A Computational Framework to Simulate the Coevolution of Language and Social Structure

Tao Gong*, Jinyun Ke, James W. Minett and William S-Y. Wang

Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong

*tgong@ee.cuhk.edu.hk

Abstract

In this paper, a multi-agent computational model is proposed to simulate the coevolution of social structure and compositional protolanguage from a holistic signaling system through iterative interactions within a heterogeneous population. We implement an indirect meaning transference based on both linguistic and nonlinguistic information in communications, together with a feedback without direct meaning check. The emergent social structure, triggered by two locally selective strategies, *friendship* and *popularity*, has small-world characteristics. The influence of these selective strategies on the emergent language and the emergent social structure are discussed.

1. Introduction

Recently, computational modeling of language evolution has grown rapidly, as exemplified by many anthologies and reviews (Standish et al 2003; Cangelosi and Parisi 2001; Wagner et al. 2003). Many computational models, based on evolutionary or artificial life theories, have been reported, such as the neural network models (e.g., Munroe and Cangelosi 2002), the vocabulary coherence model (e.g., Ke et al. 2002), and the Iterative Learning Framework (ILM) (Kirby 2002; Smith et al. 2003). These 'emergent' models (according to Schoenemann 1999) share several assumptions related to language development. However, there are still several limitations.

First, most of them assume *direct meaning transference* (excluding Munroe and Cangelosi (2002)) in the interactions among agents, i.e., the intended meanings encoded in linguistic utterances and sent by speakers are always accurately available to listeners. However, it is obvious that expression and interpretation are independent in speakers' and listeners' minds, and that there are at least no *direct* connections among them. Other channels, such as pointing while talking or primitive feedback, can only provide a certain degree of confirmation. Interpretation is a complex process requiring linguistic and nonlinguistic information. It is unrealistic to assume direct meaning transference.

Second, these models either fail to model syntax (e.g., Ke et al. 2002), build in the syntactic features (e.g., Munroe and Cangelosi 2002), or else do not adopt a coevolutionary view of the emergence of syntax and the lexicon (e.g., Smith et al. 2003). However, syntax in language is likely to have become conventionalized through language use, rather than as the result of an innate, grammar-specific module (Schoenemann 1999). The syntax is assumed to have emerged because of a pre-adapted cognitive capacity reflected in other cognitive processes, i.e. the sequencing ability, which can be attested in other primates and pre-language infants (Christiansen and Ellefson 2002). The emergence of the lexicon and the convergence of syntax should be interwoven, i.e. they should coevolve.

Third, these models often use random interactions, which disregard the influence of social structure. Although sociological research has studied structures that have emerged based on stable or global factors, very little research has touched upon the emergence of structure based on the evolution of language. Mutual understanding based on the evolving language can be a factor to trigger social structure and so is worth studying. Recently, the rapidly developing *complex networks* theory (Newman 2003) provides an efficient methodology to study it.

Fourth, most current models are based on homogeneous populations. However, sociolinguists have shown there to be dramatic variations in the speech community and various dichotomies in the learning styles of children (Shore 1995). Heterogeneity of natural characteristics and linguistic behaviors among agents should therefore be considered in the computational models that are adopted.

Addressing these limitations and based on the 'emergent' theory of Wray (2002), we present a computational model which uses an indirect meaning transference and simulates a coevolution of lexicon and syntax (in the form of simple word order) during the transition from a holistic signaling system to a compositional language. Based on mutual understanding of the evolving language and two locally selective strategies, this model also simulates the emergence of social structure during the emergence of language.

The rest of the paper is organized as follows: in Section 2, we describe the model; results and discussions are presented in Section 3; finally, we draw some conclusions and point out some future directions in Section 4.

2. Description of the model

This model is basically a linguistic communication game among independent agents in a population, focusing only on horizontal transmission. Agents express and interpret two types of meanings: “predicate<agent>”, such as “run<tiger>”, and “predicate<agent, patient>”, such as “chase<tiger, wolf>” or “eat<tiger, meat>”. Nonlinguistic information (*Cues*) is used to assist the meaning interpretation, especially meanings like “chase<tiger, wolf>” — without cues, it is not clear who is chasing whom, especially in the early stages of the language evolution. Cues, pragmatic meanings describing environmental events, are integrated meanings all with the same strength, e.g., “fight<dog, cat>” (0.5); 0.5 is the strength. Cue Reliability (*CR*) manipulates the probability that the intended meaning is contained in one of the cues.

This model uses a rule-based system to represent the language. Linguistic rules includes *lexical rules* (meaning-utterance mappings), such as holistic, phrase and word rules, and *word order rules* which cover all possible sequences to regulate utterances for expressing integrated meanings with two or three meaning constituents, such as “agent first, predicate last, patient medial” (denoted by SVO for simplicity). Rule strengths, numerically indicating the frequency of successful use of the rules, can be adjusted by self-organizing strategies in rule competition. Agents start with a holistic signaling system (sharing a set of common holistic rules) and no dominant word order (all word order rules have the same strengths). Through iterative communications, a common set of rules shared by all agents indicates the convergence of the language.

A two-scale storage system, inspired by a *Classifier System* based model (Holland 2001), is used to handle lexical rules. It includes a buffer (storing ‘previous experiences’ — meaning-utterance mappings (M-U mappings) obtained in previous communications) and a rule list (storing ‘linguistic knowledge’ — lexical rules generalized from M-U mappings in the buffer when it is full). Rules in the rule list are used to express integrated meanings and interpret utterances, together with nonlinguistic information, in future communications.

There are two mechanisms for agents to acquire new rules: *random creation* in meaning expression (as in Kirby et al.’s model), i.e., with a certain probability, speakers create lexical rules (holistic or compositional) to help their production of integrated meanings, and *rule generalization*, i.e., a flexible detection of *recurrent patterns* (recurrent constituents in meanings and recurrent syllables in utterances among two M-U mappings) without syntax or location restriction. Rule generalization occurs when the buffer is full. Figure 1 shows some rule generalization examples. By extracting recurrent patterns as new compositional rules, some holistic signals are decomposed.

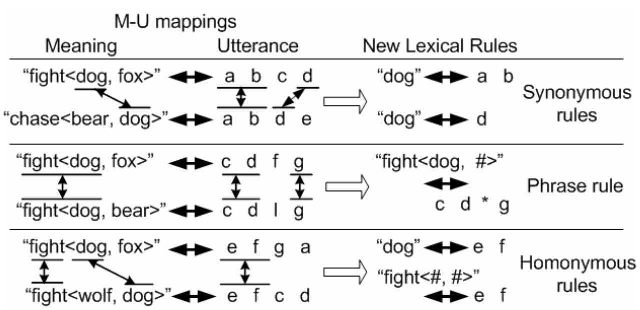


Figure 1: Examples of rule generalization. (#, *: matching pragmatic meaning items and syllable(s))

Synonymous and homonymous rules emerge inevitably during the execution of these two mechanisms because there is no clear access to other agent’s language, and flexible rule generalization does not consider the existent rules. Due to the limited size of storage, the lack of context (meanings expressed in communications are independent of each other) and unreliable cues (otherwise, it would still be *direct meaning transference*), homonym avoidance, in which a form that is used ‘successfully’ is reinforced and others which it competed are weakened, is built in. As for synonyms, agents randomly learn one form from a set of synonymous rules based on the *Principle of Contrast* (Clark 1987).

Heterogeneity means different agent can have different buffer and rule list sizes and different linguistic abilities for random creation and rule generalization.

Communications in this model are *concurrent*; during each time step, many communications between different pairs of agents happen simultaneously. An indirect meaning transference is implemented in communication. Communication proceeds as follows (summarized in Figure 2). First, the speaker selects a meaning to express. Based on his current linguistic rules, the speaker encodes the selected meaning into the utterance of his winning rules which have the highest combined rule strength, CS_{speak}, calculated from the formula

$$CS_{speak} = \begin{array}{l} \mathrm{Str}\left(\text{combinable activated rules}\right) \\ + \mathrm{Str}\left(\text{applicable word order rules}\right) \end{array} \qquad (1)$$

The utterance, built up accordingly, is transferred to the listener, who attempts to interpret the utterance. The listener sometimes also receives cues from the environment. Interpretation involves a more complex process of rule

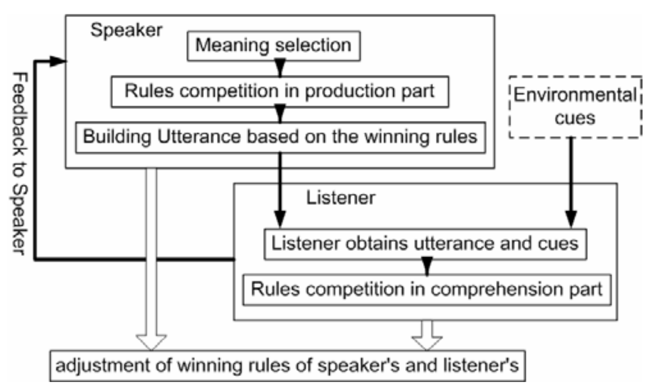


Figure 2: Indirect meaning transference.

competition, considering not only linguistic but also nonlinguistic information, in the listener's mind, based on the combined rule strength CS_{listen} of his linguistic rules:

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

The listener interprets the meaning based on his winning rules. If the combined rule strength of the listener's winning rules exceeds a certain threshold, a positive feedback is sent to the speaker indicating the listener's confidence in the interpretation. Otherwise, a negative feedback is sent, meaning that the listener was either unable to infer a meaning or else was not confident of inferring the intended meaning. Finally, based on this feedback rather than on a direct meaning check, both the speaker and the listener adjust their own rules, increasing the strengths of the winning ones and decreasing those of the losing ones if the feedback is positive. During the whole process of communication, expression and interpretation are independent and the interpretation is based on the interaction of linguistic and nonlinguistic information.

Mutual understanding based on the evolving language can influence the possibility of future communication between these two agents. A fully-connected weighted network is used to indicate the social relationships among members; see Figure 3. The connection weight, adjusted in both successful and failed communications, indicates the cumulative probability of successful communication between the two agents. Once the connection weight exceeds a threshold, a permanent edge is built. However, such permanent connections can still be broken after many failed communications. Agents permanently connected to each other are linguistic 'friends', and have a higher chance to understand each other. The number of permanent edges of one agent indicates his linguistic 'popularity', i.e., his propensity to communicate successfully with other agents. Friendship and popularity are local factors focusing on individual agents.

To enhance the realism of the model, we introduce a *local-view* assumption similar to Li and Chen's paper (2003), i.e., in one communication, one agent can only view several agents (local-view) instead of all group members and communicate with some of them.

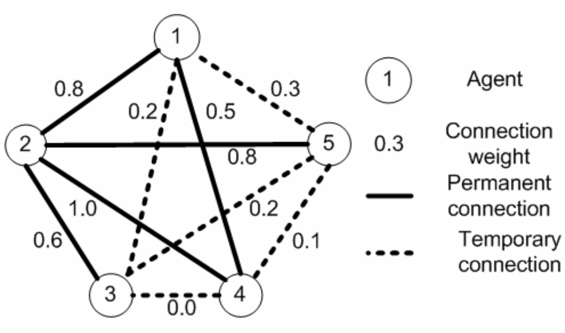


Figure 3: Social network used in this model.

We run two types of simulation. **Sim.1:** Each generation, each agent selects the agents in his local-view, those to whom he is permanently connected having a higher chance to be chosen. Agents communicate with a subset of agents in their local-view, preferring to communicate with agents having higher popularity. A new generation begins after all agents have executed this process. **Sim.2:** Each generation, agents randomly select the agents in their local-view and randomly attempt to communicate with some of them.

In this model, heterogeneities, such as different buffer or rule list size, different random creation and generalization rates (simulated by the Gaussian distribution), and different mechanisms to acquire new rules (simulated by random assignment), are allowed to make the model more realistic.

Finally, several major factors are used to study the performance: a) the *understanding rate* (*UR*), defined by

$$UR = \frac{\sum_{i,j}\left(\text{number of understandable meanings between agent i and j}\right)}{\left(\text{number of all possible pairs of } i, j\right)} \quad (3)$$

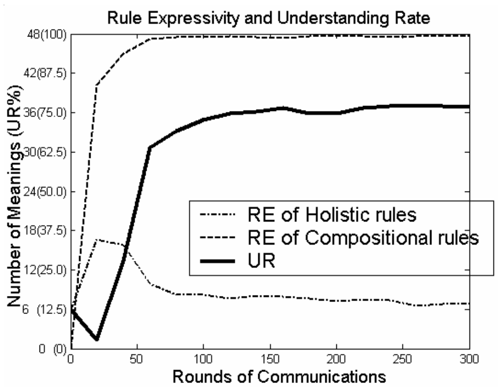
indicates the average number of meanings understandable by every pair of agents in the population based on linguistic information only — it tests the real representation ability of the acquired language, considering not only the expressivity of meanings that might not happen in immediate environment, but also the understandability of these expressions (Displacement) (Hockett 1960); b) the *degree distribution* (P_k) indicates the distribution of the number of permanent connections (degree) versus the number of agents having such degree; c) the *number of sub-clusters* of connected agents, indicates the divergence within the population. Other parameters, such as the *rule expressivity* (*RE*), *convergence time* (*CT*) (the number of rounds of communication by which the highest *UR* is reached), the *average degree* (*AD*), the *clustering coefficient* (*C*), and the *average shortest path length* (*L*) are also considered. The following simulations' conditions are: 50 agents, 500 generations, *RC*=0.8. Buffer size=35±5, Rulelist size=45±5. Random creation rate=0.5±0.2, Rule generalization rate=0.5±0.2.

3. Results and discussions

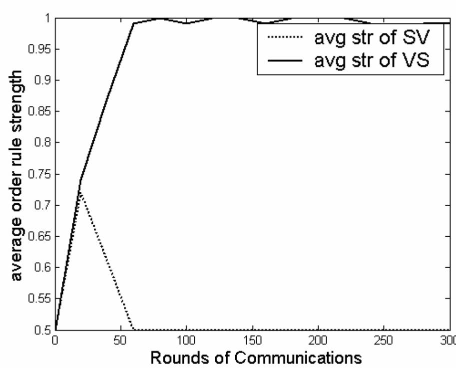
3.1 Coevolution of lexicon and simple syntax

Coevolution of the lexicon and syntax is simulated in this model, as shown in Figure 4. Figure 4(a) shows the *RE* of both holistic rules and compositional rules; the decrease of the former and the increase of the latter show the transition from initially holistic signals to a compositional language. The *UR* in Figure 4(a) shows the convergence to a common lexicon. The *UR* undergoes an S-shaped evolution, matching the result of Ke et al.'s model (2002). The *RE* of compositional rules used in combination increases rapidly, but the use of compositional rules may cause some meanings understandable when 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 the holistic rules, which finally makes possible the emergence of a common lexicon.

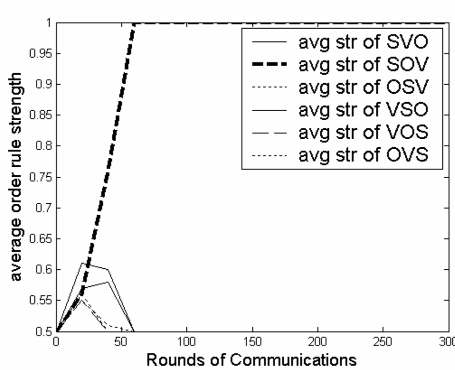
Figures 4(b–c) show the convergence of the dominant word order from all possible sequential order rules; the curves trace the average strength of each of the eight order rules. Mutual understanding requires not only common lexical rules but also a shared syntax to combine compositional rules. Two dominant word orders emerge from the initial state of no syntax, one for each of the two meaning types. There is no prior bias conferred to any particular word order; each is initially equally likely. Finally, combining Figures 4(a–c), we observe the coevolution of lexicon and syntax: the use of compositional rules triggers the convergence of syntax, which in turn boosts the convergence of the lexicon; the sharp increase of the *UR* and the dominant order rules' strengths are almost *synchronized*.



(a)



(b)



(c)

Figure 4: Coevolution of Lexicon and Syntax (local-view=10, com. of each agent=5, Sim.1 (Sim.2 is similar)). (a) Lexicon convergence; (b),(c) Syntax convergence for "predicate<agent>" and "predicate<agent, patient>" meanings.

3.2 Emergence of social structure

During the emergence of language, the selective strategies trigger a global social structure based on the mutual understanding of the evolving language. The *AD* and *C*, also following an S-curve, trace the emergence of the social structure (see Figures 5(a–b)). Due to the restriction of the selective strategies, the *AD* and *C* of Sim.1 is smaller. Besides, these strategies have their own influences. For example, the local, 'self-centered' strategy of friendship can trigger an earlier increase of *AD* and *C* compared with Sim.2. It also triggers an earlier emergence of sub-clusters (see Figure 5(c)). The high *C* and low *L* indicate that the emergent social structures of both simulations have small-world (Watts 1999) characteristics. However, their structures are different due to the influence of the friendship and popularity strategies (see Figure 5(d)). In Sim.2, a network that is almost fully-connected emerges, with most agents having the same, high degree. However, in Sim.1, the degree distribution is more uniform. Although almost all members can understand each other, the degrees of some agents do not increase much. This is because friendship tends to restrict agents to communicate with other agents belonging to their local-view and popularity only triggers a local convergence within the local-view. Agents within the local-view might have intensive connections with one another, but they do not connect to outsiders frequently. This local centralization prevents the degrees of some agents from greatly increasing.

On the other hand, different local-view sizes in the selective strategies can influence the emergent social structure. With the increase of the local-view size, the influence of friendship is gradually reduced, which breaks down the local convergence. Then, the degree of every agent increases gradually. This can be seen in Figure 6(a).

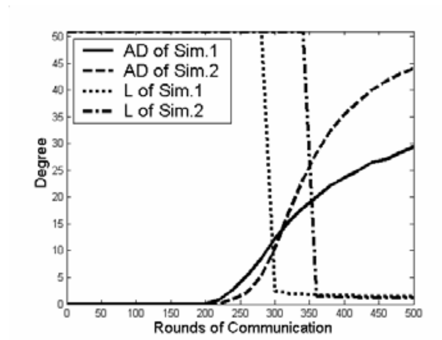
As for the language that emerges, with the increase of the local-view size, the centralization is more global; there seems to be an optimal local-view size for peak *UR* (Figure 6(b)) — either too much 'democracy' or too much 'dictatorship' cannot achieve the best *UR*. Actually, centralization around some agent(s) has two effects. First, popular agents connect many unpopular agents, like a network hub. Centralization around them can increase the chances for unpopular agents to exchange information, and thereby accelerate the convergence of linguistic rules. On the other hand, effective information transference between two agents (say, A_1 and A_2) requires direct connection or connection through a 'stable' intermediary (say, a popular agent, whose internal rules do not change much, so that the information received by A_2 via the popular agent does not change much from the original information sent by A_1). However, with the increase of global centralization, other agents have higher chances to contact the popular agent and influence his rules. This makes the popular agent unstable, i.e., although the input information is the same, the output information differs greatly from time to time. This greatly affects the information transference and the convergence of linguistic rules between A_1 and A_2 through the popular agent. To compromise these two contradictory factors, the optimum performance occurs at an intermediate level of centralization.

Finally, the structure of Sim.1 is triggered by the social strategies which are based on the evolving language. The evolution of the language has its influence on the final result; these social strategies, if based on a non-evolving language, can trigger a local-world (Li and Chen 2003) (see Figure 6(c)) or, if local-view is the whole group, a scale-free (Barabási 1999) structure. However, with an evolving language, this model has no such structures. This shows that when using language-related factors to trigger structure, one should consider the evolution of language.

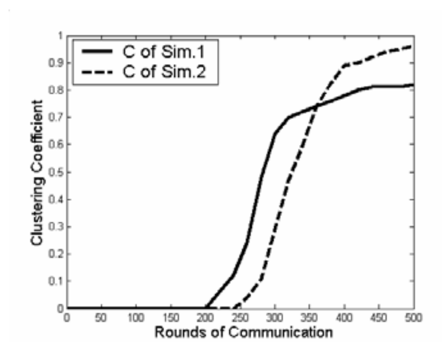
4. Conclusions and future directions

Considering the evolutionary point of view and real communication situations, coevolution and indirect meaning transference are more realistic. Mutual understanding when using the evolving language can be a factor to trigger a social structure during the emergence of language and the emergent social structure has small-world characteristics. Without considering the evolving property of the language, results are different, which indicates the influence of language's evolving property on the emergent social structure. This model imports complex networks theory to study the emergent social structure based on evolving factors using artificial life modeling, an appropriate tool to simulate and study the influences of these evolving factors.

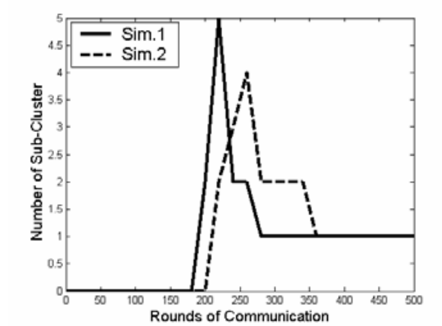
Several future directions are promising. First, the current model can be 'situated' in an artificial world, and the *Genetic Algorithm* (GA) (Holland 1975) used to evaluate the fitness, both with and without language. Second, to enhance the realism of the current model, other types of communication, such as 'one speaker, multiple listeners', need to be simulated to further study the language emergence in a more realistic situation. Third, in the social structure aspect, it is worth comparing the structures triggered by linguistic communication with those triggered by other nonlinguistic factors. Besides, when an agent chooses other agents to communication, he should consider not only linguistic-related factors such as 'friendship', 'popularity', but also other factors, such as kinship, economic status.



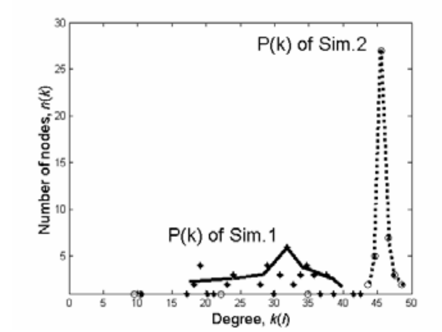
(a)



(b)

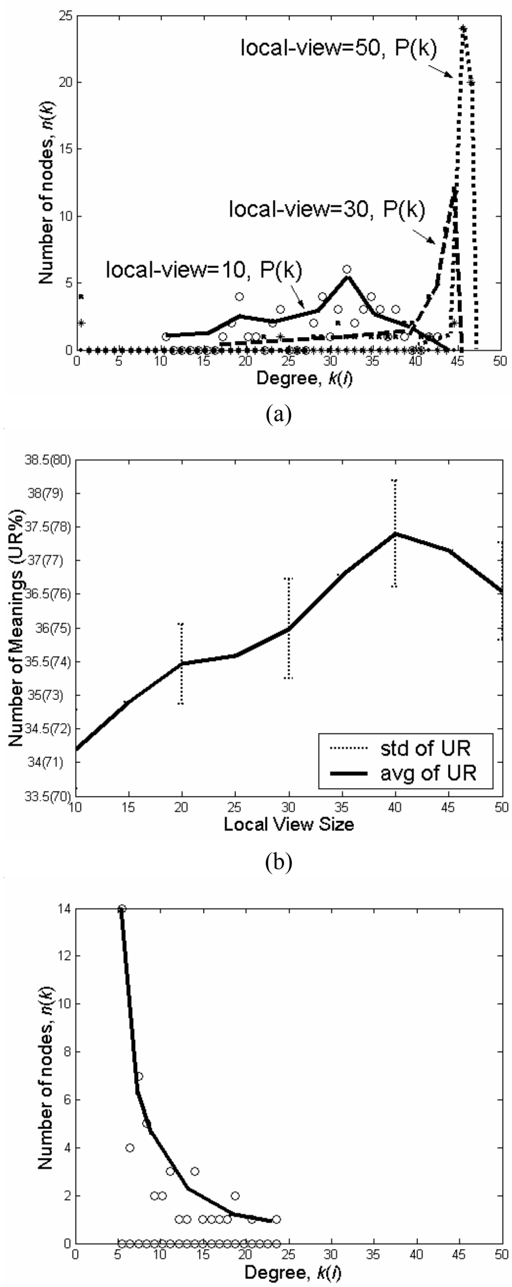


(c)



(c)

Figure 5: The emergent social structure in two simulations. (a) AD and L; (b) C ; (c) Sub-Clusters; (d) P_k.



(a)

(b)

(c)

Figure 6: Local-view size effects. (a) P_k; (b) *UR*; (c) P_k of local world structure (based on Li and Chen (2003)'s model)

Acknowledgments

This work is supported by City University of Hong Kong grants No. 9010001-570, 9040781-570. The authors offer thanks to Profs. J. Holland, G. R. Chen and T. Lee for useful discussions and suggestions.

References

Barabási, A.L. 1999. Mean-field Theory for Scale-free. Physica A 2(72): 172–182.

Cangelosi, A., and Parisi, D. eds. 2001. Simulating the Evolution of Language. London: Springer Verlag.

Christiansen, M. H., and Ellefson, M. H. 2002. Linguistic Adaptation without Linguistic Constraints: The Role of Sequential Learning in Language Evolution. In Wray, A. ed. The Transition to Language, 335–358. Oxford: Oxford University Press.

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

Hockett, C. F. 1960. The Origin of Speech. Scientific American 203: 88–96.

Holland, J. H. 1975. Adaptation in Natural and Artificial Systems. Michigan: University of Michigan Press.

Holland, J. H. 2001. Exploring the Evolution of Complexity in Signaling Networks. Complexity 7(2): 34–45.

Ke, J.; Minett, J. W.; Au, C. P.; and Wang, W. S-Y. 2002. Self-organization and selection in the emergence of vocabulary. Complexity 7(3):41–54.

Li, X. and Chen, G. R. 2003. A Local-World Evolving Network Model. Physica A 328(1-2):274–286.

Newman, M. E. J. 2003. The Structure and Function of Complex Networks. SIAM Review 45.

Munroe, S. and Cangelosi, A. 2002. Learning and the evolution of language: the role of culture variation and learning cost in the Baldwin Effect. Artificial Life 8(4):311–340.

Schoenemann, P. T. 1999. Syntax as an Emergent Characteristic of the Evolution of Semantic Complexity. Minds and Machines 9:309–346.

Shore, C. M. 1995. Individual Differences in Language Development. Thousand Oaks, CA: Sage Pub.

Smith, K.; Kirby, S.; and Brighton, S. 2003. Iterative learning: A framework for the emergence of language, Artificial Life 9:371–386.

Standish, R. K.; Bedau, M. A.; and Abbass, H. A. eds. 2003. Artificial life VIII: proceedings of the eighth International Conference on Artificial Life. Cambridge, MA: MIT Press.

Kirby, S. (2002). Learning, Bottlenecks and the Evolution of Recursive Syntax. In: Briscoe, T. ed. *Linguistic Evolution through Language Acquisition: Formal and Computational Models.* Cambridge, MA, Cambridge University Press. pp. 173–205.

Wagner, K.; Reggia, J. A.; Uriagereka, J.; and Wilkinson, G. S. 2003. Progress in the simulation of emergent communication and language. Adaptive Behavior 11(1):37–69.

Watts, D. J. 1999. Small Worlds. Princeton, NJ: Princeton University Press.

Wray, A. 2002. Formulaic Language and the Lexicon. New York: Cambridge University Press.