Connection Science
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713411269

Exploring social structure effect on language evolution based on a computational model

Tao Gong [a]; James W. Minett [a]; William S. -Y. Wang [a]
[a] Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, People's Republic of China

Online Publication Date: 01 June 2008

To cite this Article: Gong, Tao, Minett, James W. and Wang, William S. -Y. (2008) 'Exploring social structure effect on language evolution based on a computational model', Connection Science, 20:2, 135 — 153

To link to this article: DOI: 10.1080/09540090802091941
URL: http://dx.doi.org/10.1080/09540090802091941

Connection Science
Vol. 20, Nos. 2–3, June–September 2008, 135–153

Taylor & Francis
Taylor & Francis Group

Exploring social structure effect on language evolution based on a computational model

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

Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, People's Republic of China

A compositionality-regularity coevolution model is adopted to explore the effect of social structure on language emergence and maintenance. Based on this model, we explore language evolution in three experiments, and discuss the role of a popular agent in language evolution, the relationship between mutual understanding and social hierarchy, and the effect of inter-community communications and that of simple linguistic features on convergence of communal languages in two communities. This work embodies several important interactions during social learning, and introduces a new approach that manipulates individuals' probabilities to participate in social interactions to study the effect of social structure. We hope it will stimulate further theoretical and empirical explorations on language evolution in a social environment.

Keywords: computational simulation; language evolution; social structure; power-law distribution

1. Introduction

Social structures are found in human societies, primate communities, and colonies of some other species. Almost a century of study on chimpanzees and other primates (e.g. de Waal 2005) has provided ample empirical data on their social structures. From an evolutionary perspective, these social structures might resemble those of the early humans, at least as precursors (Whiten 2005), and they, existing prior to human language, could have played certain roles in language evolution. Some scholars (e.g. Dunbar 1996) have argued that it was the social structuring and related social activities (e.g. coalition and competition) that created the conditions for the emergence of human language. Many sociolinguists have taken into account the influence of social factors on language evolution, especially on language change. For instance, Labov has argued that individual linguistic behaviors could be stratified by various social factors, such as age, education level, social class, and so on. Apart from language-internal factors (Labov 1994), language-external factors, such as social networks, identity, and gender (Labov 2001) could determine linguistic variations. Thomason and Kaufman (1988) have further argued that instead of the structural linguistic relations, the social facts of particular contact situations mainly determine the contact-induced language change.

In addition to empirical studies, modelling social systems as complex systems also offers insights into the interdisciplinary social science research regarding self-organisation, emergence,

*Corresponding author. Email: gtojty@gmail.com

ISSN 0954-0091 print/ISSN 1360-0494 online

DOI: 10.1080/09540090802091941
http://www.informaworld.com

and hierarchy (Alessa et al. 2006). This modelling approach has recently become widely adopted in explaining both historical linguistic changes (reviewed by Bhattacharjee 2003) and sociological phenomena (reviewed by Malsch and Schulz-Schaeffer 2007). For instance, Livingstone (2001) used a computational model to study dialectal diversity. In his model, all agents were arranged in a single row, the ends of which were disconnected. Communications among agents were limited by predefined neighbourhoods based on distance. After a number of communications, a '*dialect continuum*' (minor changes existed within neighbourhoods, allowing successful communications therein, and major shifts and differences existed across neighbourhoods, allowing reinforcement of group identities) emerged in this system. The author then claimed that the limitation induced by neighbourhood or social distance could give rise to emergent dialects. In addition, Nettle (1999), based on some computational models derived from social impact theory (Nowak 1990), explored *the threshold problem* (how an initially rare innovation can win over a strong linguistic norm) in language change. In his models, social structure was modelled as a *weighted, regular network* (a network whose nodes have an equal number of weighted edges connecting to other nodes). In this network, any innovation at a node could affect its neighbouring (connected) nodes. This effect decreased exponentially with the increase in the distance between this node and its neighbours. These models demonstrated that successful innovations usually originate from the speakers having higher influence ('social impact') than others, as others favour learning from these influential ones. Furthermore, Ke (2004) extended Nettle's work by introducing some popular networks found in biological and social phenomena, such as the small-world (Watt and Strogatz 1998) and scale-free (Barabasi and Albert 1999) networks. She explored the diffusion of linguistic innovations in these networks, and found that the innovations that occurred in the idiolects of influential speakers (nodes having more edges) could easily diffuse to the entire population.

Two aspects of social factors were explored in these previous studies that illustrated the influence of social structures on linguistic phenomena. In Livingstone's and Nettle's studies, the social distance was the main factor to affect communications; the longer the social distance between the speaker and listener, the less the social impact they had on each other during communications. In Ke's models, the social connection was the main factor to affect communications; the more edges an agent had, the more influential its idiolect became. In addition, these studies put agents in some specific structures that were initialised before the simulation and remained stable throughout it. In Livingstone's model, agents were arranged in a row, whereas in Nettle's and Ke's studies, they were located in some regular or complex networks. Furthermore, all these studies mainly discussed the effect of social structure on lexical evolution. For instance, in Livingstone's and Ke's models, language was treated as a set of lexical items. Due to different adopted social structures, agents could either develop new lexical items or diffuse some salient ones to the whole population.

In this paper, instead of particular social structures, we concentrate on the probabilities that individuals participate in communications and discuss their effect on language evolution. Three kinds of probability, each reflecting the influence of a variety of simple social structures, are defined as follows:

(1) The probability that a particular agent (*the popular agent*) participates in communications, which is denoted by *PopRate* (PR). The higher the PR, the higher the probability that the popular agent speaks or listens to others. Similar to the α-male(s) in primate communities, an early hominid group might have contained some agent(s) who could get involved in many activities. A chief that still exists in some hunter-gatherer societies (Barnard 2003) could resemble such a popular agent. PR is subject to many factors, such as an agent's physical abilities, economic condition, and so on.

(2) The probability for every individual in the community to participate in communications, which is denoted by *Individual's Popularity*. In a large community, agents with various social

statuses could have different probabilities to communicate with each other. Many social factors, such as friendship, political influence and economic incomes, could affect an individual's popularity. The distribution of all individuals' popularities could reflect the collective effect of the social structure in that community. The difference between PR and *Individual's Popularity* is that PR only concerns the single popular agent, and treats others equally, whereas *Individual's Popularity* concerns each individual in the community, and all these popularities could follow some predefined distributions.

(3) The probability that individuals of a community choose to communicate with members of their own community, rather than members of a different community. When two communities interact with each other, there are two types of communications: *intra-community communications* among members of the same community and *inter-community communications* among members of different communities, the probabilities of which are, respectively, controlled by *IntraRate* and *InterRate*. These two probabilities sum to 1.0. They reflect the influence of geographical, economic, and political factors. For instance, enlarging the social distance may decrease *InterRate*, and enhancing the economic bond may increase *IntraRate*.

We design three experiments to explore the effect of social structure on language evolution, each manipulating one of the above probabilities. In Experiment 1, a community with a single popular agent is simulated, and this agent's popularity is regulated by PR. In Experiment 2, a community with a predefined distribution of individuals' popularities is simulated. In Experiment 3, a situation where agents from two communities interact with each other is simulated. Both Experiment 1 and Experiment 2 explore the intra-community effect of social structure, which mainly results from the non-uniformly distributed communications in the community. Experiment 3 explores the inter-community effect of social structure, which mainly results from the degree of linguistic contact and the similarities in some linguistic features.

All these experiments are based on a computational model (Gong 2008) that was originally designed to study the phylogenetic emergence of language in a population of language users. In this model, equipped with some domain-general abilities, such as sequencing and pattern detection abilities, a population of interacting agents can gradually develop a common set of lexical items and word orders through iterated communications. The model traces a coevolution of two linguistic universals, compositionality (in the form of lexical items) and regularity (in the form of simple word order), during the transition from an initial holistic signalling system to a compositional language. Compared with previous studies that mainly adopt models focusing on lexical evolution, this model gives us an appropriate level of complexity to observe the effect of social structures on both lexical and syntactic evolutions. In addition, besides the emergence of linguistic universals, this model also helps to study the effect of social structure on the maintenance of linguistic universals, which is more relevant to present-day societies. Also, a better understanding of the social structure effect can be revealed from a comparison of both the emergence and maintenance situations.

The remainder of the paper is organised as follows: Section 2 briefly describes the compositionality-regularity coevolution model; Section 3 discusses the simulation results in the three experiments; and finally, Section 4 discusses the results, gives the conclusions, and points out some future work.

2. The compositionality-regularity coevolution model

Figure 1 shows the conceptual framework of this model. Its detailed description can be found in the study of Gong (2008). The following sections only briefly describe some of its major components.

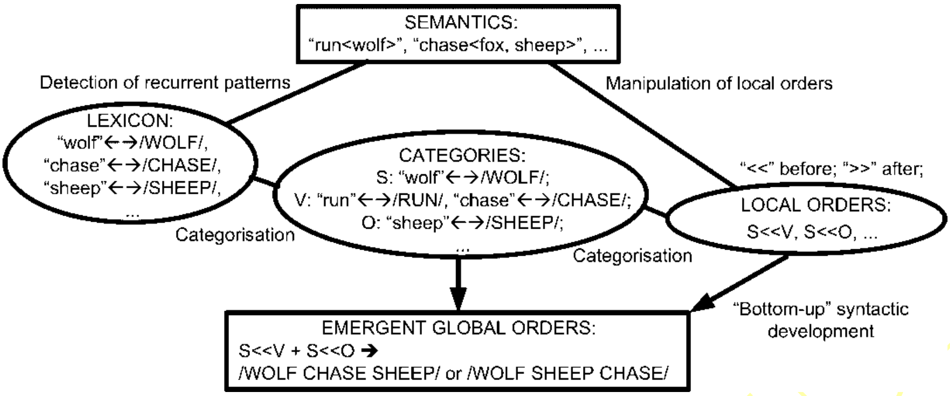

Figure 1. The conceptual framework of the compositionality-regularity coevolution model. The SEMANTICS rectangle represents the predefined semantic space, and the ovals represent the three aspects of linguistic knowledge acquired by agents based on different domain-general abilities including pattern extraction, sequential learning, and categorisation. The EMERGENT GLOBAL ORDERS rectangle represents the emergent syntactic patterns triggered by these learning abilities and correlated linguistic knowledge (indicated by arrows). Letters within ‘ ’ are semantic items, those within // are utterance syllables. ‘S’, ‘V’, and ‘O’ represent the syntactic roles of categories.

2.1. *The representation and acquisition of linguistic rules*

Language in this model is treated as a set of *meaning-utterance mappings* (*M-U mappings*). Individuals exchange two types of integrated meanings in communications: Type1: ‘Pr1<Ag>’, such as ‘run<wolf>’ meaning ‘a wolf is running’; and Type2: ‘Pr2<Ag, Pat>’, such as ‘chase<wolf, sheep>’ meaning ‘a wolf is chasing a sheep’. Here, Ag, Pr1/2 and Pat represent semantic roles, in which Ag denotes the actor of an action (agent); Pr1/2, the action (predicate); and Pat, the entity that undergoes an action (patient). These integrated meanings are mapped to utterances (strings of syllables). In the simulations of this paper, there are 12 semantic items (four can be Ag, four can be Pr1, and four can be Pr2). Also, the items that can be Ag in some integrated meanings could also be Pat in other integrated meanings. Therefore, all these items form a semantic space containing 16 (4×4) Type1 and 48 (4×4×(4-1)) Type2 meanings.

A rule-based system is adopted to represent individuals’ linguistic knowledge (Figure 2). Three types of linguistic rules are defined. *Lexical rules* record mappings between integrated meanings and strings of syllables (these are termed as *holistic rules*, see rules (a) and (b) in Figure 2) or between semantic items and strings of syllables (these are termed as *compositional rules*, see rules (c) and (d) in Figure 2). A lexical rule consists of a mapping and a strength, the latter of which numerically indicates the probability (from 0.0 to 1.0) of successfully using that mapping.

Syntactic rules record local orders (e.g. before or after, but not necessarily immediately before or after) between the strings of two sets (syntactic categories) of compositional rules in utterances (see rules (I), (II), and (III) in Figure 2). A syntactic rule contains a local order and a strength, the latter of which indicates the probability of successfully applying this local order on two lexical rules.

Syntactic categories associate a set of lexical rules whose semantic expressions have the same semantic roles (Pr1/2, Ag, or Pat) in some integrated meanings, and their utterances are *similarly used* (have an identical local order with respect to the utterances of other lexical rules) in some sentences. The identical local order is also associated with the same category as a syntactic rule. A syntactic category contains a syntactic role (for convenience, syntactic categories are labelled with the syntactic roles to which they correspond in simple declarative sentences in English, i.e. S, Subject; V, Verb; and O, Object), a list of lexical rules encoding items with the same semantic role in integrated meanings, and a list of syntactic rules encoding local orders between lexical rules of this category and those of others, or between lexical rules of this category and some

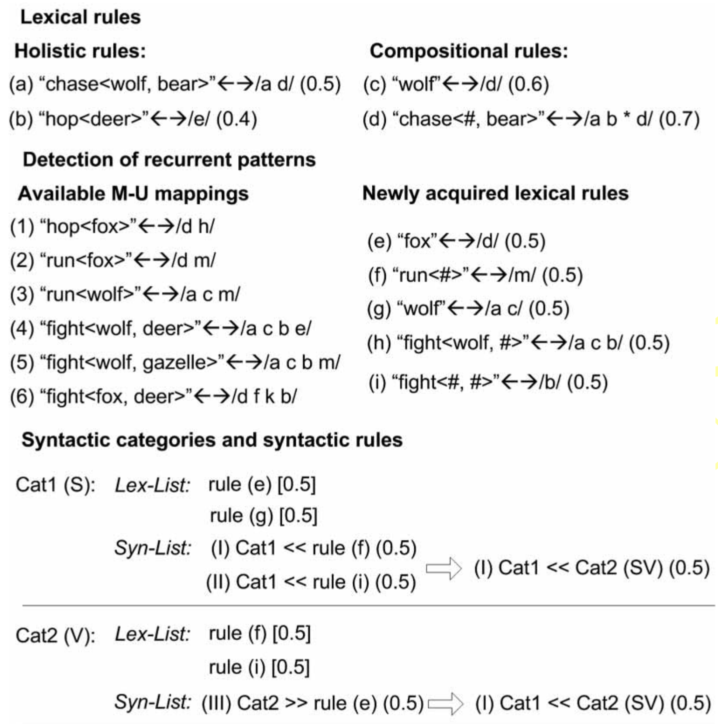

Figure 2. The representation and acquisition of linguistic knowledge: ‘#’ can be replaced by other semantic items, and ‘*’ by other syllables. Lexical rules are itemised by letters, M-U mappings by Arabic numerals, and syntactic rules by Roman numerals. Numbers enclosed by ( ) denote rule strengths, and those by [ ] denote association weights. ‘<<’ indicates the local order ‘before’.

other lexical rule(s). *An association weight* is defined to numerically indicate the probability for a lexical rule to follow the syntactic rules of a category. A lexical rule can be associated with many categories having identical syntactic roles but with different association weights. Lexical rules encoding items being Ag or Pat can be associated with both S and O categories.

Lexical rules are acquired through the detection of recurrent patterns. Each agent has a buffer storing some previous experience (a finite list of M-U mappings obtained in its previous communications with others). Newly acquired M-U mappings are compared with those stored in the buffer before they too are inserted into the buffer. *A recurrent pattern* is defined as one or more semantic item(s) and one or more syllables that appear recurrently in at least two M-U mappings in the buffer. For instance, in Figure 2, by comparing M-U mapping (2) with (1), the recurrent pattern ‘fox’ ←→ /d/ is detected, and so acquired as a lexical rule, whose initial rule strength is set to 0.5.

During the acquisition of lexical rules, syntactic rules and categories are also acquired. Evident in the previous experience (M-U mappings (2) and (3) in Figure 2), the syllables /d/ of rule (e) and /ac/ of rule (g) precede the syllable /m/ of rule (f). Since ‘wolf’ and ‘fox’ share the same semantic role (Ag) in these integrated meanings, rules (e) and (g) are associated into a new category, labelled S (Cat1). The association weights are all set initially to 0.5. Meanwhile, the local order (before) with respect to rule (f) is acquired as a syntactic rule (I) in this category. It indicates that the syllables of lexical rules from this category should precede the syllable of rule (f). Similarly, checking M-U mappings (5) and (6), another syntactic rule (II) with respect to rule (i) is acquired. Furthermore, checking M-U mappings (2) and (6), the syllables /m/ of rule (f) and /b/ of rule (i) are found to follow the syllable /d/ of rule (e). Consequently, a new V category

(Cat2) associating rules (f) and (i), which share the same semantic role 'Pr1/2', is created together with a new syntactic rule (III). Now, since rules (f) and (i) are already associated into a category, all these syntactic rules are updated as one rule 'Cat1≪Cat2 (SV) (0.5)' in both categories. It indicates that the syllables of lexical rules from the S category should precede those of lexical rules from the V category. In addition to the creation of categories, if rules (f) and (i) already belonged to different V categories, this previous experience would trigger a merging of these V categories into one category comprising their lexical and syntactic members. Without directly acquiring the global orders in sentences that encode Type2 meanings, agents can use their local orders to regulate the syllables of compositional rules in pairs to build up these sentences. For example, to express 'fight<fox, sheep>' based on the lexical rules (i), (e), and another lexical rule expressing 'sheep', the SV local order in syntactic rule (I) can be used for regulating the syllables of rules (i) and (e). Meanwhile, another local order, say SO, from another syntactic rule is used for regulating rule (e) and the rule expressing 'sheep'. Then, the global order based on these local orders can be either SVO or SOV. A similar example is shown in Figure 1, in which under the regulation of two local orders SV and SO, the utterance encoding 'chase<wolf, sheep>' could be either /WOLF CHASE SHEEP/ or /WOLF SHEEP CHASE/.

The formation of global orders based on local information introduces a certain degree of imprecision: the combination of some local orders can lead to multiple global orders (e.g. SV plus SO lead to either SVO or SOV, as shown above), and a particular global order can be represented by the combination of different local orders (e.g. SVO can be represented by SV plus VO or SO plus VO). This imprecision increases both the difficulty for a population of agents to acquire a common global word order and the probability of word order change (Minett 2006).

The rule strengths and association weights make possible the competition of linguistic knowledge. Individual's linguistic knowledge also suffers a forgetting process; after a communication, agents subtract a small amount (0.05) from the strengths of their linguistic rules and the association weights of their lexical rules. Lexical rules having negative strengths or association weights after the subtraction are removed from the individual's rule list or the syntactic categories containing these rules. Syntactic rules having negative strengths are also removed from the related categories. The rule competition and forgetting strengthen and maintain the frequently used linguistic knowledge and cause language to self-organise.

2.2. *The communication scenario*

This model simulates dyadic communication. Besides linguistic information, environmental cues as one type of non-linguistic information can assist the comprehension of heard utterances. An environmental cue is simulated as an integrated meaning with a fixed strength (0.75). Cues are not always reliable; otherwise, the learning procedure would involve mind-reading. The probability that one cue corresponds to the speaker's intended meaning is represented by *Reliability of Cues* (RC). In this model, RC is set to an intermediate value, 0.6. Introducing environmental cues and manipulating their reliability make this model semi-situated and provide the physical grounding of social interactions.

A dyadic communication between two randomly chosen agents (one is speaker and the other is listener) contains multiple (20) rounds of utterance exchange. During an utterance exchange, first, the speaker chooses an integrated meaning from the semantic space to express. Then, it activates some lexical or syntactic rules and related categories with which to encode this integrated meaning. Through a strength-based competition, it identifies the winning rules, builds up the utterance accordingly, and transmits the utterance to the listener. If lacking a set of rules to represent all the semantic items contained in the chosen meaning, the speaker may occasionally (the random creation rate, 0.25) create a holistic rule to express the whole meaning. After the speaker produces the utterance the listener receives the utterance and one cue from the environment. Then, it activates both lexical rules whose syllables fully or partially match the heard utterance and related categories. It selects the set of rules that allows it to comprehend an integrated meaning with the highest combined strength. The calculation of the combined strength considers the strengths of both linguistic rules and available cues. If the combined strength of the winning rules exceeds a *Confidence Threshold* (CT, set to 1.5 in the experiments reported here), the listener transmits a positive feedback to the speaker, and both agents reward their winning rules by increasing their strengths. Otherwise, a negative feedback is sent, and these rules are penalised by decreasing their strengths. The amount by which the rule strengths are increased or decreased is 0.1.

Throughout the utterance exchange, there is no direct check whether the speaker's intended meaning matches the listener's comprehended one. The listener's comprehension considers both linguistic and non-linguistic information. This provides the opportunity for developing reliable linguistic knowledge to withstand the interference from cues that do not match the speaker's intended meaning. This communication scenario can trigger a reliable language capable of describing events not happening in the immediate space or time.

All three types of linguistic rules participate in utterance exchange. For example, as shown in Figure 3, during production, the speaker first activates the lexical rules that can be combined to encode the chosen meaning. Then, based on the syntactic categories of these rules, it activates the syntactic rules (OS and VO) by which these lexical rules can be regulated. Then, it judges which set of linguistic rules can win the strength-based competition, and produces an utterance (/abcdef/) accordingly. Similarly, during comprehension, after the listener identifies the lexical rules whose syllables partially match the heard utterance, the local orders (VS and OS) that are consistent with the locations of the syllables in the heard utterance (/ab/ before /ef/ and /d/ before /ef/) are detected. If these local orders match the syntactic rules of the categories to which these lexical rules belong, both the categories and their syntactic rules are activated. Then, based on the activated categories, the semantic roles of these lexical items in the comprehended meaning are specified ('lion' is Ag, 'fox' is Pat, and 'fight<#, #>' is Pr2). The calculation of the combined strength also considers the strength of the cue that matches this comprehended meaning ('fight<lion, fox>'). The listener then judges which set of linguistic rules can win the strength-based competition and determines the feedback accordingly.

In this example, the speaker's intended meaning 'chase<lion, wolf>' encoded in /abcdef/ is misinterpreted by the listener as 'fight<lion, fox>'. Nevertheless, it is shown that categories play an important role in communication. In production, they link semantic structures with syntactic structures, and based on local orders in their syntactic members, global orders are formed. In comprehension, they transcribe syntactic structures into semantic structures, and based on the local

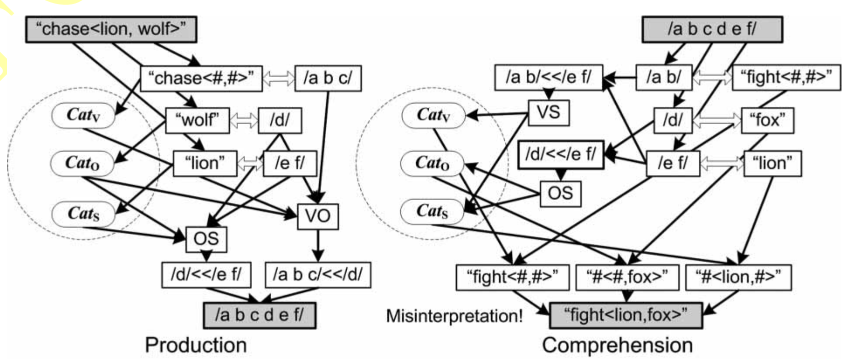

Figure 3. The interaction of linguistic rules in production and comprehension: Cat_S, Cat_V, Cat_O represent syntactic categories with syntactic roles of S, V, and O, and '<<' represents the local order 'before'.

orders in their syntactic members, semantic roles of lexical rules are specified. The whole process embodies how the conceptual-symbolic system (lexical items) and the regulatory system (syntactic categories and local orders) work together closely to process complex linguistic expressions.

The coevolution of compositionality and regularity takes place in iterated communications. During iterated communications, through the above learning mechanisms, individuals first learn some lexical rules and induce some independent categories to associate these lexical rules. Then, based on the M-U mappings stored in their buffers, they acquire more lexical rules, expand the lexical and syntactic members of their categories, and gradually merge categories containing lexical items with identical semantic roles and having the same usage. Finally, they develop a communal language in which all lexical items that have the same semantic role belong to a single category. Moreover, by using syntactic rules to regulate local orders of lexical members of different categories, individuals form a common, emergent global word order. Together with the development of compositionality, the development of global word order follows a 'bottom-up' routine based on the local and partial information contained in linguistic instances. This coevolution results from interactions of many factors, such as lexicon and syntax, linguistic materials and non-linguistic cues, and individual learning and social interactions. The evolution process also depends on some hierarchical relations. For example, the acquisition of lexical items provides a basis for the development of local orders, and both relevant lexical items and local orders provide a basis for selecting related categories to mediate semantic and syntactic roles. These aspects make the model suitable for exploring some embodiment-related problems.

2.3. *The indices to evaluate the performance*

Several indices are defined to evaluate the performance of the model. *Understanding Rate* (UR) calculates the average percentage of integrated meanings that are understandable to each pair of agents in the community based on their linguistic knowledge only. A high value of UR indicates that the communal language has high understandability.

$$\mathrm{UR} = \frac{\sum_{i,j} \text{Understood integrated meanings between agents } i \text{ and } j}{\text{Number of pairs of } i,\ j\ \times \text{ number of integrated meanings}}$$

Convergence Time (CT) calculates the average number of rounds of communication that is required to achieve a language with certain UR (here, 0.8; if the highest UR throughout the simulation is smaller than 0.8, directly use that UR). CT indicates the time efficiency of language emergence.

In addition, $\mathrm{UR_{ser}}$ is defined as UR between agents at time step (round) i and those at time step $i+1$ (these agents are the same, but their linguistic knowledge might differ). $\mathrm{UR_{ini}}$ is defined as UR between agents at the starting time step and those at time step i. To measure $\mathrm{UR_{ser}}$, all agents at time step i talk to those at time step i+1, and the percentage of accurately understood integrated meanings is calculated. To measure $\mathrm{UR_{ini}}$, all agents at the beginning of the simulation talk to those at time step i, and the percentage of accurately understood integrated meanings is calculated. A high value of $\mathrm{UR_{ser}}$ indicates high understandability of the communal languages across a limited time, and a high $\mathrm{UR_{ini}}$ indicates a high possibility of maintaining an initial language for a relatively long time.

3. The simulation results of the three experiments

This paper adopts the compositionality-regularity coevolution model to study the effect of social structure on both language emergence and maintenance. In the simulations on language emergence, all agents initially share eight holistic rules to encode eight integrated meanings in the semantic space. However, in the simulations on language maintenance, a compositional language that can consistently express all integrated meanings in the semantic space is initially shared among agents. This language consists of 12 compositional rules to encode all semantic items, a set of S, V, and O categories to associate corresponding lexical rules, and three syntactic rules (SV, VO, and SO) to form a consistent global order (SVO) at the sentence level. In each situation of the experiments, the results of 20 simulations are collected for statistical analysis.

3.1. *Experiment 1: a community with a single popular agent*

In Experiment 1, a 10-agent community is simulated, and the total number of communications is 6000, covering 600 rounds. There are two forms of communication: (1) communications between the popular agent (Agent 1) and others (Agents 2–10), with probability PR; and (2) communications not involving the popular agent, with probability 1−PR. In this 10-agent community, PR lies in the interval [0.1 1.0]. If it equals 0.1, all agents have equal probabilities to communicate with each other, which is similar to the random communication situation; if it equals 1.0, all communications involve the popular agent.

Figures 4 and 5 show the statistical results of the simulations on language emergence and maintenance under different values of PR. Figure 4a shows the average and standard deviation of the highest UR, and Figure 4b the average and standard deviation of CT. Figure 5a shows the average and standard deviation of last UR after 600 rounds of communications, and Figure 5b and 5c the average and standard deviation of UR_{ser} and UR_{ini}, in which avg UR_{ser} (UR_{ini}) indicates the average UR_{ser} (UR_{ini}) throughout 600 rounds of communications, and last UR_{ser} (UR_{ini}) indicates the UR_{ser} (UR_{ini}) at the end of 600 rounds of communications).

PR indicates the degree of centralisation around the popular agent. This centralisation has two effects:

(1) *The acceleration effect*. The popular agent, like a hub in a network, connects to many others. It provides a conduit for other agents to exchange information. Centralisation around it can accelerate conventionalisation of idiolects and reduce the CT of the communal language. Some research on complex networks (e.g. Crucitti et al. 2003) has shown that introducing hubs can accelerate information transmission among nodes and the synchronisation of the whole network.

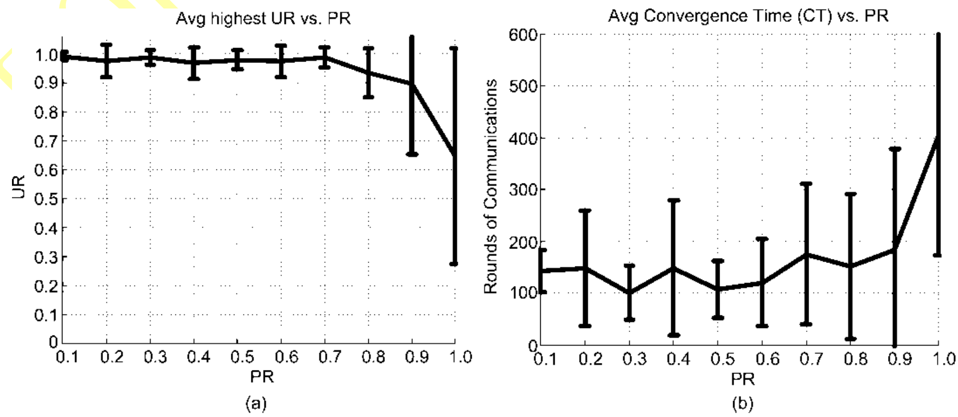

Figure 4. The statistical results in Experiment 1 on language emergence: the UR (a) and CT (b) of the emergent languages under different PR. The distance between a pair of error bars above and below a data point is twice of the standard deviation.

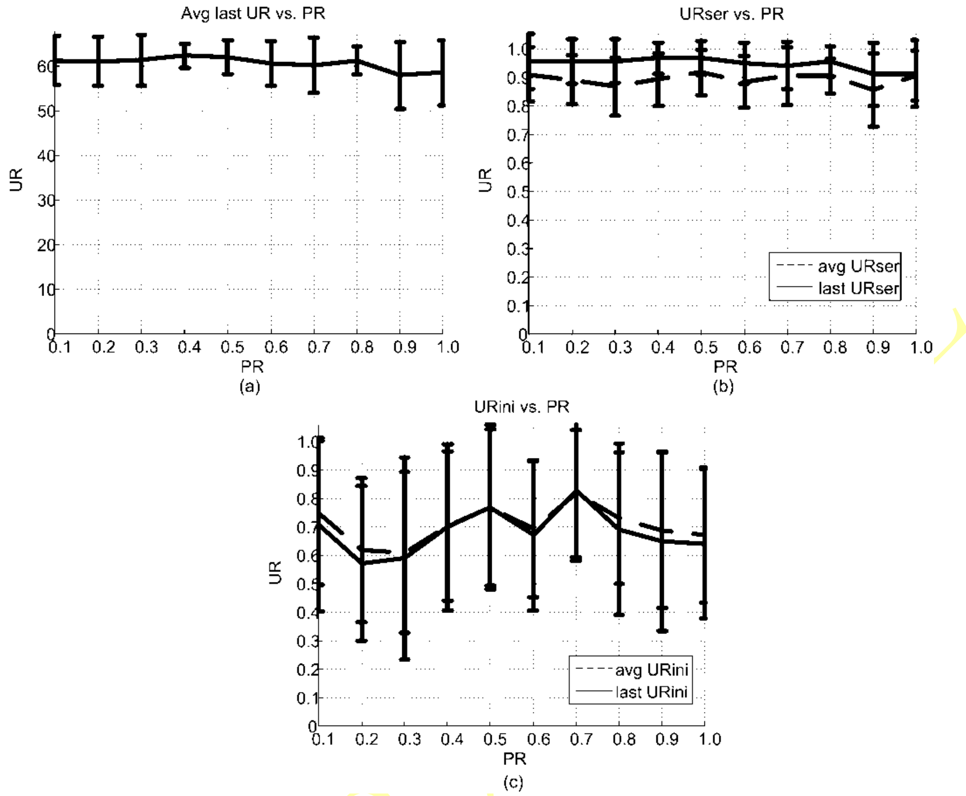

Figure 5. The statistical results in Experiment 1 on language maintenance: the UR (a), UR_{ser} (b), and UR_{ini} (c) of the communal languages under different PR.

(2) *The deceleration effect.* For agents to efficiently transmit information via a hub, this hub has to be stable. In this model, agents not only exchange linguistic information, but also modify their linguistic knowledge during communications. If the popular agent's linguistic knowledge is stable, it can transmit that knowledge unchanged to multiple agents. In other words, information transmission via stable intermediaries is effective. However, if the popular agent frequently modifies its linguistic knowledge, the knowledge that it transmits to other agents is inconsistent. In other words, information transmission via unstable intermediaries is not always effective. In addition, the syntactic knowledge in this model may not easily converge; different previous experiences may cause agents to develop different local orders based on the same global order, and the global order may in turn greatly shift when agents change one of their local orders as a result of communications with others. Via the popular agent, the syntactic information in the form of local orders may not always be efficiently transmitted to others. Both the instability of the popular agent and the imprecision introduced by local orders may affect the conventionalisation of linguistic knowledge in the population.

In Experiment 1, both of these effects coexist and compete with each other. On the one hand, in the simulations on language emergence, the popular agent, like others, continues updating its idiolect during communications. If its linguistic knowledge is changed as a result of listening to some agents whose idiolects are quite different from others, then, in the future, when it talks to others, the communal language shared by these agents might be affected. Not only does the understandability of the communal language drop, but also more communications are necessary either to recover the originally shared linguistic knowledge or to diffuse the newly acquired knowledge. With the increase in PR, other agents will have higher chances to contact the popular agent and affect its idiolect, which may enhance the deceleration effect. As shown in Figure 4, when PR is low, both effects cannot greatly affect language emergence. But when it becomes high, the average UR of the emergent languages begins to drop, the average CT increases, and the standard deviations of both increase, too. All these illustrate the deceleration effect. On the other hand, in the simulations on language maintenance, all agents already share a communal language. In this situation, the popular agent does not frequently update its idiolect, and communications via this stable intermediary are effective. As shown in Figure 5, both of the effects do not greatly affect the maintenance of the communal language; UR, UR_{ser}, and UR_{ini} all remain high under different PR.

These two effects of the popular agent on language emergence were first explored by Gong and Wang (2005), in which, due to different simulation details, the competition between the acceleration and deceleration effects resulted in an optimum UR occurring at a PR lower than 1.0. It showed that neither absolute 'democracy' (the situation of random communication) nor absolute 'dictatorship' (the situation where PR is 1.0) can efficiently achieve a communal language with a high UR. Although such optimum UR does not show up in this paper, the following conclusion still holds: if PR is too high, both the understandability and emergent process of the communal language become fluctuated.

3.2. *Experiment 2: a community with a given distribution of individuals' popularities*

Experiment 1 focuses on a 10-agent community with a single popular agent. In larger communities, stratification could cause different agents to have different probabilities to participate in communications. In Experiment 2, we use the distribution of all agents' popularities to represent the effect of the whole social structure. In addition to a 10-agent community, we also study the effect of individuals' popularities in communities having larger sizes.

Sociological research has discovered that instead of uniformity, in many social phenomena (e.g. sexual contact (Lijeros et al. 2003), vote distributions in legislator elections (Situngkir 2004), spread of rumours (Moreno 2004), Worldwide Web (Broder et al. 2000), and many others), the elements and interactions among these elements usually follow *power-law relations* (reviewed by Newman 2005). It is also evident in linguistic phenomena. For instance, the frequency of usage of any word in a corpus is approximately inversely proportional to its frequency rank (Zipf's law), and the rank of a language family (based on its size) and its size (the number of languages it contains) also follow a power-law relation (Stauffer et al. 2006). These power-law distributions are also characteristic in many self-organising systems (Bak 1996), and many factors such as preferential attachment (Barabasi amd Albert 1999) and geographical constraints (Warren et al. 2002) can help to explain the formation of power-law distribution in both social and linguistic phenomena. Considering these, in Experiment 2, the distribution of agents' popularities is assumed to follow the power-law distribution.

A power-law relation of two scalar quantities x and y is mathematically defined as follows:

$$y = ax^{-\lambda}$$

where a is a scale parameter, x represents an element or interaction in a given phenomenon, and y the frequency of this element or interaction. Drawn on log–log axes, a power-law distribution appears as a straight line, whose slope increases with the value of λ. The power-law distributions in different phenomena may have different λ values. As reported by Newman (2003), the λ value in the film actor collaboration network is 2.3, 2.0 in the email message network, and 2.1 in the telephone call network. In linguistic phenomena, Zipf's law has a λ value of 1.0, and the λ value in the distribution of language families is approximately 2.0.

In the power-law distributions of Experiment 2, a is chosen so that the sum of all probabilities is 1.0, x represents agent index from 1 to N (the number of agents), and y calculates the probability for an agent with a given index to participate in communications. The λ value lies in the interval [0.0, 3.0], which dovetails with the λ values observed in many real-world power-law distributions. If λ equals 0.0, all agents have an equal probability of communicating with others. In Experiment 2, the sampling λ values include 0.0, 1.0, 1.5, 2.0, 2.5, and 3.0, and the community sizes include 10, 30, and 50. As an example, Figure 6 shows the individuals' popularities under different power-law distributions in a 10-agent community. On the log–log axes, as λ increases, the distribution line becomes steeper, indicating that agents with bigger indices are significantly less popular than those with smaller indices, and the latter ones will communicate frequently with each other and occasionally with the former ones.

In our model, during a single round of communications, many communications are assumed to take place simultaneously among different pairs of agents. The number of these communications is proportional to the community size. In the simulations on language emergence, in order to let all agents have sufficient opportunities to develop their idiolects, the number of communications per round is set to scale to the square of community size: No.Com $\propto N^2$. Therefore, for a 10-agent community, the total number of communications is 6000 (600 rounds); for a 30-agent community, it is 54,000 (1800 rounds); and for a 50-agent community, it is 150,000 (3000 rounds). In the simulations on language maintenance, as a communal language is already shared, the number of communications per round only scales to the community size: No.Com $\propto N$. Therefore, for a 10-agent community, the total number of communications is 6000 (600 rounds); for a 30-agent community, it is 18,000 (600 rounds); and for a 50-agent community, it is 30,000 (600 rounds). Figure 7 shows the results of language emergence under different power-law distributions of individuals' popularities in different communities, in which panels a and b record the average and standard deviation of the highest UR and the last UR in these situations; panel c displays the average and standard deviation of CT in these situations. Figure 8 shows the results of language maintenance in these situations, in which panel a traces the average and standard deviation of the last UR in these situations; panels b and c display the average and standard deviation of UR_{ser} and UR_{ini} in these situations.

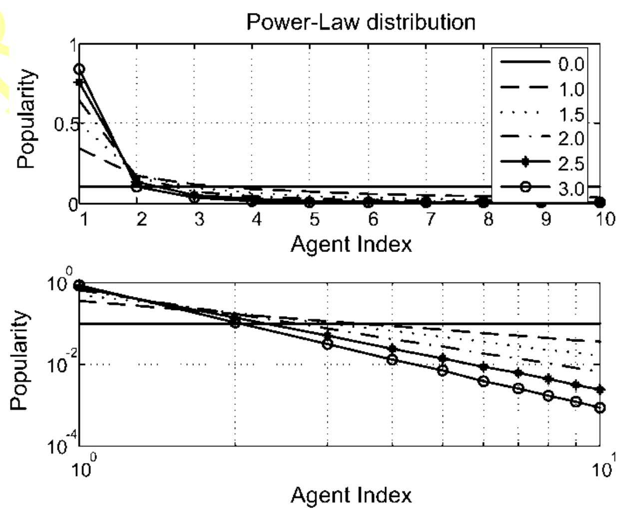

Figure 6. The individuals' popularities in different power-law distributions. The top figure is in normal axes, and the bottom one in log–log axes.

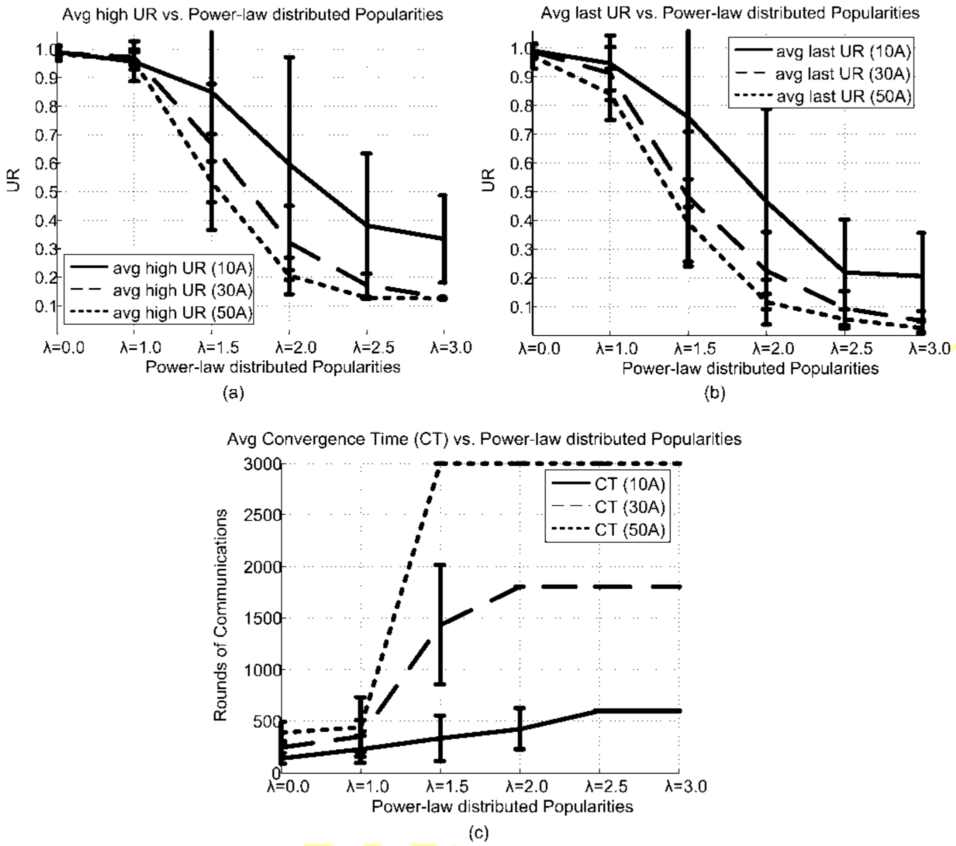

Figure 7. The statistical results in Experiment 2 on language emergence: the high UR (a), last UR (b), and CT (c) of the emergent languages in communities of different sizes and with different power-law distributions.

As shown in Figure 7, for a fixed community size, when λ is small (0.0 or 1.0), most emergent languages have a high UR exceeding 0.8. With the increase in λ, UR drops and CT increases. When λ is greater than 2.0, a communal language with a high UR rarely emerges. This tendency is more obvious in communities with bigger sizes. Meanwhile, for a fixed power-law distribution, with the increase in community size, UR drops. But when λ is 0.0 or 1.0, in all communities, agents can develop a communal language with a high UR. Similarly in Figure 8, for a fixed community size, with the increase in λ, the initial communal language gradually disappears. Both UR_{ser} and UR_{ini} decrease under power-law distributions with high λ values. Meanwhile, for a fixed λ value, with the increase in community size, both UR and UR_{ser} gradually drop. But when λ equals 0.0 or 1.0, in all communities, the initial communal languages are maintained to a certain extent (UR_{ini} is around 0.6), and both UR and UR_{ser} remain high (around 0.8). These results illustrate a *boundary* λ value (1.0, beyond which the understandability of the communal language starts to decrease) in a community with power-law distribution of individuals' popularities.

This boundary λ value results from the competition between two social trends. On the one hand, individuals in better economic or political conditions (members in higher positions of the social hierarchy) tend to become more popular and get involved in more social activities. On the other hand, members in lower positions of the social hierarchy also require sufficient social activities to maintain the communal language of the community. Without this language, the interactions between the members in different positions of the social hierarchy cannot well proceed. Then, the whole community may split up, and the social hierarchy could be demolished. The results in

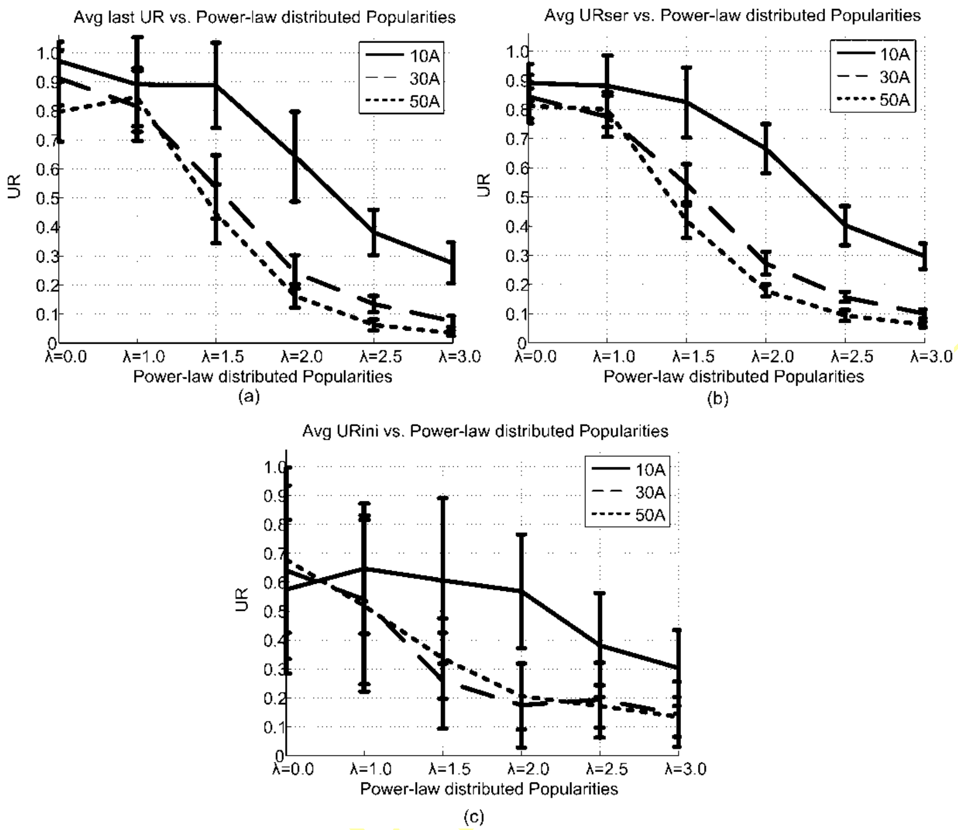

Figure 8. The statistical results in Experiment 2 on language maintenance: the UR (a), UR_{ser} (b), and UR_{ini} (c) of the communal languages in communities of different sizes and with different power-law distributions.

Experiment 2 show that a power-law distribution with an intermediate value of λ could emerge as a compromise to maintain a certain degree of both social hierarchy and mutual understanding.

The λ value in a power-law distribution of individual's popularities and the λ' value in the distribution of communications (activities) among individuals have the following relation:

$$\lambda' = 1 + \frac{1}{\lambda}$$

The mathematical proof is roughly shown below:

$$\left.\begin{array}{l} r(k) \approx \int_k^\infty p'(k) = \int_k^\infty k^{-\lambda'} dk \approx k^{1-\lambda'} \\ p(r(k)) = r(k)^{-\lambda} \approx k \end{array}\right. \Longrightarrow -\lambda = \frac{1}{1-\lambda'}$$

in which $r(k)$ is the rank of agent k, $p'(k) = k^{-\lambda'}$ is the power-law distribution of activities and $p(r) = r^{-\lambda}$ the power-law distribution of individuals' popularities, and the scaling factors are disregarded.

Considering this relation, the boundary λ value (1.0) in this model is commensurate with the λ' value (2.0) in the distribution of communications. As reviewed by Newman (2003), the power-law distributions in many social activity networks, such as the email exchange network (59,912 nodes and λ' is 2.0) and the telephone call network (47,000,000 nodes and λ' is 2.1), all have their λ' values approximately to equal 2.0. This provides an empirical support for the boundary λ value in this model. In addition, as a theoretical model, our results further predict that a social structure with a power-law distribution having bigger λ or λ' values is insufficient to trigger and maintain a relatively high level of mutual understanding. Furthermore, our finding also reflects some other simulation studies (e.g. Yang et al. 2008) which have shown that sufficiently transmitting information among nodes requires a scale-free network to have an optimal structure with a certain value of λ.

3.3. *Experiment 3: the linguistic contact between two communities*

In Experiment 3, agents from two communities can participate in either intra-community or inter-community communications, the percentages of which are controlled by *IntraRate* and *InterRate*. The simulations on language emergence explore the effect of inter-community communications on the mutual understandability of the emergent languages in these communities; the simulations on language maintenance investigate the effect of lexicon and syntax on the convergence of communal languages in these communities. In both types of simulations, two 10-agent groups are simulated, and the total number of communications is 6000 (600 rounds), in which intra-community and inter-community communications are interwoven.

Figure 9 shows the mutual understandability of the communal languages emerging in the two groups under different ratios between *IntraRate* and *InterRate*. These ratios include 100:0 (all communications are intra-community), 80:20 (roughly 480 rounds of intra-community communications and 160 rounds of inter-community communications), 50:50, 20:80, and 0:100. Figure 9a shows the average and standard deviation of the highest $\mathrm{UR}_{\mathrm{Group1}}$ (UR among agents in Group 1), $\mathrm{UR}_{\mathrm{Group2}}$, and $\mathrm{UR}_{\mathrm{Cross-group}}$ (UR between agents in Group 1 and those in Group 2) in simulations with different ratios of *IntraRate* over *InterRate*, and Figure 9b displays the average and standard deviation of CT in these simulations.

In these simulations, each community can develop its own communal language with a high UR. The mutual understandability of these communal languages is determined by the percentage of inter-community communications in all communications. As this percentage increases, $\mathrm{UR}_{\mathrm{Cross-group}}$ begins to grow, indicating that agents from the two communities can understand each other better. Although these results are consistent with previous studies showing that the degree of linguistic contact can affect the convergence of communal languages, they further illustrate that linguistic contact can take effect quite early during language emergence.

Besides the percentage of inter-community communications, similarities in certain linguistic features of communal languages may also affect the convergence of communal languages. Based on the compositionality-regularity coevolution model, we focus on two linguistic features: lexical

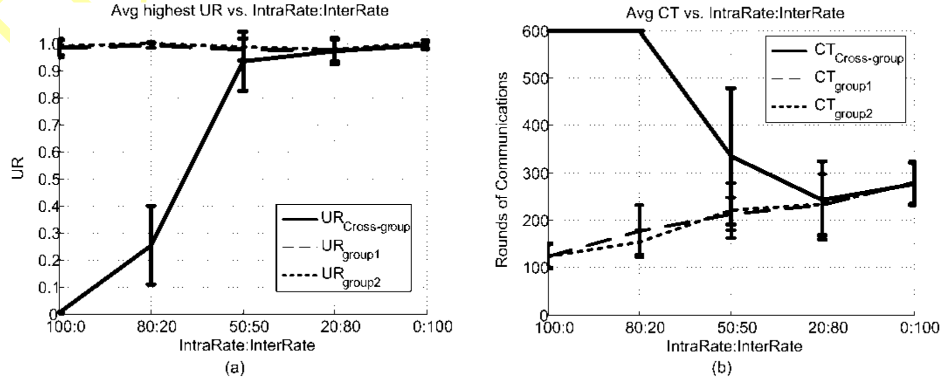

Figure 9. The statistical results in Experiment 3 on language emergence: the UR (a) and CT (b) of the communal languages in different communities under various degrees of linguistic contact.

Table 1. The four situations for the simulations in Experiment 3 on language maintenance.

Case	Lexical items	Simple syntax
1 (LS&SS)	Similar	Similar
2 (LS&SD)	Similar	Different
3 (LD&SS)	Different	Similar
4 (LD&SS)	Different	Different

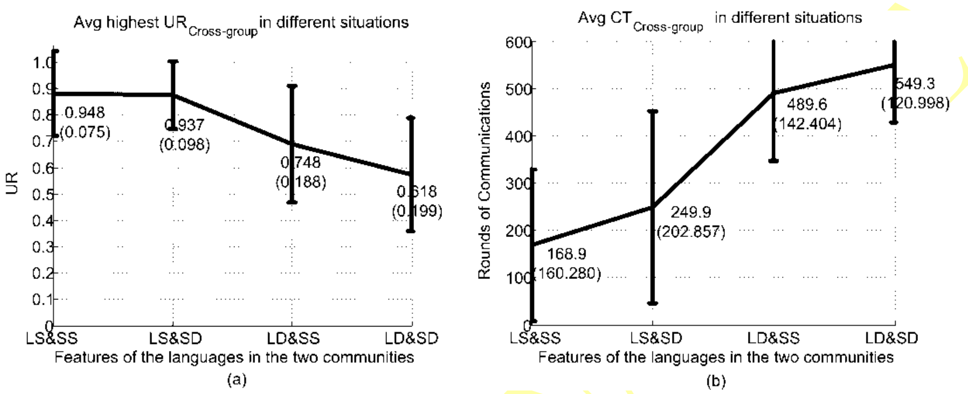

Figure 10. The statistical results in Experiment 3 on language maintenance: the $UR_{Cross-group}$ between two communal languages sharing different linguistic features. Numbers outside brackets are average values, and numbers inside brackets standard deviation values.

items and simple syntax. There are four situations to be considered, as listed in Table 1. Two communal languages are said to *have similar lexical items*, if 80% of their lexical rules are identical; they are said to *have similar syntax*, if two of their local orders are identical, e.g. SV, VO, and SO in one, and SV, OV, and SO in the other; they are said to *have different syntax*, if all their local orders are different, e.g. SV, VO, and SO in one, and VS, OV, and OS in the other. The indices to evaluate the convergence of communal languages include $UR_{Cross-group}$ and $CT_{Cross-group}$ (the average number of rounds of communications required for $UR_{Cross-group}$ to reach its highest value). Figure 10 shows the simulation results of these four situations for a given ratio between *IntraRate* and *InterRate* (60:40).

As shown in Figure 10, the average UR in Cases 1 and 2 is higher than that in Cases 3 and 4, and the average CT in Cases 1 and 2 is smaller than that in Cases 3 and 4, which suggest that for a fixed degree of inter-community communications, sharing lexical items is more efficient than sharing syntactic rules to converge communal languages. This difference is caused by linguistic features and language processing mechanisms. In the model used here, there are two types of integrated meaning: ‘Pr1<Ag>’ and ‘Pr2<Ag, Pat>’. For Type1 meanings, once the related lexical items have been learned, the whole integrated meanings can be produced or comprehended. For Type2 meanings, however, the syntax mainly plays the role of distinguishing Ag and Pat. This role must be fulfilled after the related lexical items have been learned. Therefore, the convergence of communal languages sharing similar lexical items is easier than the convergence of communal languages sharing similar local orders.

4. Discussions and conclusions

Language evolves primarily via social contact among a finite number of individuals. In this paper, based on a computational model that simulates individual learning mechanisms to acquire some linguistic features, we briefly discuss the effect of social structure on both language emergence and maintenance in three experiments.

In Experiment 1, the acceleration and deceleration effects of a single popular agent are pointed out, and their collective impact on language evolution is discussed. It shows that a biased social learning towards a particular agent has both advantages and disadvantages. Neither totally unbiased nor totally biased social learning can result in a good level of mutual understanding in the population.

In Experiment 2, the impact on language evolution of individuals having power-law distributed popularities is discussed. It demonstrates that in both small and large communities, in order to maintain both an efficient communication system and a degree of social hierarchy, the λ value of the power-law distribution should not go very high. This result also reflects the finding in Experiment 1 that highly biased social learning (with higher λ values) can impair the level of mutual understanding.

In Experiment 3, social learning involving the contact between two communities is discussed. It demonstrates that different degrees of linguistic contact can affect the convergence of a communal language. And under a given pattern of linguistic contact, some linguistic features such as lexical items can affect the convergence of communal languages.

Without implementing any particular social structure, in these experiments, we mainly control the probabilities for individuals to participate in communications. This approach has some advantages for studying the effect of social structure on language evolution. First, from a macro perspective, social structures with various connection patterns may share similar characteristics that can cast their influence on linguistic exchange. The approach of manipulating only the probabilities for agents to participate in communications can summarise these general principles of the social structure. For instance, as discussed in this paper, many social structures exhibit power-law distributed language-related interaction – the results obtained for these distributions could be independent of specific social structures. Second, instead of considering many parameters to define a specific social structure, our approach focuses on a limited number of probability parameters. The approximate trajectory of language evolution can be easily analysed through manipulating these few parameters.

Following the embodiment perspective, our behavioural model takes account of several important interactions during social learning. For example, the involvement of non-linguistic information during social interactions provides the physical grounding of social learning. Also, the evolution of language through iterated social interactions concerns the development of both lexical and syntactic items, as well as their interactions. These aspects make the process of social interactions in our study more realistic and illustrate that the social learning process is typically dependent on many linguistic and non-linguistic factors. Moreover, our work touches upon the interaction of internal linguistic features and external social learning and exemplifies the influence of some linguistic features on the efficiency of social learning. This aspect could be insightful for empirical studies of the competition among languages sharing similar linguistic features.

Despite these advantages, some extensions based on the current approach can help to systematically study the effect of social structure on language evolution. For example, due to frequent changes in individuals' activities and communication patterns, the associated social networks are subject to constant evolution (Palla et al. 2007). The probabilities that individuals participate in communications should be frequently updated to reflect changes in social structures. In addition, PR, *Individual's Popularity*, *IntraRate*, and *InterRate* only determine the probability for agents to participate in communications, without specifying their roles (speakers or listeners) in communications. As shown in some theoretical simulations, these roles can affect language evolution. For instance, as studied by Baronchelli et al. (2006), when the hubs in a scale-free network are frequently chosen as speakers, they tend to easily propagate successful words to others; the resultant convergence of language is faster than when the hubs are frequently chosen as listeners.

Furthermore, apart from power-law distributed popularities, many social structures exhibit *the small-world phenomenon* (the principle that most of us are linked by short chains of acquaintances (Watt and Strogatz 1998), such as acquaintance networks (Bernard et al. 1998), scientific citation networks (Nowak et al. 1990), and so on). A key feature of small-world networks is the existence of 'shortcuts' between certain nodes. These shortcuts can affect information transmission (Newman 2000). To embody the effect of shortcuts, we can introduce additional parameters that specify the probabilities of 'shortcut' communications among particular agents. All these extensions pave the way for future explorations of the effect of social structure on language evolution.

Acknowledgement

This work was conducted at the Language Engineering Laboratory (LEL) in the Chinese University of Hong Kong (CUHK), with support from the Research Grants Council of Hong Kong. Some results were presented in IEEE World Congress on Computational Intelligence 2006. We are indebted to Professors Thomas Lee from CUHK, Vittorio Loreto from 'La Sapienza' University of Rome, Italy, and LEL colleagues F. Wong, L. Shuai, and H.-Y. Zheng for the useful suggestions and discussions. This article has also greatly benefited from the comments of its anonymous reviewers.

References

Alessa, L.N., Laituri, M., and Barton, M. (2006), "An 'All Hands' Call to the Social Science Community: Establishing a Community Framework for Complexity Modeling Using Agent Based Models and Cyberinfrastructure," *Journal of Artificial Societies and Social Simulation*, 9, available at http://jasss.soc.surrey.ac.uk/9/4/6.html.

Bak, P. (1996), *How Nature Works: The Science of Self-Organized Criticality*, New York: Copernicus.

Barabási, A.-L. and Albert, R. (1999), "Emergence of scaling in random networks," *Science*, 286, 509–512.

Barnard, A. (ed.) (2003), *Hunter-Gatherers in History, Archaeology and Anthropology*, New York: Berg, Oxford.

Baronchelli, A., Loreto, V., Dall'Asta, L., and Barrat, A. (2006), "Bootstrapping Communication in Language Games: Strategy, Topology and All That," in *The Evolution of Language: Proceedings of the 6th International Conference*, eds. A. Cangelosi, A. D. M. Smith, and K. Smith, London: Word Scientific Publishing Co., pp. 11–18.

Bernard, H.R., Kilworth, P.D., Evans, M.J., McCarty, C., and Selley, G.A. (1988), "Studying Social Relations Cross-Culturally," *Ethnology*, 27, 155–179.

Bhattacharjee, Y. (2003), "From Heofonum to Heavens," *Science*, 303 , 1326–1328.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. (2000), "Graph Structure in the Web," *Computer Networks: The International Journal of Computer and Telecommunications*, 33, 309–320.

Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2003), "Efficiency of Scale-Free Networks: Error and Attack Tolerance," *Physica A*, 320, 622–642.

de Waal, F.B.M. (2005), "A Century of Getting to Know the Chimpanzee," *Nature*, 437, 56–59.

Dunbar, R. (1996), *Grooming, Gossip, and the Evolution of Language*, Cambridge, MA: Harvard University Press.

Gong, T. (2008), *Computational Simulation in Evolutionary Linguistics: A Study on Language Emergence,* Frontiers in Linguistics, Monograph IV. Taiwan: Academia Sinica.

Gong, T. and Wang, W.S.-Y. (2005), "Computational Modeling on Language Emergence: A Coevolution Model of Lexicon, Syntax and Social Structure," *Language and Linguistics*, 6, 1–42.

Ke, J.-Y. (2004), "Self-Organization And Language Evolution: System, Population and Individual," Doctoral Dissertation, City University of Hong Kong.

Labov, W. (1994), *Principles of Linguistic Change: Internal Factors*, Oxford, UK: Basil Blackwell.

———. (2001), *Principles of Linguistic Change: Social Factors*, Oxford, UK: Basil Blackwell.

Lijeros, F., Edling, C.R., and Amaral, L.A.N. (2003), "Sexual Networks: Implications for the Transmission of Sexually Transmitted Infections," *Microbes and Infection*, 5, 189–196.

Livingstone, D. (2001), "The Evolution of Dialect Diversity," in *Simulating the Evolution of Language,* eds. A. Cangelosi and D. Parisi, London: Springer-Verlag, pp. 99–117.

Malsch, T. and Schulz-Schaeffer, I. (2007), "Socionics: Sociological Concepts for Social Systems of Artificial (and Human) Agents," *Journal of Artificial Societies and Social Simulation*, 10, available at http://jasss/soc/surrey.ac.uk/10/1/11.html.

Minett, J.W., Gong, T., and Wang, W.S.-Y. (2006), "A Language Emergence Model Predicts Word Order Bias," in *The Evolution of Language: Proceedings of the 6th International Conference*, eds. A. Cangelosi, A. D. M. Smith, and K. Smith, London: World Scientific Publishing Co., pp. 206–213.

Moreno, Y., Nekovee, M., and Pacheco, A.F. (2004), "Dynamics of Rumor Spreading in Complex Networks," *Physical Review E*, 69, 1–7.

Nettle, D. (1999), "Using Social Impact Theory to Simulate Language Change," *Lingua*, 108, 95–117.

Newman, M.E.J. (2000), "Models of the Small World: A Review," *Journal of Statistical Physics*, 101, 819–841.
———. (2003), "The Structure and Function of Complex Networks," *SIAM Review*, 45, 167–256.
———. (2005), "Power Laws, Pareto Distributions and Zipf's Law," *Contemporary Physics*, 46, 323–351.
Nowak, M.A., Szamrej, J., and Latané, B. (1990), "From Private Attitude to Public Opinion: A Dynamical Theory of Social Impact," *Psychological Review*, 97, 362–376.
Palla, G., Barabási, A.-L., and Vicsek, T. (2007), "Quantifying Social Group Evolution," *Nature*, 446, 664–667.
Situngkir, H. (2004), *Power Law Signature In Indonesian Legislative Election 1999–2004*, Research Paper WPL2004, Bandung Fe Institute.
Stauffer, D., Schulze, C., Lima, F.W.S., Wichmann, S., and Solomon, S. (2006), "Non-Equilibrium and Irreversible Simulation of Competition Among Languages," *Physica A*, 371, 719–724.
Thomason, S.G. and Kaufman, T. (1988), "*Language Contact, Creolization, and Genetic Linguistics*," Berkeley, Los Angeles, Oxford: University of California Press.
Warren, C.P., Sander, L.M., and Sokolov, I.M. (2002), "Geography in a Scale-Free Network Model," *Physical Review E*, 66, 1–5.
Watt, D.J. and Strogatz, S.H. (1998), "Collective Dynamics of Small-World Networks," *Nature*, 393, 440–442.
Whiten, A. (2005), "The Second Inheritance System of Chimpanzees and Humans," *Nature*, 437, 52–55.
Yang, H.-X., Wang, W.-X., and Wang, B.H. (2008), "Asymmetric negotiation in structured languages games," *Physical Review E*, 77, 027103.