessary as agents come to use more compositional rules to encode meanings, so triggering the convergence of syntax. The *UR* relies increasingly on the compositional rules, although the use of compositional rules may cause some meanings that were initially understandable when

[3] Without further statement, the simulation condition is: 10 agents, 500*5 communications, *RC*=0.7

expressed by holistic rules to be misunderstood, causing the *UR* to briefly drop slightly. However, the recurrence of these compositional rules in successive communications allows them to win the competition with holistic rules, and finally makes possible the emergence of a common lexicon and with dominant word orders. Furthermore, if the meaning space is increased in size gradually as extant meanings become understandable, the acquisition of linguistic rules to express salient meanings follows a similar curve.

Taken together, Figures 4(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 use of compositional rules triggers the convergence of syntax, which in turn boosts the convergence of the lexicon; the sharp increase of *UR* and the strengths of the dominant order rules are almost synchronized.

3.3 Homophone avoidance and reliability of cues (RC)

Several constraints are necessary to acquire the above results. Because of the existence of unreliable cues (without which the feedback would still be equivalent to direct meaning transference) and the lack of context (meanings expressed in communications are independent of each other), some internal strategies are required to avoid ambiguity in the utterances caused by homonymous rules, which inevitably emerge during rule acquisition. We assume that homonym avoidance is one such strategy. Such homonym avoidance can be traced in some research (e.g., [21]) on children's language acquisition — 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 ("agent"/"patient" or "predicate").

In this model, homonym avoidance is implemented in rule adjustment: one randomly selected form among a set of homonymous rules within the same semantic category has its strength increased, while the homonymous rules with which it competes decreased. Statistical results show that without homonym avoidance, the *UR* is very low (see Table 1). Table 2 shows the common rules acquired under homonym avoidance. Though some homonymous rules with meaning constituents in different categories exist, the high *UR* indicates less misunderstanding.

[Table 1 is about here. Table 2 is about here.]

A high *RC* is obviously necessary for the emergence of language. However, even when the nonlinguistic information is highly reliable, with no internal homonym avoidance, a mature language cannot be acquired. Figure 5 shows the *UR* for different values of *RC*, both with and without homonym avoidance. With homonym avoidance, *UR*, increasing along with *RC*, is higher than that without homonym avoidance. Indeed, without homonym avoidance, the peak *UR* is not very high, even when the cues are always reliable (i.e., *RC* =1.0). Although accurate interpretation is possible with the aid of highly reliable nonlinguistic information, homonymous rules can still cause misunderstandings when such information is absent. The language that then emerges does not exhibit the characteristics of displacement.

[Figure 5 is about here.]

The above discussions suggest that linguistic communication through indirect meaning transference requires internal strategies to avoid ambiguity in the comprehension of utterances, homonym avoidance being one such strategy. In addition, in order to efficiently acquire a language, highly reliable nonlinguistic information is desirable.

3.4 Influences of heterogeneous population

The influences of two types of heterogeneity on the language emergence are discussed here.

1. Storage capacity of the buffer and rule list. Heterogeneous capacity is simulated by treating the size of the buffer and of the rule list as random variables, each having a Gaussian distribution (rounded off to the nearest integer value). We set the average value of each distribution to the same values as was used in the homogeneous situation in which all agents have same capacity and the variance to 5. Figure 6(a) shows *CT* for different buffer capacities but a fixed rule list capacity (40). Figure 6(b) shows *CT* for different rule list capacities but a fixed buffer capacity (40). The solid line indicates the homogeneous condition and the dashed line indicates the heterogeneous condition.

[Figure 6 is about here.]

The buffer capacity affects the rule generalization. When abundant M-U mappings can be stored in the buffer, the probability for recurrent patterns among them is high and many new rules can be generalized simultaneously. However, a bigger buffer needs more communications to fill it, which may delay the rate at which rules are updated. The *CT* curve in Figure 6(a) shows that with the increase of the buffer capacity, *CT* increases limitedly. This indicates that the second effect is stronger than the first, but not explicitly. For most reasonable sizes of the buffer capacity, most simulations converge to languages with a high *UR*, even in heterogeneous conditions.

The rule list capacity determines rule storage. The bigger the capacity, the more easily agents store new rules. This can accelerate the convergence process. However, for a fixed semantic space, large rule capacity introduces redundancy. For example, in the current model, 12 word rules are sufficient to express all integrated meanings. However, when the system converges, there are empty spaces in some agents' rule lists, and some rules which are not shared by all agents are still stored. As shown in Figure 6(b), although with the increase of the rule list's capacity, *CT* decreases limitedly indicating the first effect, it is not explicit. For a reasonable size of the rule list capacity (over 12), most simulations converge to languages with high *UR*, even in heterogeneous conditions.

2. Rule acquisition ability. These linguistic behaviors include the ability of random creation (*Creating Rate* (*CR*), which controls the rate of creation of salient linguistic utterances), and the ability to detect recurrent patterns (*Detecting Rate* (*DR*), which controls the rate of acquiring new rules from available M-U mappings). Heterogeneous abilities are simulated using the Gaussian distribution — both *CR* and *DR* have their mean value set to the same value as in the homogeneous condition, and their variance set to 0.2. Figure 7(a) shows *CT* for different values of *CR* but a fixed *DR* (0.5). Figure 7(b) shows *CT* for different values of *DR* but a fixed *CR* (0.5).

[Figure 7 is about here.]

Obviously, without random creation no salient linguistic utterance will emerge. Also, without detection of recurrent patterns, although many linguistic rules might emerge, similarities among them will only occur by chance, the result being that emergence of a common lexicon will be virtually impossible! Second, given a certain value of *DR*, if *CR* is too small, although agents can detect recurrent patterns, acquisition of new rules will be delayed by the insufficient number of randomly created lexical rules. Similarly, given a certain value of *CR*, if *DR* is too small, although many M-U mappings might be available, few recurrent patterns are extracted and so few holistic rules are decomposed. This delays the convergence of the system. Except in these extreme cases, however, most simulations converge to languages with high *UR*, even in heterogeneous conditions.

To summarize, we infer that a population with certain types of heterogeneity (e.g. storage capacity and linguistic behavior) can still allow a mature language to emerge. This shows the robustness of the self-organization process in our model, that is, a language can be effectively acquired by a population of agents, withstanding interference caused by either external noise or internal parameters.

4 Conclusions and future directions

In this paper, a computational model of language emergence is presented addressing the limitations of current computational models. A coevolution of the lexicon and syntax at the protolanguage level is simulated by concurrent linguistic communication using indirect meaning transference and feedback without direct meaning check. In order to avoid ambiguity in utterances, strategies, such as homonym avoidance, are found to be necessary. In addition, we show that certain heterogeneities do not significantly influence the likelihood of emergence of a compositional language using this framework.

There are still important aspects of language emergence still to be studied: First, this model shows that an emergent language can be acquired through iterative communications, to enhance its realism, agent replacement and teaching, as well as a separation of rule lists into production and comprehension (e.g., [5]), should be considered. Second, homonym avoidance is directly built into the current model, but it would be more realistic to observe this as an emergent phenomenon of the model. Third, embedding and recursion, rarely touched by current models (exc. Kirby's [8]), are important features of language. Modeling the emergence and interpretation of embedded meanings, based on linguistic and nonlinguistic information, would be a major advance. Finally, it would be worthwhile to simulate more realistic communication, such as one-speaker-many-listener, and inter/intra-group communication.

Acknowledgements

This work is supported by City University (Hong Kong) grant No.9010001-570, 9040781-570. The authors offer thanks to Prof. T. Lee for helpful discussions and fruitful suggestions.

Appendix A: An example of rule competition in production

Assume that the speaker wants to express “fight<dog, fox>”, according to his own rule list, there are three ways to express this meaning and related rules are activated: 1) using one holistic rule, no word order rule is considered; 2) using three word rules, all six possible word order rules are applicable, so the strongest one (*SVO*(0.5)) is chosen; 3) using one word rule and one phrase rule, the utterance of the phrase rule restricts that only *VSO*(0.4) and *OSV*(0.3) are applicable word order rules, so the strongest one *VSO*(0.4) is chosen. In each condition, CS_{speak} is calculated. The strongest CS_{speak} (CS3) is chosen. Then, the lexical rules and word order rules (Bold ones in Figure A.1) contribute to *CS3* are the speaker’s winning rules, the utterance based on them are built up and sent to the listener.

Figure A.1 is about here.

References

1. Gyori, G. Ed. Language Evolution: Biological, Linguistic and Philosophical Perspectives. Frankfurt am Main: New York, 2001.
2. Christiansen, M.; Kirby, S. Eds. Language Evolution. Oxford University Press: New York, 2003.
3. Cangelosi, A.; Parisi, D. Eds. Simulating the Evolution of Language; Springer Verlag: London, 2001.
4. Wagner, K.; Reggia, J. A.; Uriagereka, J.; Wilkinson, G. S. Progress in the Simulation of Emergent Communication and Language. Adaptive Behavior 2003, 11, 1; pp. 37–69.
5. 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; pp. 41–54.
6. Munroe, S.; Cangelosi, A. Learning and the Evolution of Language: The Role of Culture Variation and Learning Cost in the Baldwin Effect, Artificial Life 2002, 8, 4; pp. 311–340.
7. 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.
8. Smith, K.; Kirby, S.; Brighton, S. Iterative learning: A Framework for the Emergence of Language, Artificial Life, 9; pp. 371–386.
9. 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.
10. Kirby, S. Natural Language from Artificial Life. Artificial Life 2002, 8, 2; pp. 185–215.
11. Quine, W. V. Word and Object. MIT Press: Cambridge, MA, 1960.

12. Romaine, S. Language in society, an introduction to sociolinguistics. Oxford University Press: Oxford, 1994.
13. Shore, C. M. Individual Differences in Language Development. Sage Pub: Thousand Oaks, Calif., 1995.
14. 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.
15. Calvin, W. H.; Bickerton, D. Lingua ex machine: reconciling Darwin and Chomsky with the human brain. MIT Press: Cambridge, MA, 2000.
16. Wray, A. Formulaic Language and the Lexicon. Cambridge University Press: New York, NY, 2002.
17. Schoenemann, P. T. Syntax as an Emergent Characteristic of the Evolution of Semantic Complexity. Minds and Machines 1999, 9; pp. 309–346.
18. Hauser, M. D. The Evolution of Communication. MIT Press: Cambridge, MA, 1996.
19. 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.
20. Wray, A. Protolanguage as a Holistic System for Social Interaction. Language & Communication 1998, 18; pp. 47–67.
21. 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.
22. 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, California, 1986; pp. 593–623.

23. Clark, E. The principle of contrast: A constraint on language acquisition. In: Mechanisms of language acquisition; MacWhinney, B. Ed.; Lawrence Erlbaum Assoc.: Hillsdale, NJ, 1987; pp. 1–33.

24. Hockett, C. F. The Origin of Speech. Scientific American, 203; pp. 88–96.

Figure and Table Captions

Figure 1: 2-level memory system.

Figure 2: Rule generalization examples.

Figure 3: Communication through indirect meaning transference.

Figure 4: Coevolution of lexicon and syntax: (a) Lexicon emergence; (b) Syntax convergence for "predicate<agent>" meanings; (c) Syntax convergence for "predicate<agent, patient>" meanings.

Figure 5: *UR* under different *RC* with and without homonym avoidance.

Figure 6: Storage capacity: (a) *CT* under different buffer's capacity and a fixed rule list's capacity; (b) *CT* under a fixed buffer's capacity and different rule list's capacity.

Figure 7: Linguistic behavior: (a) *CT* under fixed *DR* and different *CR*. (b) *CT* under different *DR* and fixed *CR*.

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

Table 1: Statistical results of *UR* with and without homonym avoidance.

Table 2: Common rules among all agents in one simulation with homonym avoidance.

Figures and Tables

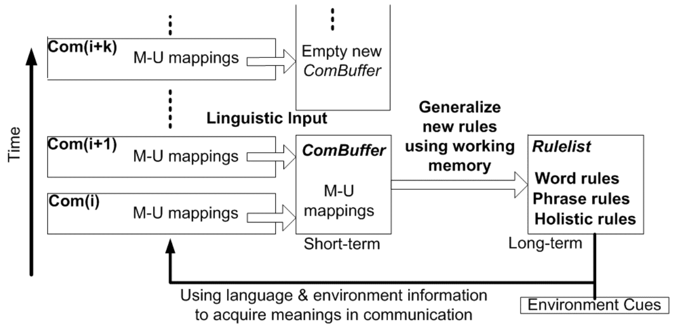


Figure 1: 2-level memory system.

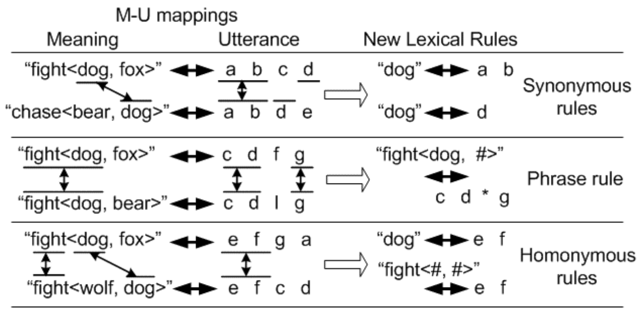


Figure 2: Rule generalization examples.

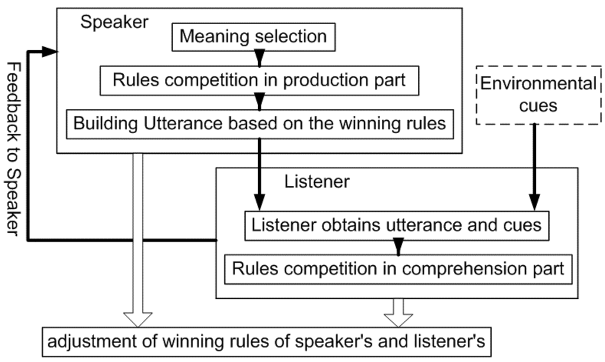
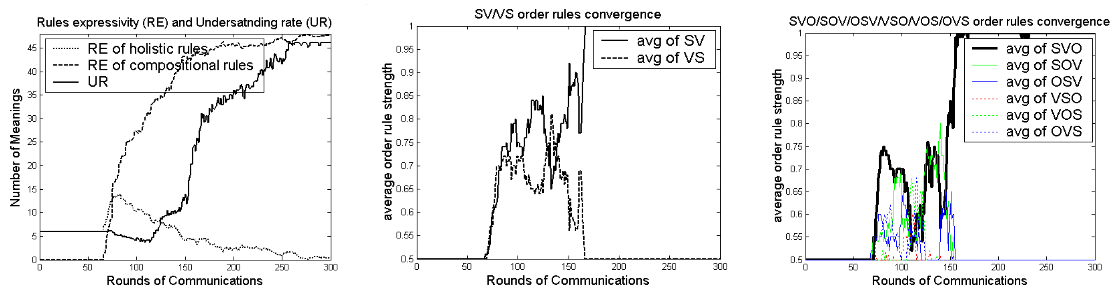


Figure 3: Communication through indirect meaning transference.



(a) (b) (c)

Figure 4: Coevolution of lexicon and syntax: (a) Lexicon emergence; (b) Syntax convergence for “predicate<agent>” meanings; (c) Syntax convergence for “predicate<agent, patient>” meanings.

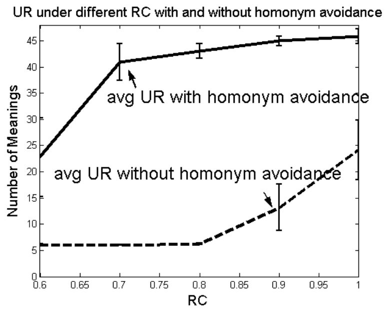
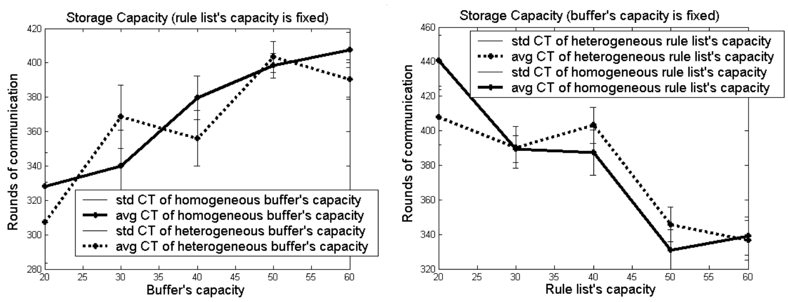
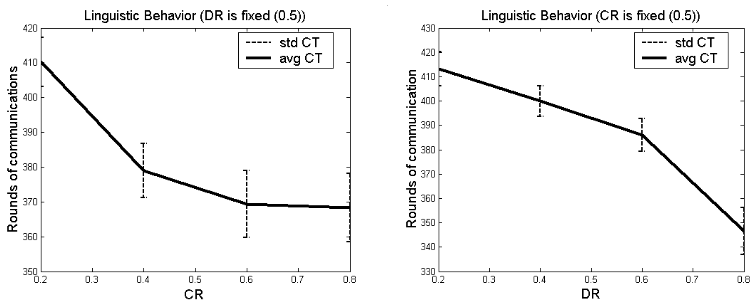


Figure 5: *UR* under different *RC* with and without homonym avoidance.



(a) (b)

Figure 6: Memory capacity: (a) *CT* under different buffer’s capacity and a fixed rule list’s capacity; (b) *CT* under a fixed buffer’s capacity and different rule list’s capacity.



(a) (b)

Figure 7: Linguistic behavior: (a) *CT* under fixed *DR* and different *CR*; (b) *CT* under different *DR* and fixed *CR*.

Production Part Meaning to express: *" dog fight fox"*

	Activated rules	Applicable word order rules	Combined Strength (CS)
1 holistic rule	*"fight<dog, fox>"←→/a b/* (0.6)		*CS1 = 0.6*
3 word rules	*"dog"←→/b/* (0.8) *"fight<#, #>"←→/c e/* (0.5) *"fox"←→/g/* (0.3)	*SVO* (0.5)	CS2 = 1/2(1/3(0.8+0.5+0.2)+0.5) = 0.5
1 word rule 1 phrase rule	***"dog"←→/b/* (0.8)** ***"fight<#, fox>"←→/e * f/* (0.8)**	***VSO* (0.4)** *OSV* (0.3)	**CS3 = 1/2(1/2(0.8+0.8)+0.5) = 0.65**

Utterance built up: */ e b f/*

Figure A.1: Example of rule competition in production

Table 1: Statistical results of UR with and without homonym avoidance.

```
No. of Simulation: 10
(1): using homonym avoidance
(2): not using homonym avoidance
    avg UR   std UR
(1) 40.3022  2.2201
(2) 15.7744  8.3388
```

Table 2: Common rules among all agents in one simulation with homonym avoidance.

```
Rounds of Communication = 294

Common Rules for all agents: 13
" cry<#>" -> /24 11/
" run<#>" -> / 5/
" suck<#, #>" -> / 2/
" chew<#, #>" -> /21/
"fight<#, #>" -> /18/
"chase<#, #>" -> / 9/
"  dog" -> /16/
" bear" -> /23/
"  fox" -> / 2/
" wolf" -> /20/
" wolf" -> / 5/
"water" -> /21/
" meat" -> /15  2/

Rules Expressivity = 48.00 ;
Average Understanding Rate = 47.64 (99.26%)
```