

Language Change and Social Networks

Jinyun Ke^{1,*}, Tao Gong² and William S-Y Wang²

¹ *English Language Institute, University of Michigan, Ann Arbor, MI 48104-2028, USA.*

² *Language Engineering Laboratory, Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong.*

Received 5 August 2007; Accepted (in revised version) 27 August 2007

Communicated by Dietrich Stauffer

Available online 11 December 2007

Abstract. Social networks play an important role in determining the dynamics and outcome of language change. Early empirical studies only examine small-scale local social networks, and focus on the relationship between the individual speakers' linguistic behaviors and their characteristics in the network. In contrast, computer models can provide an efficient tool to consider large-scale networks with different structures and discuss the long-term effect of individuals' learning and interaction on language change. This paper presents an agent-based computer model which simulates language change as a process of innovation diffusion, to address the threshold problem of language change. In the model, the population is implemented as a network of agents with age differences and different learning abilities, and the population is changing, with new agents born periodically to replace old ones. Four typical types of networks and their effect on the diffusion dynamics are examined. When the functional bias is sufficiently high, innovations always diffuse to the whole population in a linear manner in regular and small-world networks, but diffuse quickly in a sharp S-curve in random and scale-free networks. The success rate of diffusion is higher in regular and small-world networks than in random and scale-free networks. In addition, the model shows that as long as the population contains a small number of statistical learners who can learn and use both linguistic variants statistically according to the impact of these variants in the input, there is a very high probability for linguistic innovations with only small functional advantage to overcome the threshold of diffusion.

AMS subject classifications: 91.D30

PACS: 89.65.-s, 89.75.Hc

Key words: Language change, social network, agent-based modeling.

*Corresponding author. *Email addresses:* jyke@umich.edu (J. Ke), gtojty@gmail.com (T. Gong), wsywang@ee.cuhk.edu.hk (W. S. Y. Wang)

1 Introduction

Social network is considered as a determining factor in language change, contact, maintenance and shift, etc. (Labov 2001, de Bot & Stoessel 2002). In sociolinguistics, empirical studies of social network often examine in detail the networks of small communities and focus on the relation between individuals' social network properties and their linguistic performance (Milroy 1980/1987, Eckert 2000). A classic study in this area done by Milroy and colleagues (Milroy 1980/1987) examined three stable inner-city communities of Belfast in Britain, and found that the working class communities have "close-knit" social networks in common. These networks are of high density and multiplex: individuals usually have multiple relationships, being relatives, neighbors, friends, and/or colleagues, and they vary in their degrees of integration into the community, some having very few links with individuals outside their social group, while others having fewer links within the group but more links outside. Studies have quantitatively shown that individuals' linguistic behaviors are highly correlated with their degrees of integration into the network: in situations where linguistic variations are present in the community, the more integrated an individual is into the community, the less variation (s)he has, and the better (s)he conforms to the speech norm of the community.

Most of such empirical studies only focus on synchronic linguistic variations in small communities, and few have touched upon the question on how different social networks affect language change at a larger historical scale. In fact, social network has created a *paradox* in the study of language change: although intuitively one would think that social network should be an important factor in determining language change, very few empirical data have been able to show the effect of social network quantitatively over long periods of time (de Bot & Stoessel 2002). It is hardly possible to get a clear picture of the social structure of a large community at present with respect to individuals' linguistic behaviors, not to mention the social structure in the past.

This gap can be filled by computer simulation which provides a convenient platform to systematically study the effect of social network under controlled conditions (Gong 2007, Parisi & Mirolli 2007). Computer simulation can manipulate various parameters, such as population size, network connectivity, and so on, and it is particularly effective in addressing problems at a large time-scale beyond empirical studies of social networks. These problems include how the dynamics differ in populations with different social structures, how the structure of social network affects the rate of change, and so on.

However, despite of these advantages, existing computer models of language change either do not consider the actual population structure (Niyogi & Berwick 1997, Niyogi 2006), or simply assume the population structure as regular or random networks (Nettle 1999a). For instance, in Nettle's model of language change, the population structure is implemented as a weighted regular network, as shown in Fig. 1, in which each agent is connected to all his neighbors, and the strength of connection is inversely proportional to the distance between two agents.

Recent studies on large-scale complex networks in the real world (Barabási 2002) re-

5	4	3	2	1	1	2	3	4	5	5	4	3	2	1	1	2	3	4	5
5	4	3	2	1	1	2	3	4	5	5	4	3	2	1	1	2	3	4	5
5	4	3	2	1	1	2	3	4	5	5	4	3	2	1	1	2	3	4	5
5	4	3	2	1	1	2	3	4	5	5	4	3	2	1	1	2	3	4	5

Figure 1: The social network in Nettle's model of language change (1999). The numbers represent agents at different age stages. Nodes 1 and 2 represent infants and children, 3, 4 and 5 represent adults.

veal that most sparsely connected networks, such as the Internet, scientist collaboration networks, friendship networks, etc., are neither regular nor random. Two important features have been discovered in these networks: *small-world* (Watts & Strogatz 1998) and *scale-free* (Barabási & Albert 1999). Recently a few computer models studying the evolution of vocabulary show that random networks and scale-free networks produce different outcome in the convergence of vocabulary in the population (Dall'Asta & Baronchelli 2006, Dall'Asta et al. 2006, Kalampokis et al. 2007). In this paper, we will examine the effect of social networks on the dynamics and outcome of language change, using a computer model that simulates language change as a process of innovation diffusion.

2 Language change as a diffusion process

Language change can be viewed as a diffusion process of some new linguistic elements (*linguistic innovations*) in a language community (Shen 1997, Nettle 1999a, Wang et al. 2004). This process can be divided into two sub-processes: "innovation" and "diffusion" (or "propagation") (Croft 2000)[†]. In this paper, we focus on the diffusion process and assume that the innovation is present at the beginning of the process without considering its source.

It has been largely accepted that most adults change their language little after childhood, and language change mainly happens through children's learning (c.f. see a critic in Croft 2000). In biological evolution, when a new mutant trait arises in an individual, it has a good chance to be passed on to the offspring of that individual, as long as the mutant is not severely deleterious or actually lethal. But linguistic transmission is different from genetic transmission. Instead of inheriting genes from *one or two parents*, a language learner samples at least a proportion of the language community, which may include *a fairly large number of people* in the generations above him as well as in his peer group. Therefore, the innovation, or the mutant, being the minority at the beginning, is unlikely to be learned by the next generation. This is the "threshold problem" of language change (Nettle 1999b). In order for an innovation to spread and become the new norm in a language community, it must pass "*a threshold of frequency*" (ibid).

[†]There have been similar proposals of dividing the process of language change into two sub-processes, for example, Weinreich et al. (1968)'s "actuation" and "transmission", and Chen & Wang (1975)'s "actuation" and "implementation", though there are minor differences in the definitions of these terms (Croft 2000).

There are two possibilities for the innovation to overcome the threshold (Nettle 1999b). One is “functional selection”, i.e., there is a *functional bias* toward the innovation over the original norm. Studies on language universals and language evolution have proposed various functional accounts, such as perceptual salience, production economy, markedness, iconicity, etc. (Croft 1990/2003, Kirby 1999). The other possibility to cross the threshold is “social selection”, in which the innovation originates from some influential speakers who have higher influence, or “social impact”, than others, and learners may favor learning from them.

Nettle proposed a model to study the threshold problem in language change. This model is adapted from the Social Impact Theory that simulates attitude change in social groups (Nowak et al. 1990). In the model, the population is structured with age and social status. The language learner chooses one of the competing linguistic variants by evaluating their impact in the community. Individuals within a shorter social distance or with a higher social status have a stronger impact on the learner. This model demonstrates that in a community homogeneous in social status, the functional bias needs to be unrealistically high in order for an innovation to spread successfully; but with social selection, in a community heterogeneous in social status, an innovation with a very small functional advantage has a high chance to spread. Concluding from these simulation results, Nettle suggests that functional biases may affect the direction of language change, but cannot provide a sufficient condition for change to occur. “Without the potential for change provided by differences in social influence, functionally favored variants might never overcome the threshold required to displace prior norms” (Nettle 1999b: 116).

However, Nettle’s conclusion that language change requires the existence of super-influential agents to ensure the diffusion of an innovation faces a challenge in explaining “changes from below” (Labov 2001): there are a lot of changes in which innovations diffuse, not from the highest social class, but from the upper working class or lower middle class, who are considered as having less social impact. In addition, the regular population structure in Nettle’s model is also problematic. In this paper, we will present a new model with various network structures and different types of learners. The model shows that these factors can affect not only the dynamics but also the threshold for innovation to diffuse.

3 The model

In our model, the population is represented as a network containing N nodes (agents) and some connections among them. Each agent has one of the two linguistic states, using either unchanged form “U” or the innovation, i.e. the changed form “C”, based on which form they have learned. The model adopts the age structure in Nettle’s model, but unlike Nettle’s grid structure, children and adults are randomly distributed in the network. Each agent has an age stage ranging from 1 to 5. Agents at stage 1 are infants who can only learn from their connected teachers; agents at stage 2 can both learn and

teach others. Therefore, agents at stage 1 and 2 are learners, while agents at stage 3-5 are adults who can only teach learners and cannot change their own states. The ratio between adults and learner is 3:2. After each time step, all agents advance in age by one stage, and those at stage 5 will be replaced by new infants who inherit the connections of the adults they replace, but have their linguistic states undetermined till they start to learn.

Here we give an illustration on how a learner may learn from his connected neighbors. When there are both "U" and "C" present in the input, the learner will learn a form that has a higher fitness. The fitness of a form is measured by a function of incorporating the functional value and the frequency of that form, as represented by the equations below:

$$F(U) = f_U q_U, \quad (3.1)$$

$$F(C) = f_C q_C, \quad (3.2)$$

in which f_U and f_C are the functional values of the two forms U and C respectively, while q_U and q_C are their frequencies in the learner's connected neighborhood. The state of the learner L is determined by

$$S(L) = \begin{cases} U & \text{if } F(U) > F(C), \\ C & \text{if } F(U) \leq F(C). \end{cases} \quad (3.3)$$

For example, in a network with 10 agents, a learner is connected with 4 agents, three of which use "U" and one uses "C". If the functional values for U and C are 1 and 4, then $F(U)=3$ and $F(C)=4$. Therefore the learner will learn the "C" form. For the sake of convenience, hereafter we will assume $f_U = 1$, and use a parameter called the functional bias β , which measures the functional advantage of C over U, i.e., $\beta = f_C / f_U$.

The model compares the diffusion processes in four different kinds of network structure: random, regular, small-world and scale-free networks. A regular network is built as a ring, each node having an equal number of connections to its nearest neighbors, and a random network is set up by connecting two nodes based on a probability which is determined by the given connectivity of the network. A small-world network is built based on the model proposed by Watts & Strogatz (1998). It starts from a regular network, and rewires a number of regular links randomly according to a probability p . The rewiring probability p determines statistically how many numbers of regular connections are changed into shortcuts. In this study, p is set as 0.01, which is within the range where the small-world characteristics are best represented. A scale-free network is built following the growth model proposed by Barabási & Albert (1999).

4 The effect of different types of networks

We first simulate the diffusion process in the above four types of networks, each for 20 runs, with the same set of conditions: the network size N is 500, the average degree $\langle k \rangle$

is 20 and the functional bias β is 20. The innovation C is randomly generated from one adult; all other agents are set as U . The simulation results are shown in Fig. 2, in which each curve in a graph tracks one diffusion process. Under this condition, diffusion is successful in all runs, but the diffusion curves are different in different networks. The innovation diffuses in a linear way in a regular network, while in other networks, the diffusion follows a sharp S-curve. The diffusion rate is much slower in a regular network than in other networks, which can be explained by its lack of "short-cut" connections between distant nodes in the population, while the other three networks have these "short-cuts".

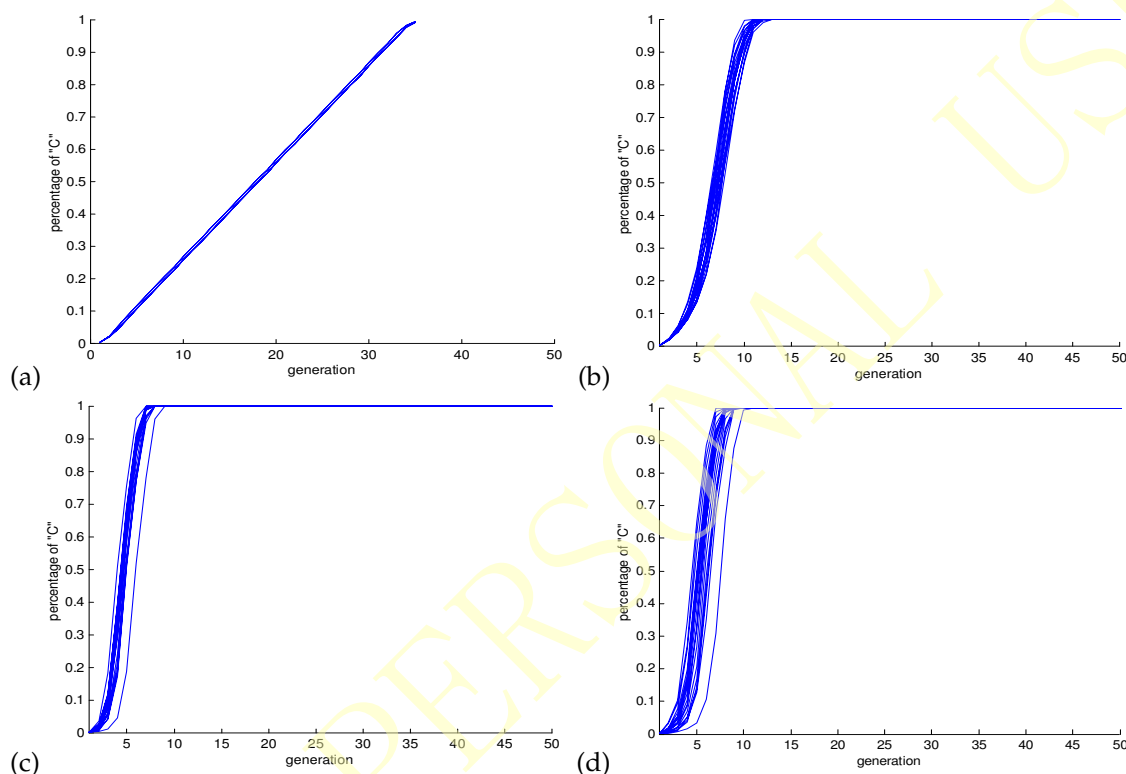


Figure 2: Diffusion dynamics in four types of networks in 20 runs (x axis: the number of generations, y axis: the percentage of changed form used in the population) (population size $N=500$, average degree $\langle k \rangle=20$, functional bias $\beta=20$, and number of innovators $I=1$). (a) regular network; (b) small-world network; (c) random network; (d) scale-free network.

It is obvious that when the functional bias is large enough, diffusion always completes, reaching the whole population. Fig. 3 shows the diffusion dynamics under another set of condition: functional bias decreases from 20 to 10, and the number of innovators increases to 10, randomly distributed. Regular and small-world networks show similar gradual diffusion, while random and scale-free networks still exhibit the rapid diffusion. The fact that the small-world networks now show a pattern different from the first condition indicates that number of short-cuts in a small-world network is still small,

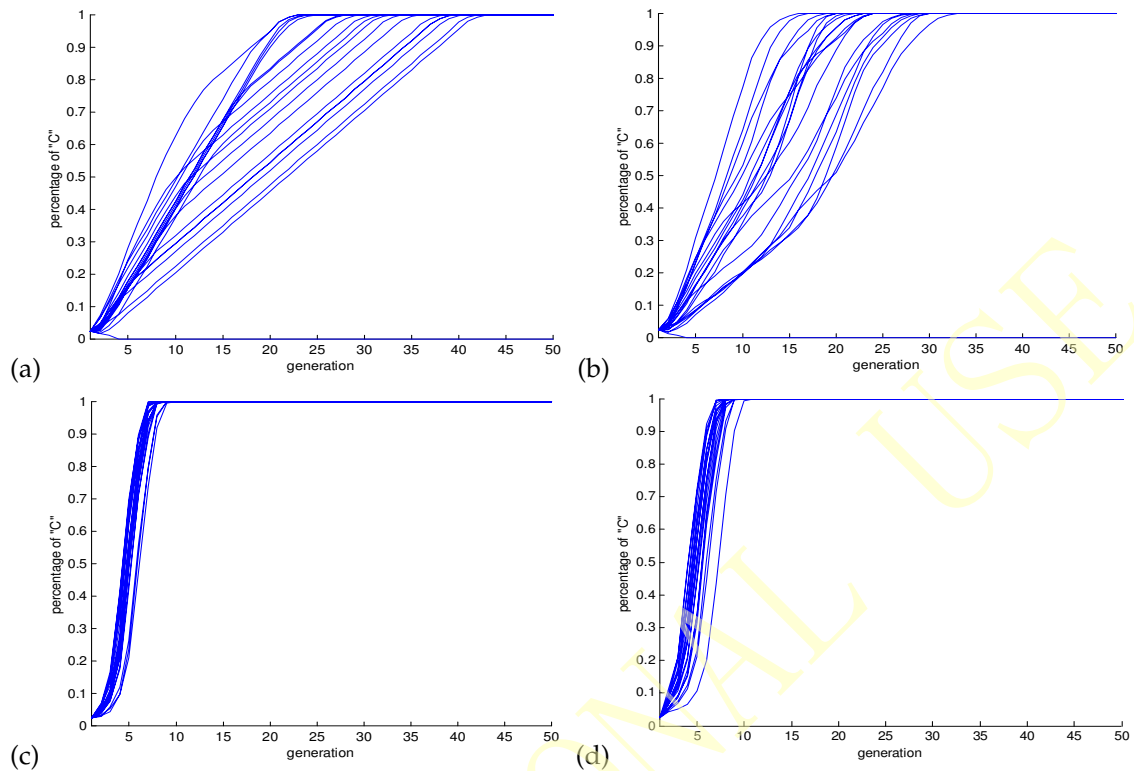


Figure 3: Diffusion dynamics in four types of networks in 20 runs under another set of condition: $N=500$, $\langle k \rangle=20$, $\beta=10$, $I=10$. (a) regular network; (b) small-world network; (c) random network; (d) scale-free network.

and when the functional bias is not big enough, the C has to diffuse slowly into the whole population in a small-world network. In contrast, the random and scale-free networks have a large number of shortcuts, and the diffusion in them is insensitive to the functional bias, always appearing in a sharp form, which we will continue to see in other sets of conditions in the following.

When the functional bias is further decreased to 2, the innovation cannot spread unless there are a large number of innovators. Fig. 4 shows the results when there are 100 innovators. This situation may correspond to the case of a massive immigration flow. It is clear that regular and small-world networks again show gradual diffusion, and random and scale-free networks show rapid diffusion in a sharp S-curve. Moreover, there are a large number of runs with unsuccessful diffusion in the latter two types of networks; in both cases, the rate of unsuccessful diffusion is about 85%. If real social networks are like random or scale-free networks, the different outcomes under the same condition may explain why while similar linguistic innovations constantly appear, only a small number of them successfully diffuse into the population.

In the following, we statistically test the effect of functional bias and number of in-

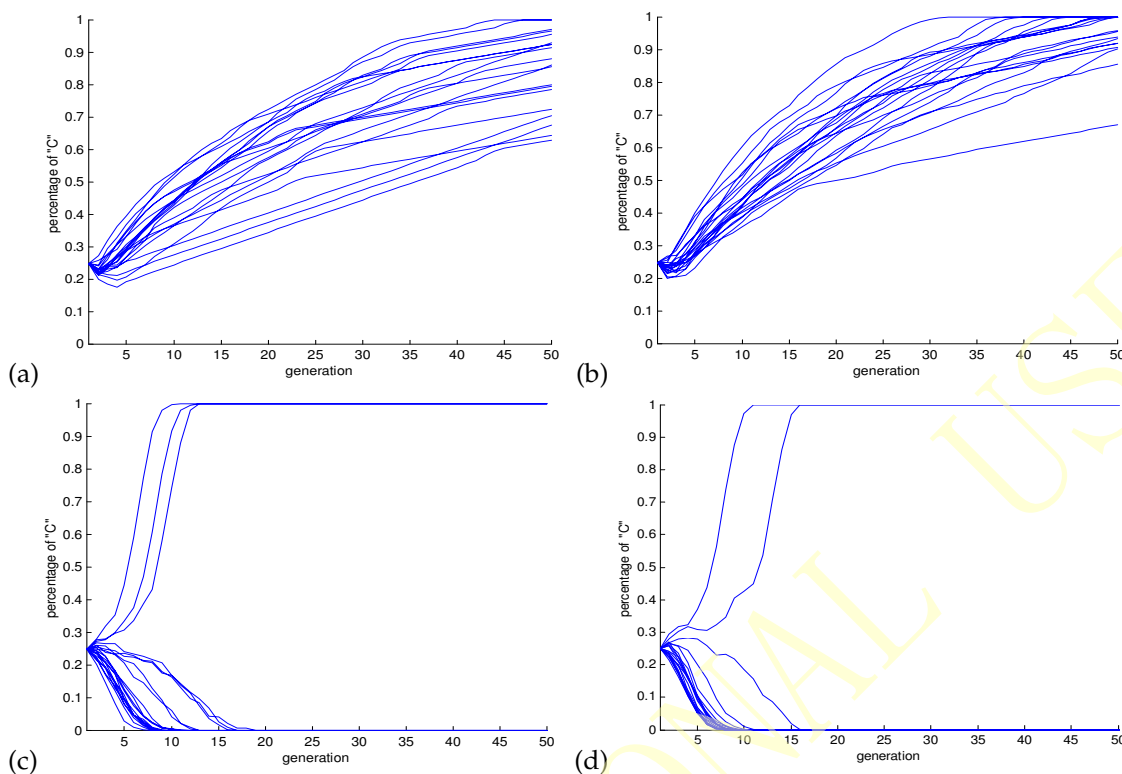


Figure 4: Diffusion dynamics in four types of networks in 20 runs with a small functional bias ($\beta=2$) but a large number of innovators ($I=100$). (a) regular network; (b) small-world network; (c) random network; (d) scale-free network.

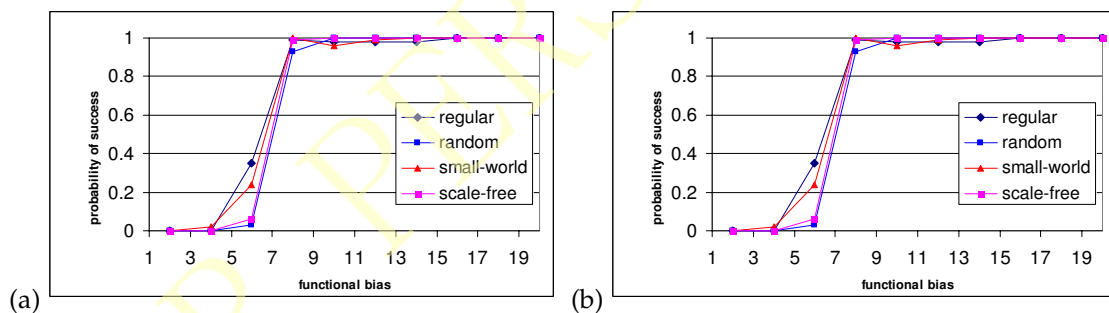


Figure 5: (a) Probabilities of successful diffusion under different functional biases; (b) Average diffusion time over 100 runs ($N=400$, $\langle k \rangle=20$, $I=10$).

novators in these 4 types of networks. Fig. 5 gives the probability of successful diffusion over 100 runs under different functional biases. Fig. 5(a) shows that when functional biases are small ($\beta < 3$), there is no diffusion in all networks; for a range of functional biases ($\beta=3 \sim 7$), regular and small-world networks have higher probabilities of successful diffusion than scale-free and random networks (Fig. 5(a)). In other words, the last two types of

networks have a higher threshold of functional bias for successful diffusion. Meanwhile, the last two types of networks take much less time to complete the diffusion than the first two, as indicated by the average diffusion time over 100 runs shown in Fig. 5(b). When the functional bias is high enough ($\beta > 7$), there are little differences between the four types of networks. Therefore, within a range of small functional bias and small number of innovators, the four types of networks exhibit different characteristics. The dynamics in small-world networks is similar to that in regular networks, i.e., high success probability, but slow diffusion rate. The dynamics in scale-free networks is similar to that in random networks, i.e., fast diffusion rate, but lower success probability.

The simulation results of the two types of dynamics in these four types of networks illustrate the importance and necessity of identifying the actual linguistic interaction network in reality.

5 Effect of two types of learners

In the above simulations, the agents in the model, faced with the presence of the two competing variants, learn and use only one of them. This is inconsistent with the empirical findings from the studies of on-going language change, in which the co-existence of variation in individual speakers is prevalent. Ke (2004) proposed that there could be two types of learners, “categorical” and “statistical”, in terms of their capacity to accommodate competing variants. A categorical learner only learns and uses the form which is encountered often and early enough during the acquisition period, while a probabilistic learner acquires both forms and uses them in proportion to their frequency in the input[‡]. Recent empirical studies on learning artificial language also show that there exist two types of learners given input with variation: some learn the probabilistic characteristics of the variation, while some others tend to generalize and show a categorical learning outcome (Hudson Kam and Newport 2005). Therefore, in our model, we take account of this distinction in agents’ learning behaviors.

The learners in our model learn from all connected neighbors. At age stages 1 and 2, learners evaluate the impact of the forms, if they encounter more than one during their learning period. The impact of the variant form is measured by the product of its functional bias and frequency as given in Eqs. (3.1)-(3.2). A categorical learner adopts only the form which has higher impact, while a probabilistic learner may adopt both forms and use them probabilistically proportional to their impact.

At the beginning when the innovation “C” is still rare in the population, learners will most likely only encounter “U”, and therefore, they only learn and use “U”. But at later stages of the change when the innovation has diffused to more speakers, learners are likely to be exposed to both “U” and “C”. If a learner has encountered “U” three times and “C” twice from his teachers, and if the functional bias β is 2, then a categorical

[‡]These two types of learning, categorical and probabilistic, may correspond to the majority and voter models in the physics literature (Liggett 1985).

learner will use “C” consistently in his adulthood, while a probabilistic learner will use both forms with different probabilities, $prob(U) : prob(C) = 3 : 4$.

With the existence of probabilistic learners, the innovation with a small functional bias can spread much more easily than in a population with only categorical learners. Fig. 6 shows the diffusion of an innovation with a functional bias of value 2, starting from only one adult, in a small-world and a scale-free network. Under this condition, the innovation has no chance to diffuse at all in a population with all categorical learners. If the learners are all probabilistic learners, diffusion is possible. And consistent to earlier findings, small-world networks ensure more successful diffusions, but require longer time to complete than scale-free networks do.

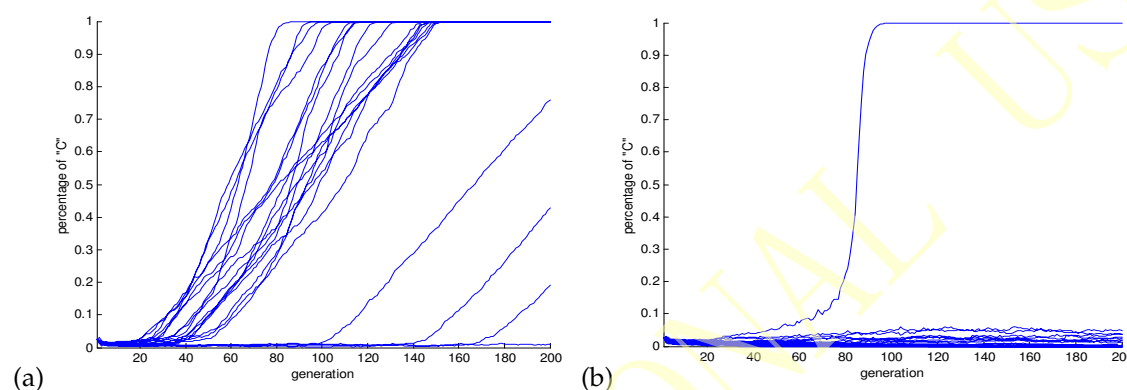


Figure 6: The diffusion dynamics in a population with all probabilistic learners in two networks. ($N=500$, $\langle k \rangle=20$, $\beta=2$, $l=1$). (a) small-world network; (b) scale-free network.

We compare the effect of probabilistic learners under conditions of different functional biases, shown in Fig. 7. In a small-world network, if all learners are categorical, the threshold of functional bias for successful diffusion is 13, but with 10% of probabilistic learner, the threshold drops to 5. If half of the population is probabilistic learners, then an innovation with functional bias of 2 will have more than 50% of chance to diffuse successfully. If the whole population is probabilistic learners, a functional advantage of 1.3 will allow 99% successful diffusion. Similar phenomena can be observed in scale-free networks.

From the simulation results, we suggest that it is because of the existence of probabilistic learners that language change is so frequent; many innovations can successfully spread as long as they have a small functional bias to replace the original norm.

6 Effect of different population size

Nettle (1999b) suggests using simulation results from his model that a larger community requires longer time for changes to complete, and thus fewer changes will occur, which seems to support the data of the linguistic diversity in the world. In his model, however,

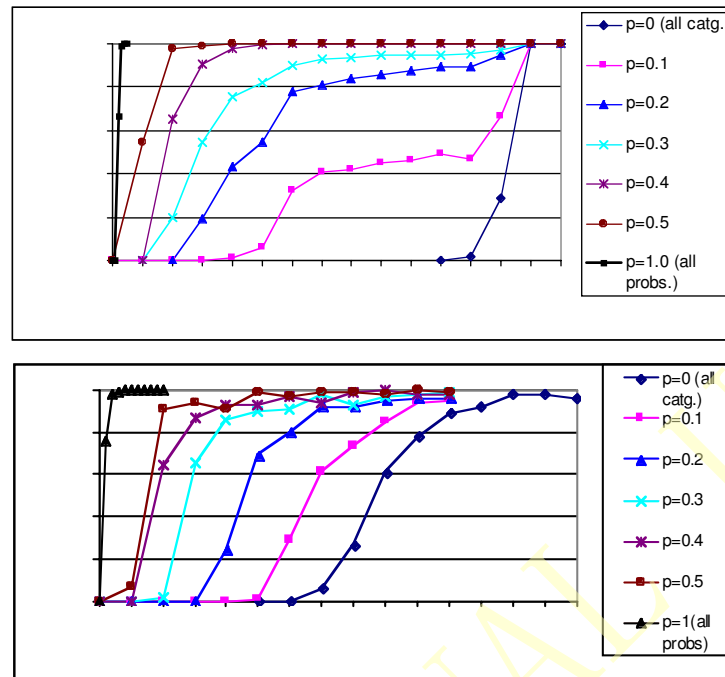


Figure 7: Probability of successful diffusion in populations with different proportions of probabilistic learners, under different functional biases. ($N=500$, $\langle k \rangle=20$, $l=1$). Upper panel: small-world network; lower panel: scale-free network.

as mentioned earlier, the social network is a kind of weighted regular network. Here we compare regular networks with other three types of networks using the model with 50% probabilistic learners in the population. As shown in Fig. 8, in regular networks, the diffusion increases almost linearly with the increase of population size, similar to what Nettle's model suggests. However, the other three types of networks show different results. There is little increase in the diffusion time, compared to the regular network.

One possible explanation for the differences between regular and other networks is that the latter three types of networks have short average path length independent of the network size, as shown in Fig. 9. In other words, the increase of the network size does not lead to the increase of the distance of any two nodes in these networks, and therefore the time for diffusion to take place is hardly affected. Dall'Asta et al. (2006) report a similar discussion in their model of the emergence of linguistic convention, in which "the small-world property holds when the diameter of the network grows slowly, i.e. logarithmically or slower, with its size N . This ensures that every part of the network is rapidly reachable from any other part, in contrast to what happens in regular lattices" (p10).

Wichmann et al. (ms) discuss the rate of language change with respect to population size using a computer model that differentiates internal change, diffusion and shift in agents. They show that the rate of change is not correlated with the population size in

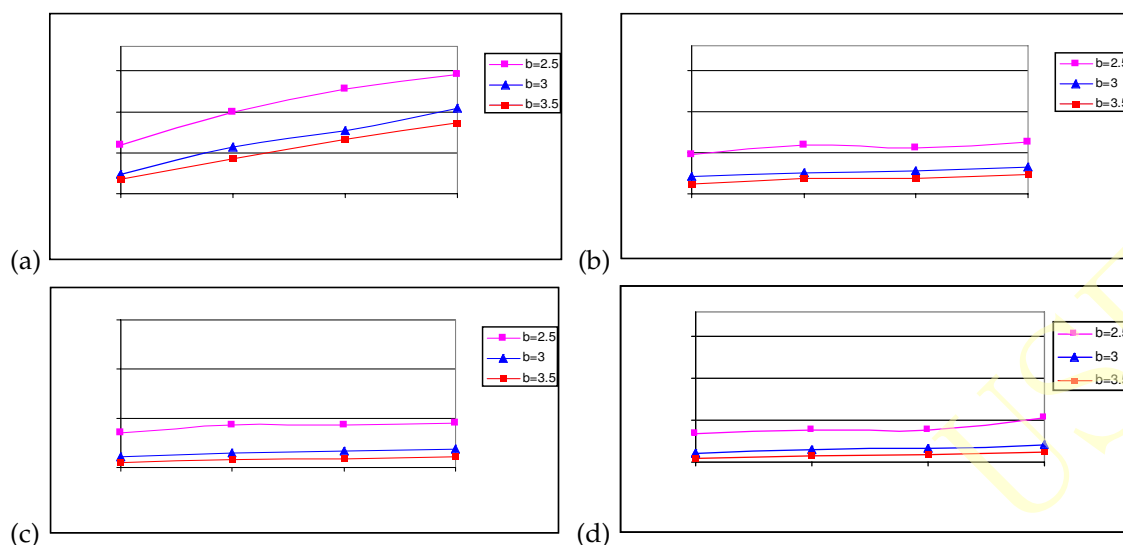


Figure 8: The relation between population size and the rate of change in four types of networks ($N=500$, $\langle k \rangle=20$, $I=1$, 50% probabilistic learners). (a) regular network; (b) small-world network; (c) random network; (d) scale-free network.

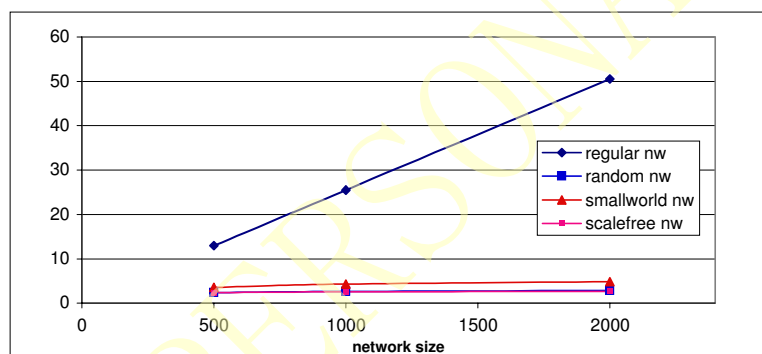


Figure 9: The average path length with respect to different network sizes in the four types of networks.

most cases, except under the condition when the probability for diffusion is high and the diffusion is globally based. Nettle (1999b)'s claim that smaller languages change faster may be valid under some special conditions.

7 Discussion for the models of language change and future works

We have presented a computer model to study language change. The model simulates language change as a process of innovation diffusion. We have examined four typical types of networks and their effect on the diffusion dynamics. When the functional bias is

sufficiently high, innovations always spread in a linear manner in *regular* and *small-world* networks, but diffuse quickly in a sharp S-curve in *random* and *scale-free* networks. The success rate of diffusion is higher in regular and small-world networks than in random and scale-free networks.

These two types of dynamics lead to questions for both empirical and modeling studies. On the one hand, the model raises the question for empirical studies on the relation of language change and social networks, to explain historical data: did social networks at different historical periods differ a lot? So far there have been few data of large-scale social networks with linguistic behavioral data. This question is of particular interest in studies of the various situations in language contact. More systematic analysis of the social networks in those contact situations may provide useful data for network modeling to build upon.

On the other hand, it is pressing to find out which type of networks is more appropriate for representing the real population structure. Which type of network is more realistic, the small-world or the scale-free network, or a new type of network model that can incorporate both the small-world and scale-free, and other features? The model needs to accommodate the presence of a large proportion of local regular connections for the majority of agents, the presence of some hubs in the network, the age structure, and social class distinctions that have often shown significant effect in language change. Newman and Park (2003) discuss various features that social network differ from other types of networks. For instance, social networks have high clustering and transitivity, and exhibit assortative mixing. Several proposals have been reported for modeling, such as Newman and Girvan (2003)'s model of community structure. Schnegg (2006) proposes a model that considers the reciprocity characteristics in human interactions, i.e. the tendency to give to those from whom one has received in the past, based on the B-A scale-free growing model, to account for the significantly lower scaling exponents in social networks shown in ethnographic data. A model proposed by Schwämmle (2005) for language competition, which takes into account the ageing and reproduction in a changing population, may be considered with modification. Other network formation models proposed in economics may also be candidates for further exploration (Jackson 2004).

Our model suggests a new answer to the threshold problem of language change. While functional and social selection may account for the successful diffusion of an innovation to replace the old norm, these conditions may not be as stringent as early models suggest (Nettle 1999b). Our model shows that as long as the population contains a small number of statistical learners who can learn and use both linguistic variants statistically according to the impact of these variants in the input, there is a very high probability for a linguistic innovation with only a small functional advantage to diffuse. This may in part explain why language changes are so prevalent.

The current model is very simplistic regarding the representation of the object of change, which is only in the form of an abstract innovation. When more realistic linguistic features are taken into account in the model, it can be used to simulate different situations of change, such as lexical diffusion, chained change, and different conditions of

language contact. The computer modeling studies on language competition (Schulze and Stauffer 2006) developed in recent years will also be a fruitful area to explore by taking into account more realistic linguistic representations and social considerations discussed above, to address questions such as the rate of language change, linguistic diversity, and so on.

Acknowledgments

We thank Prof. Dietrich Stauffer for the encouragement of publishing this article. An early version of this paper was presented in the Conference on Language Evolution in Leipzig in 2004. This work has been supported in part by grants from 1224/02H and 1127/04H awarded by the Research Grants Council of Hong Kong. The first author would like to thank Santa Fe Institute for the support during July-August 2007 at the time for the writing of this paper.

References

- [1] Barabási, Albert-László 2002. *Linked: the New Science of Networks*. Cambridge, Mass.: Perseus Pub.
- [2] Barabási, Albert-László and Réka Albert. 1999. Emergence of scaling in random networks. *Science*, 286: 509-512.
- [3] Chen, Mathew Y-C. and William S-Y. Wang. 1975. Sound change: actuation and implementation. *Language*, 51(1):255-281.
- [4] Croft, William. 1990/2003. *Typology and universals*. Cambridge University Press. Cambridge, England.
- [5] Croft, William. 2000. *Explaining Language Change: an Evolutionary Approach*. Harlow, England; New York: Longman.
- [6] Dall'Asta, L., Baronchelli, A., Barrat, A., and Loreto, V. 2006. Non-equilibrium dynamics of language games on complex networks. *Physical Review E*, 74:036105.
- [7] Dall'Asta, L. and Baronchelli, A. 2006. Microscopic activity patterns in the Naming Game. *Journal of Physics A: Mathematical and General*, 39(48):14851-14867.
- [8] de Bot, Kees and Saskia Stoessel. 2002. Special issue on Language Change and Social Networks, *International Journal of the Sociology of Language*, no. 153.
- [9] Eckert, Penelope. 2000. *Linguistic Variation as Social Practice: the Linguistic Construction of Identity in Belten High*. Oxford ; Malden, Mass.: Blackwell Publishers.
- [10] Gong, Tao. 2007. *Language Evolution from a Simulation Perspective: On the Coevolution of Compositionality and Regularity*. Monograph 4 in the series *Frontiers in Linguistics*. Academia Sinica: Institute of Linguistics (in press).
- [11] Hudson Kam, C. L. and Elissa Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1:151-195.
- [12] Jackson, Matthew O. 2004. A survey of models of network formation: Stability and efficiency. in *Group Formation in Economics; Networks, Clubs and Coalitions*, edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press, Cambridge U.K.

- [13] Kalampokis, A., Kosmidis, K., and Argyrakis, P. 2007. Evolution of vocabulary on scale-free and random networks. *Physica A: Statistical Mechanics and its Applications*, 379(2):665–671.
- [14] Ke, Jin-Yun. 2004. Self-organization and Language Evolution: System, Population and Individual. unpublished PhD dissertation, Hong Kong: City University of Hong Kong.
- [15] Kirby, Simon. 1999. *Function, Selection and Innateness: The Emergence of Language Universals*. Oxford University Press, New York.
- [16] Labov, William. 2001. *Principles of Linguistic Change (II): social factors*. Oxford, UK: Blackwell.
- [17] Milroy, Lesley. 1980/1987. *Language and Social Networks*. Oxford: B. Blackwell.
- [18] Nettle, Daniel. 1999a. Using social impact theory to simulate language change, *Lingua*, 108:95-117.
- [19] Nettle, Daniel. 1999b. *Linguistic Diversity*. Oxford: Oxford University Press.
- [20] Newman, M. and J. Park. 2003. Why social networks are different from other types of networks. *Physical Review E*, 68:036122.
- [21] Newman M. and Girvan, M. 2003. Mixing patterns and community structure in networks, in *Statistical Mechanics of Complex Networks*, R. Pastor-Satorras, J. Rubi, and A. Diaz-Guilera (eds.), Berlin: Springer.
- [22] Niyogi, P. 2006. *The Computational Nature of Language Learning and Evolution*. MIT Press.
- [23] Nowak, Andrzej, Jacek Szamrej and Bibb Latané. 1990. From private attitude to public opinion: A dynamical theory of social impact. *Psychological Review*, 97:362–376.
- [24] Parisi, D. and Mirolli, M. 2007. The emergence of language: How to simulate it. in *Emergence of communication and language*, C. Lyon, C. L. Nehaniv, and A. Cangelosi (eds.), London: Springer-Verlag, 269–285.
- [25] Schneck, M. 2006. Reciprocity and the emergence of power laws in social networks. *International Journal of Modern Physics C* 17:1067-1076.
- [26] Schulze, Christian and Dietrich Stauffer. 2006. Recent developments in computer simulations of language competition. *Computing in Science and Engineering*, 8.3: 60-67.
- [27] Schwämmle, Veit. 2005. Simulation for competition of languages with an ageing sexual population. *International Journal of Modern Physics C*, 16(10):1519-1526.
- [28] Shen, Zhong-Wei. 1997. Exploring the dynamics aspect of sound change. *Journal of Chinese Linguistics*, monograph.
- [29] Wang, W.S-Y., J. Ke and J.W. Minett. 2004. Computational studies of language evolution. 65-106 in *Computational linguistics and Beyond*. Huang, C.R. and W. Lenders eds. Academia Sinica: Institute of Linguistics.
- [30] Watts, Duncan J. and Steven H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440-442, 1998.
- [31] Weinreich, Uriel, William Labov and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In L. P. Lehmann and Y. Malkiel (eds.) *Directions for Historical Linguistics*. Austin: University of Texas Press.
- [32] Wichmann, Søren, Dietrich Stauffer, Christian Schulze, and Eric W. Holman. (submitted manuscript) Do language change rates depend on population size? (available in <http://arxiv.org/abs/0706.1842>)