Language Emergence and Simple Social Structure

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

Language Engineering Laboratory, City University of Hong Kong, Hong Kong, PRC
*50005488@student.cityu.edu.hk

Abstract

Addressing the limitations of current computational models, in this paper, we present a multi-agent computational model to simulate the coevolution of lexicon and syntax (simple word order) during the transition from a holistic signaling system to a compositional language through iterative interactions within a heterogeneous population. An indirect meaning transference based on both linguistic and nonlinguistic information in communications, together with a feedback without direct meaning check, is implemented in communications. Based on this model, the influences of simple exogenous social structures, such as structure with popular agent and inter-group communication, on language emergence are studied. Besides, under the assumption of geographic limited communication, a phenomenon of "global polarization, local convergence" is simulated during language emergence. In the end, necessary linguistic and social structure related future directions are pointed out.

Key words: Language emergence, multi-agent model, coevolution, lexicon, syntax, indirect meaning transference, heterogeneity, social structure

1 Limitations of current computational models

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; Steels et al. 2002), and the iterative learning framework (e.g., Smith et al. 2003). Based on whether *agents* (language users) are situated in an artificial world and whether the communication acts use single or several unstructured tokens versus structured utterances composed of multiple tokens, most of current models can be divided into 4 types (according to Wagner et al. 2003):

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

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

All of these “emergent” models (according to Schoenemann 1999) view language evolution as a *Complex Adaptive System* (*CAS*) and share several assumptions which shed light on the real language development (e.g., interactions between agents and learning through generations drive the emergence of language; language-specific syntactic predispositions are unlikely, etc.). However, there are still several limitations in these models.

First, most of these models (excluding Cangelosi and Parisi 2002) assume *direct meaning transference* in interactions among agents: 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 these utterances express (Kirby 2002). Other channels, such as pointing while talking or feedback, can only provide certain degree of confirmation. Quine’s question (1960), regarding pointing, is a good counterexample: If someone points to a dog and says: “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 accurately interpreted 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 failed to model syntax (e.g., Ke et al. 2002; Steels et al, 2002), build in the syntactic features (e.g., Cangelosi and Parisi 2002), or else do not adopt a coevolutionary view of the emergence of syntax and lexicon (e.g. Smith et al. 2003). From an evolutionary point of view, the emergence of 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. Many scholars assumed that language emergence relies on or contributes to certain social factors (Romaine 1994; Dunbar 1998; Knight et al. 2000). On the one hand, *exogenous structures*, formed irrespective of language but dependent on biological and/or socio-economic factors, such as kinship and social classes, place constraints on possibilities and probabilities of interactions between agents. Different social structures may have different consequences in language acquisition and language change. On the other hand, *endogenous structures*, formed based on mutual understanding of language, may emerge along with the emergence of language. It is worth studying influences of both exogenous and endogenous structures on language emergence and vice versa.

Fourth, many models are built upon homogeneous agents, assuming each agent having identical characteristics and consistent linguistic abilities. However, sociolinguists have observed dramatic variations in speech communities (Romaine 1994), and studies on language acquisition have revealed various dichotomies in children’s learning styles (Shore 1995). Therefore, it is more realistic for computational models to take account of heterogeneity.

Addressing these limitations and based on the “emergent” scenario of Wray (2002), we present a "semi-situated, structured" computational model to study language emergence at macrohistory level (Wang 1991). It simulates an indirect meaning transference and shows a coevolution of lexicon and syntax (simple word order) during the transition from a holistic signaling system to a compositional language. A brief description of this model and coevolution process of lexicon and syntax are shown in Section 2. In Section 3, based on this model, we discuss two exogenous structures' influence on language emergence. A scenario, assuming a geographic limited communication, is implemented and the phenomenon of "global polarization, local convergence" (Axelrod 1997) during the emergence of language is discussed in Section 4. Finally, conclusions and some future directions are pointed out in Section 5.

2 Coevolution of lexicon and syntax (simple word order)

2.1 Description of the model

This model is basically a linguistic communication game, focusing on horizontal transmission among group of agents. Using linguistic utterances, agents express and interpret two types of integrated meanings, "predicate<agent>", such as "run<wolf>", and "predicate<agent, patient>", such as "eat<wolf, meat>" and "chase<wolf, dog>". Utterances, strings of syllables for encoding and decoding meanings, can be mapped to either entire integrated meanings or meaning constituents and utterances mapped to constituents can be combined under the regulation of syntax (simple word order) to map integrated meanings.

Each agent uses a rule-based system to indicate language. Linguistic rules include *lexical rules*, which are mappings between meanings and utterances, and *word order rules*, which are optional sequences to regulate utterances for expressing meanings with two or three constituents. Each rule has a condition part and strength, the latter numerically indicates the frequency of successful use of this rule and rule competition and adjustment are based on rule strength. 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)

0.4 is the strength of this rule. Compositional rules include both word and phrase rules. *Word rules* are mappings between a single constituent and utterance. For example:

"eat<#, #>"↔/d e/ (0.3) or "dog"↔/c/ (0.5)

where "#" can be replaced by constituents to form an integrated meaning. *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)

Considering domain general *sequencing ability* (Christiansen and Ellefson 2002) found in both human and some other animals and one language predispositions — *declarative memory* (Terrace 2002), agents are assumed to use sequencing ability to regulate linguistic utterances. Word order rules are like these:

"utterance for predicate precedes that for agent"(0.4), to simplify, SV(0.4) or

"utterance for agent first; that for predicate second; that for patient last"(0.2), SVO(0.2)

Agents start with a holistic signaling system (sharing only several common holistic rules) without dominant word order (all word order rules have same initialized strength (0.5)). Sharing a set of common linguistic rules indicate the emergence of common language among group of agents.

Each agent uses a two-level storage system to handle lexical rules (see Fig.1), inspired from a

Classifier System based model (Holland 2001). This system includes a buffer (storing "previous experiences" — meaning-utterance mappings (*M-U mappings*) obtained in previous communications) and a rule list (storing 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.

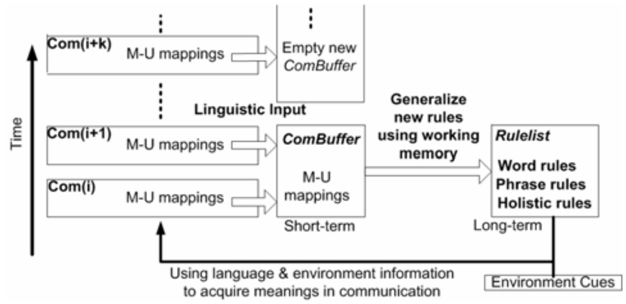


Fig. 1 Two-level storage system

Each agent has two linguistic behaviors to acquire new rules. One is *random creation* in meaning expression (similar to Kirby's model (1998)), under certain possibility, speakers can create lexical rules (holistic or compositional) to help their production of integrated meanings. The other is *rule generalization* through detecting *recurrent patterns* (recurrent meaning constituents in meanings and recurrent syllables in utterances among two *M-U mappings*) in rule update. Fig. 2 shows some examples of rule generalization. Through this flexible search of recurrent patterns without syntactic or location constrains, some holistic signals are decomposed into compositional rules. Synonymous and homonymous rules emerge inevitably during the execution of these two behaviors due to the lack of direct access to other agent's language and the flexible search of recurrent pattern. As for synonyms, agents randomly learn one form from a set of synonymous rules based on the *Principle of Contrast* (Clark 1987). As for homonyms, we will show that the effective transition to compositional language requires homonym avoidance in later Section.

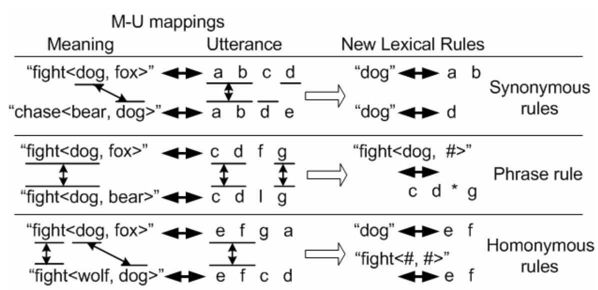


Fig. 2 Examples of rule generalization

This model allows that heterogeneity, i.e., each agent can have different storage capacity (S_{buffer} and $S_{rule\ list}$) and different linguistic ability in random creation and rule generalization ($R_{random\ creation}$ and $R_{rule\ generalization}$). In order to simplify, all agents' S_{buffer}, $S_{rule\ list}$, $R_{random\ creation}$, $R_{rule\ generalization}$ follow correlative Gaussian distributions.

Linguistic communication is only activity of agents. In this *concurrent* communication (on one time step, many communications between different pairs of agents happen simultaneously) system, an indirect meaning transference is implemented in communications. *Cues* are used to represent nonlinguistic information, which indicate events happening in environment. Cues are integrated meanings with same strength, e.g., "eat<wolf, meat>"(0.5), as assistance to listener's interpretation. Cues are unreliable, whether the speaker's intended meaning is contained in one of the cues is manipulated by *Reliability of Cues* (*RC*).

Communication in this model is summarized in Fig.3. 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

$$CS_{speak} = Str(\text{combinable activated rules}) + Str(\text{applicable order rules}) \tag{1}$$

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

$$CS_{listene} = LangWeight \begin{Bmatrix} Str(\text{combinable activated rules}) \\ + Str(\text{applicable order rules}) \end{Bmatrix} + Env\{Str(\text{related cues})\} \tag{2}$$

The listener interprets the meaning based on his winning rules. If the combined 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 under positive feedback and vice versa under negative feedback. During the whole process of communication, expression and interpretation are independent and the interpretation is based on the interaction of linguistic and nonlinguistic information. An example of communication with indirect meaning transference is shown in Appendix.

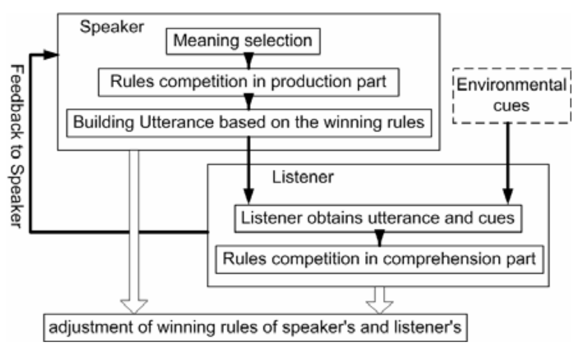


Fig.3 Communication with indirect meaning transference

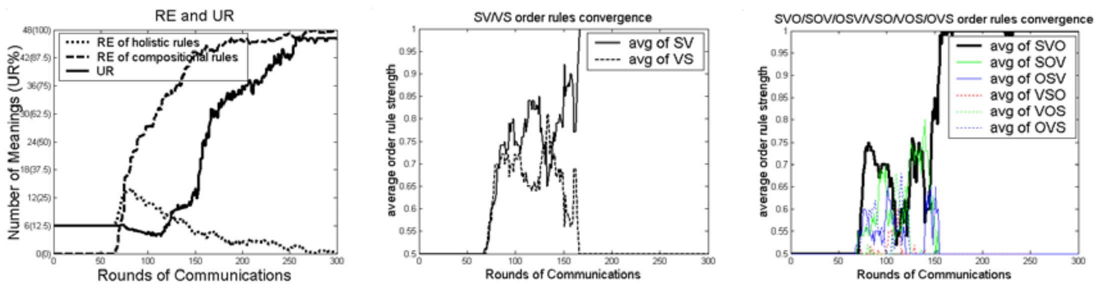
2.2 Coevolution of lexicon and syntax

Several parameters are used to test language emergence: 1) *Rule expressivity* (*RE*), the average number of meanings that all agents can express; 2) *Understanding rate* (*UR*), the average proportion of the number of meanings understandable to every pair of agents based on linguistic information only:

$$UR = \frac{\sum_{i,j} \text{number of understable meanings between agent i, j}}{\text{number of all possible pairs of i, j}} \tag{3}$$

UR evaluates the real representation ability (*Displacement* (Hockett 1960)) of the emergent language, not only the RE, but also whether such expressions are understandable using linguistic information only; 3) *Convergence time* (*CT*), number of iterations required on average to achieve a language where UR $\geq$ 80%.

Using these parameters, coevolution of lexicon and syntax is traced. Fig. 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 Fig. 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]. Fig. 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. Combine Fig. 4(a-c), a coevolution process is shown: mutual understanding requires not only common lexical rules but also a shared syntax to regulate utterances; The sharp increase of *UR* and strengths of the dominant order rules are almost *synchronized*: the use of compositional rules triggers the convergence of syntax, which in turn boosts the convergence of the lexicon.



(a) (b) (c)

Fig.4 Coevolution of lexicon and syntax (Condition: 10 agents, 500*5 communications, *RC*=0.8, S_{buffer}=35±5, $S_{rule\ list}$=45±5. $R_{random\ creation}$=0.5±0.2, $R_{rule\ generalization}$=0.5±0.2).

These results show that during language emergence, the transition from holistic signaling system to a compositional language with dominant word order might be a process of coevolution of lexicon and syntax. Through comparison, allowing certain heterogeneity does not significantly affect the emergence process.

On the other hand, this effective transition is found to be possible when adopting certain external and internal requirements. Externally, high reliable cues are required to effectively achieve a common language. Internally, strategies to avoid ambiguity are required. Homonym avoidance is one of such strategy that works well in this model. To test this, we trace the average *UR* under different *RC* with and without homonym avoidance in Fig. 5. Without homonym avoidance, *UR* is not very high even *RC* is 1.0; with homonym avoidance, a high *UR* is achieved under not high *RC*. Due to the limited size of storage, the lack of context (meanings expressed in communications are independent of each other) and the unreliable cues (otherwise, it would still be *direct meaning transference*), homonym avoidance is necessary in this model. However, many researches in real language acquisition and interpretation (e.g., Li et al. 2001) show that homonym avoidance does play an important role when human brain processes languages.

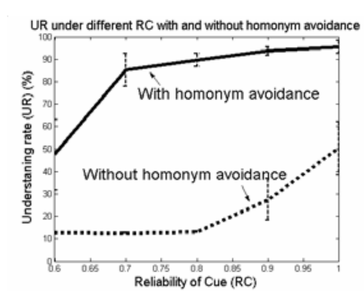
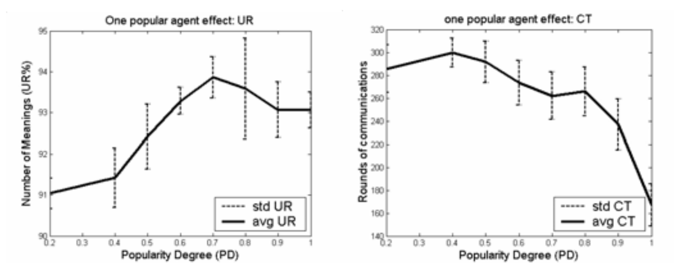


Fig. 5 *UR* under different *RC* with and without homonym avoidance (Condition: 10 agents, 500*5 communications, *RC*=0.8, S_{buffer}=35$\pm$5, $S_{rule\ list}$=45$\pm$5. $R_{random\ creation}$=0.5$\pm$0.2, $R_{rule\ generalization}$=0.5$\pm$0.2, 10 simulations).

3 Exogenous social structure's influence on language emergence

3.1 Popular agent's effect

One of the possible exogenous social structures in early human development might be the social structure with popular agent(s) (leader(s)) (Labov 2001). *Popular agent* is the agent having more chances to participate in communications than other *unpopular agents*. In this social structure, we focus on communications between popular and unpopular agent and communications between two unpopular agents. *Popularity degree* (*PD*), the probability for a random selected normal agent would communicate with the popular agent, is used to indicate the popularity of the popular agent. With the increase of *PD*, more and more communications involve the popular agent and unpopular agents gradually surround the popular agent. From this point of view, the *PD* indicates the level of global centralization. Fig. 6 shows the *CT* and *UR* under different *PD*. High *PD* reduces *CT*. Besides, there seems to be an optimal *PD* (0.7) where the highest *UR* is achieved.



(a) (b)

Fig. 6 Popular agent's effect: (a) *UR*; (b) *CT* (Condition: 10 agents, 500*5 communications, *RC*=0.8, S_{buffer}=35$\pm$5, $S_{rule\ list}$=45$\pm$5. $R_{random\ creation}$=0.5$\pm$0.2, $R_{rule\ generalization}$=0.5$\pm$0.2, 10 simulations).

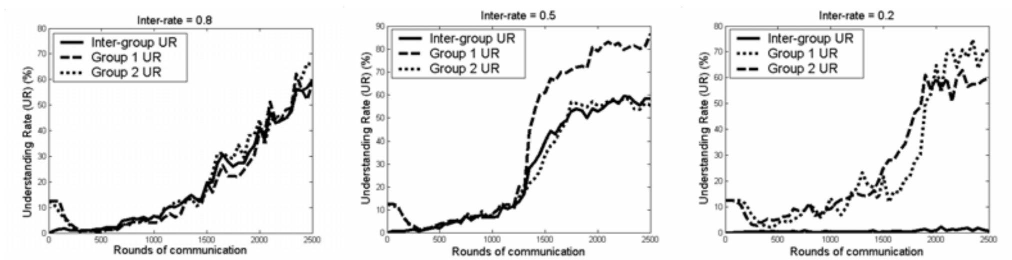
Global centralization around some agent(s) has two effects. One is *Acceleration effect*: popular agent(s) connects many unpopular ones, like a network *hub*. Centralization around it can increase the chances for unpopular agents to exchange information, and then accelerate the convergence of common rules, thus reducing the *CT*. The other is *Deceleration effect*: effective information transference between two agents (say, Agent1 and Agent2) requires a direct connection or a connection through a *stable intermediary* (i.e., if the popular agent's internal rules do not change much, the information received by Agent2 via the popular agent will not change much from the original information sent by Agent1). However, with the increase of global centralization, other unpopular 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. So, it greatly affects the information transference and the convergence of common rules between Agent1 and Agent2 via the popular agent. Balancing these two contradictory factors, the optimum UR happens at an intermediate level of global centralization; absolute "democracy" (fully-connected network) and absolute "dictatorship" (star network) will not achieve the best performance.

3.2 Inter/Intra-group communication

Another exogenous social structure relates to communications between two groups. Two types of communications are considered: communication among agents insider one group (*Intra-rate* indicates the frequency of this type of communication) and communication among agents in different group (*Inter-rate* indicates the frequency of this type of communication) (*Intra-rate* + *Inter-rate* = *1.0*). *Intra-rate* and *Inter-rate* indicate the closeness inside and between the groups. *Cross-group Understanding Rate* ($UR_{cross\text{-}group}$) is introduced to test the language similarity among agents in different groups.

$$UR_{cross-group} = \frac{\sum_{i,j} \text{number of understable meanings between agent i, j}}{\text{number of all possible pairs of agent i (in group1) and agent j (in group2)}} \quad (4)$$

Fig. 7 shows the $UR_{cross\text{-}group}$ and *UR* within two groups under different *Inter-rate* in one simulation.



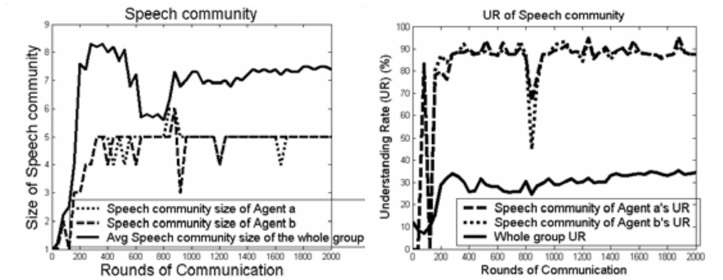
(a) (b) (c)

Fig. 7 Two group communication: (a) *Inter-rate*=0.8; (b) *Inter-rate*=0.5; (c) *Inter-rate*=0.2 (Condition: 2 groups, 10 agents in each group, 500*5 communications, *RC*=0.8, S_{buffer}=35$\pm$5, $S_{rule\ list}$=45$\pm$5. $R_{random\ creation}$=0.5$\pm$0.2, $R_{rule\ generalization}$=0.5$\pm$0.2, 10 simulations).

The frequency of inter-group communication decides the emergence process and the similarity of the language between the two groups. The higher of the *Inter-rate*, the higher the similarity of the languages emerged in two groups. When the *Inter-rate* is low, two groups can develop different lexical rules, let alone dominant orders. In real history, communications between two groups usually happened between their leaders or few representatives, which provided a chance of language divergence. On the other hand, on the boundary of the group, communications between agents in two groups were inevitable, which provided a chance of language convergence or introducing innovations into each other's group, then drove language change. This matches Milroy's hypothesis (1980).

4 Global polarization and local convergence during language emergence

In this part, a preliminary study of the emergence of endogenous social structure based on mutual understanding of evolving language during language emergence is presented. Agents are put into a 2D world and moving randomly, which simulates the early hominid clans distributed around the globe. Communications are assumed to happen within a geographic area. *Successful communication* (most of the speaker's intended meanings are accurately interpreted by the listener) can link the speaker and the listener as *Speech Community* and they tend to move together in future. However, failed communication will break this connection. Fig. 8(a) traces speech community sizes of two randomly selected agents and the average speech community size of the whole group in one simulation, Fig. 8(b) traces *UR* of these two speech communities and average *UR* of the whole group.



(a) (b)

Fig. 8 Speech community sizes and UR inside speech community (Condition: 180*80 world, 20 agents, 2000*20 com., RC=0.8, S_{buffer}=35±5, $S_{rule\ list}$=45±5. $R_{random\ creation}$=0.5±0.2, $R_{rule\ generalization}$=0.5±0.2, 10 simulations).

During language emergence, agents in the whole group are gradually forming up different speech communities and even reach a stable state when the sizes of speech communities are stable. Mutual understanding is the force for community cohesion: inside speech communities, *UR* is very high; however, *UR* of the whole group is very low. Low *UR* of the whole group and high UR in different speech community indicate a set of efficient languages mutually understood by members of same speech community emerged in different speech community, but these languages are hardly similar to one another. Such linguistic feature of "global polarization, local convergence" is formed up during the emergence of language, which extends Axelrod's (1997) conclusions.

5 Conclusions and linguistic and social structure related future directions

In this paper, addressing current computational models' limitations, we present a multi-agent model to study language emergence and show a coevolution of lexicon and syntax during the transition from holistic signaling system to compositional language with dominant word order. This model implements an indirect meaning transference and considers certain heterogeneity, which is more realistic. Meanwhile, based on this model, we concentrate on social structure's influence on language emergence, which is always disregarded by many models on language emergence, and discuss certain exogenous social structures' effects and also make a preliminary research on the emergence of endogenous social structures based on mutual understanding of the evolving language. However, this model still requires further modifications.

In linguistic part, language emergence is not achieved in one generation's communication, besides horizontal transmission, vertical transmission (a focus of Smith et al.'s model (2003)) also plays important roles and should be considered. In social structure's part, recently, the developments on complex networks (Wang 2002; Newman 2003) provide us an effective methodology, a theoretic guideline and useful tools to incorporate various social structures into the models of language evolution. More realistic social structures can be implemented and in-depth study of coevolution of language emergence and corresponding endogenous social structure is possible.

Acknowledgement

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.

Reference

Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. *The Journal of Conflict Resolution,* 41(2): 203–226.

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.

Dunbar, R. I. M. 1998. Grooming, Gossip, and the Evolution of Language. Cambridge, Mass.: Harvard University Press.

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

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.

Kirby, S. (2002). Natural Language from Artificial Life. *Artificial Life* 8(2): 185–215.

Knight, C.; Studdert-Kennedy, M; and Hurford, J. R. eds. 2000. The Evolutionary Emergence of Language: Social function and the origins of linguistic form. Oxford: Oxford University Press.

Li, P.; Sha, H.; Yip, M.; Zhang, Y.; and Tang, Y. 2001. Lexical Ambiguity in Sentence Processing: Evidence from Chinese. In Nakayama, M. ed., *Sentence Processing in East Asian Languages*, 111–129. CSLI Publications.

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

Milroy, Lesley. *Language and Social Networks*. Oxford: B. Blackwell, 1980/1987

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.

Quine, W. V. 1960. *Word and Object*. Cambridge, MA, MIT Press.

Romaine, S. 1994. *Language in society, an introduction to sociolinguistics.* Oxford, UK, Oxford University Press.

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.

Steels, L.; Kaplan, K.; McIntyre, A.; and Looveren, J. V. Crucial Factors in the Origins of Word-Meaning. In: Wray, A. ed., *The Transition to Language*, 252–271. Oxford: Oxford University Press.

Terrace, H. S. 2002. Serial Expertise and the Evolution of Language. In: Wray, A. ed., *The Transition to Language*, 64–90. Oxford: Oxford University Press.

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.

Wang W. S-Y. 1991. The Three Scales of Diachrony. In: Wang, W. S-Y. ed. *Explorations in Language*, 60–71. Taiwan: Pyramid Press.

Wang, X. F. 2002. Complex networks: topology, dynamics and synchronization. *International Journal of Bifurcation and Chaos* 12(5): 885–916.

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

Appendix: Example of communication with indirect meaning transference

Suppose a speaker, who wants to express the integrated meaning “fight<dog, fox>”, according to his own rule list, has three ways to express it as shown in Fig. A.1: 1) using a holistic rule, no word order rule is considered; 2) using three word rules, six possible word order rules are applicable, so the strongest one (*VSO*(0.6)) is chosen; 3) using a word rule and a phrase rule, the utterance of the phrase rule restricts the applicable word order to only *VSO*(0.6) and *OSV*(0.5), so the strongest one *VSO*(0.6) is chosen. In each condition, CS_{speak} is calculated. The combination of lexical and word order rules (bold text in Figure A.1) having the strongest CS_{speak} (*CS3*) are chosen as the speaker’s winning rules. The utterance “e b f” is built up accordingly, and sent to the listener.

Production Part — Meaning to express: *" fight<dog, fox>"*

	Activated rules	Applicable word order rules	Combined Strength (CS)
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)	*VSO* (0.6)	*CS2 = 1/2(1/3(0.8+0.5+0.2)+0.6) = 0.55*
1 word rule 1 phrase rule	***“dog”←→/b/* (0.8)** ***“fight<#, fox>”←→/e * f/* (0.8)**	***VSO* (0.6)** *OSV* (0.5)	***CS3 = 1/2(1/2(0.8+0.8)+0.6) = 0.7***

Utterance built up: */e b f/*

Fig. A.1: Example of indirect meaning transference: Production

Comprehension Part — Utterance heard: */ e b f/*

	Activated rules	Related cues	Detectable word order rules	Combined Strength (CS)
1 holistic rule	*“eat<dog, meat>”←→/e b f/* (0.4)	*“eat<dog, meat>”* (0.5)		*CS1 = 1/2(0.5)+1/2(0.5) = 0.45*
3 word rules	***“cat”←→/e/* (0.7)** ***“fight<#, #>”←→/b/* (0.8)** ***“dog”←→/f/* (0.6)**		***SVO* (0.6)**	***CS2 = 1/2(1/3(0.7+0.8+0.6)+0.6)+1/2(0) = 0.65***
1 word rules	*“run<#, #>”←→/b f/* (0.7)	*“run<cat>”* (0.5)		*CS3 = 1/2(0.3)+1/2(0.5) = 0.4*

Interpreted Meaning: *" fight<cat, dog>"*

Fig. A.2: Example of indirect meaning transference: Comprehension

Suppose the listener, in comprehension (Fig. A.2), hears the speaker’s utterance /e b f/ sometime selects some cues (Cue1: “eat<dog, meat>” (0.5) and Cue2: “run<cat>” (0.5), neither of which contains the speaker’s intended meaning “fight<dog, fox>”). Then, in the listener’s rule list, lexical rules whose utterance parts partially or fully match the heard utterance are activated. He has three ways to decode the utterance and CS_{listen} is calculated in each way (both *LangWeight* and *CueWeight* in (2) are 0.5): 1) using one holistic rule, since Cue1’s meaning is same to that of the holistic rule and no applicable word order can be detected, both the strength of holistic rule and that of Cue1 are used to calculate CS_{listen}; 2) using three word rules, no cue is related, but the word order *SVO* can be detected by using these word rules, so the strengths of the three word rules and that of *SVO* are used to calculate CS_{listen}; 3) using one word rule, and since Cue2’s meaning is related to it but no clear order rule can be detected, both the strength of word rule and that of Cue2 are used to calculate CS_{listen}. The combination of lexical and word order rules (bold text in Figure B.1) having the strongest CS_{listen} (*CS2*) are chosen as the listener’s winning rules. The interpreted meaning is built up accordingly. Since *CS2* exceeds the confidence threshold (0.5), a positive feedback is sent back to the speaker. Then, both the speaker and the listener increase the strengths of their winning rules, and decrease the strengths of homophonic rules with meaning parts in same category in the speaker’s and the listener's rule lists.

In this example, the interpreted meaning is misinterpreted by the listener, but the primitive feedback still makes the speaker believe that the listener understands what he says. After certain times of iterative communications, such misunderstanding can be gradually avoided and a common language can still be achieved among agents.