Title
Investigating the effects of prestige on the diffusion of linguistic variants

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Tamariz, Monica
Gong, Tao
Jaeger, Gerhard

Publication Date
2011

Peer reviewed

Investigating the effects of prestige on the diffusion of linguistic variants

Mónica Tamariz (monica@ling.ed.ac.uk)
Language Evolution and Computation Research Unit, The University of Edinburgh,
3 Charles Street, Edinburgh, EH8 1AD, UK

Tao Gong (gtojty@gmail.com)
Department of Linguistics, University of Hong Kong
Pokfulam Road, Hong Kong

Gerhard Jäger (gerhard.jaeger@uni-tuebingen.de)
Department of Linguistics, University of Tübingen,
Wilhelmstraße 19, 72074 Tübingen, Germany

Abstract

Language, arguably the cognitive capacity that distinguishes humans, is a dynamic complex adaptive system whose structure and evolution is influenced by a host of factors. This paper takes a population dynamics approach to investigate the diffusion of linguistic variants in populations, focusing on the effect of differential prestige of linguistic variants and of speakers. A novel method that combines computer simulation with mathematical modeling is applied to the specific aim of identifying factors that formally constitute selective pressures on variant diffusion. Of the factors studied, only the intrinsic prestige of variants is found to pose selective pressure, while speakers' prestige merely modulates variant spread.

Keywords: Language evolution, Price equation, Pólya urns, population dynamics.

Introduction

Language, arguably the cognitive capacity that defines humans, is a dynamic complex adaptive system (Beckner *et al*. 2009) whose structure and evolution is influenced by a host of factors. We apply the principles of population genetics (Fisher, 1930; Wright, 1984) to language, and focus on one aspect of language evolution, i.e., the changes in the proportions of linguistic variants in a linguistic community. Such changes are usually achieved via diffusion of various (phonetic, lexical, syntactic, etc.) variants. At the population level, *linguistic diffusion* (henceforth simply "diffusion") can be viewed as the shift in the proportions of linguistic variants used by a population over time (Nakamura *et al.*, 2007). Some well-documented examples of diffusion include the *Great Vowel Shift* in English occurring from the 14th to 16th century (e.g., Wolfe, 1972), other sound changes in modern languages (e.g. Shen, 1997; Labov, 2001), and lexical borrowing among languages (e.g., Bloomfield, 1933; Cheng, 1987).

Examining the mechanisms for diffusion can shed light on questions concerning the cognitive capacities for language and the effects of linguistic or socio-cultural constraints on language evolution (Wolfe, 1972; Pinker & Bloom, 1990; Croft, 2000; Hauser *et al.*, 2002; Tomasello, 2008). Empirical studies from historical linguistics and sociolinguistics have revealed that linguistic, individual learning and socio-cultural factors could all affect diffusion (e.g., Labov, 1994, 2001; Shore, 1995; Fisiak, 1995; Croft, 2000), and recently, mathematical analysis and computer simulation have been used to quantitatively analyze the effects of these factors on diffusion. By quantifying the contact patterns and constraints within or across populations, mathematical analysis helps to predict the influence of these factors (e.g., Nowak *et al.*, 2002; Abram & Strogatz, 2003; Wang *et al.*, 2004; Dall'asta *et al.*, 2006; Kalampokis *et al.*, 2007; Minett & Wang, 2008); by simulating individual behaviors during linguistic interactions, computer modeling helps to trace how interactions among individuals spur the origin of a common set of form-meaning mappings (e.g., Steels, 1995; Ke *et al.*, 2002), how processing constraints lead to linguistic regularities (e.g., Kirby, 2002; Gong *et al.*, 2009), and how social connections affect diffusion within and across communities (e.g., Nettle, 1999; Ke *et al.*, 2008).

Diffusion can be driven by *chance factors*; then the process is called *drift* and follows the neutral model of evolution (Kimura, 1968). It can also be driven by *selection*, in which case, a feature of the linguistic variants (e.g. ease of pronunciation, cognitive salience, social prestige) increases its fitness, i.e., makes it more likely to be used and learned, and therefore diffused among speakers, than alternative variants.

The present study focuses on three factors, namely variant prestige, model bias and transmission error, and seeks to establish whether each of them poses selective pressure on the evolution of linguistic variants. How these factors relate to linguistic behavior is illustrated next: First, linguistic variants possess feature values which may affect the probability with which the variants carrying them are adopted and used, in other words, they affect the variant's *prestige*. A physical feature, such as the ease in perception, for instance, could confer high prestige to a certain variant (Labov, 1994), meaning that it will be more likely to be produced than other variants. Cheng (1987) describes lexical variants borrowed from other languages as high-prestige when they are more salient to hearers than the existing variants for the same meaning. Second, in human society, ordinary people preferentially copy from individuals (models) of higher social, political, or economic status (Labov, 1963; Johnstone, 2010); this is called individual bias or *model bias,* and has been used to explain the spread of certain cultural variants (Boyd & Richerson, 1985; Henrich and McElreath 2003; McElreath and Henrich 2007). Third, *transmission error* happens simply when speakers make production or perception errors. We do not expect transmission error to constitute a selection pressure, due to its random nature (as implemented in our model).

In this paper we apply *the Price equation* (Price, 1970, 1972) to quantitatively identify *the selective pressures* on diffusion, i.e., the factors that successfully lead to the spread of certain type of variants in a population. Though originally proposed in biological terms, this equation can also be applied to any group entity that undergoes socio-cultural transmission (Gardner, 2008; Jäger, 2008). We also employ multi-agent computer model simulations following the *Pólya urn dynamics* (Johnson & Kotz, 1977; Marshall & Olkin, 1993). This model simulates production, perception, and update of variants in linguistic interactions. The results of the simulations are analyzed with the Price equation in order to determine whether the factors studied constitute selective pressures on the cultural diffusion of linguistic variants.

The rest of the paper is organized as follows: first, we describe the computer model and the Price equation; second, we analyze the effects of variant prestige, transmission error and model bias on diffusion; and third we conclude the paper and point out some promising future work deriving from this study.

Methods

The Computer Simulations

The Pólya urn model was first designed to study contagion. In its original implementation, it consists of an urn containing a number of red and green balls; at each timestep, a ball is randomly drawn from the urn, and then returned to it together with a number of balls of the same colour. Such drawing and returning processes repeat themselves, causing the distribution of variant types in the urn to change over time. In our model, an urn is initiated with V tokens, each belonging to a particular type (v_1, v_2, … v_V) and having a quantifiable *feature* x_i (all feature values form F). At a time step, a token v_i is drawn randomly from the urn and returned with another token of the same type.

Our computer model, inspired on the prototypical Pólya urn model, contains N agents (individuals), each denoted by an urn. Variant types and tokens represent linguistic types and tokens. During an interaction between two or more individuals, a token is drawn from one of the urns (the speaker); this corresponds to *production.* Token(s) of the same type are then added to another urn(s) (the listener(s)); this corresponds to *perception* and *update of knowledge*.

Prestige is implemented as follows: when hearing a high-prestige variant, a higher number of tokens are added to the urn – at each time step, a token v_i is drawn randomly from the urn and returned with p_i (the *prestige* of v_i, all prestige values form P) tokens of the same type. *Model bias* is implemented thus – when hearing a token used by the high-status individuals, hearers add more tokens of that type to their urns than if the token is produced by a low-status individual. *Transmission error*, or *mutation*, occurs when a token is returned with some token(s) of different type(s). The probability of mutation is a parameter in the model.

The Price Equation

The aim of the present study is to examine whether a number of factors constitute selection pressures on variant diffusion. A variant may come to dominate in a population for several reasons: it may have intrinsic properties that make it adaptive in its environment and it may therefore be selected for. Alternatively, the random dynamics of evolutionary drift may increase the frequency of the variant. The virtue of the Price equation (Eq. 1), a tool from evolutionary biology, is precisely that it splits change into two components: selection and transmission, allowing us to identify which one is causing evolutionary change.

$$\Delta x = Cov(s_i / s, x_i) + E(\Delta x_i \times s_i / s) \qquad (1)$$

Here, x_i is the feature value of v_i, s_i is the fitness of v_i, s is the average fitness, Δx_i is the feature discrepancy of v_i between time steps, and Δx is the expectation of feature value change.

We apply the equation to trace change in the average value of a quantifiable feature in a population between two consecutive time steps in the computer model and calculate the two terms: 1) The *covariance* between the feature value x_i and the fitness ratio s_i/s measures selection, or evolution caused by fitness differences between different types of variants. Consistent non-zero covariance values over the course of a computer simulation indicates that feature x_i is under selective pressure. 2) The *expectation* of the product of the fitness ratio s_i/s and the feature discrepancy Δx_i measures evolution occurring at transmission, in other words differences between parents and offspring variants. Consistent non-zero expectation values indicate that feature x_i is undergoing transmission error such as mutation.

By quantifying a feature relevant for diffusion and analyzing the average values of the components in the Price equation over many simulation runs, we can identify the selective pressures on this feature.

Identifying Selective Pressures on Linguistic Diffusion

For the sake of simplicity, our models contain only two variants, each characterized by a quantifiable feature F={1, 2}. Example quantifiable features include vowel length, consonant voicing onset time or lexical item recall rate. A simulation has a 100-agent population and 2000 interactions among them (20 interactions per agent on average). We conducted 1,000 simulations in each of four conditions:

- Variant prestige with and without transmission error
- Model bias with and without variant prestige

The result of each simulation consists of a record of the proportions of variants of each type in each urn at each timestep. On this data, we calculate the Covariance and the Expectation terms of the Price equation at 20 sampling points evenly distributed along 2000 interactions. To complement the Price equation, which traces *changes*, rather than *proportions*, of variant types, we also calculate *Prop* (see Equation 2) as the proportion of the majority variant type at each sampling point.

$$Prop(t) = \max_{i=1,2}(proportion(v_i, t)) \qquad (2)$$

By illustrating whether one type of variants gradually diffuses to the population, the average *Prop* of the 1000 simulations helps to evaluate the conclusions drawn from the Price equation.

Variant Prestige with and without Transmission Error

Variant prestige encompasses intrinsic properties of the variant – and *not* of the individuals carrying the variant – that makes them more likely to be adopted by individuals. Henrich and Gil-White (2001)'s study of prestige in cultural transmission do not find an effect of variant prestige on diffusion, although admittedly their focus was on model bias, and variant prestige was implicitly subsumed within that focus. In our simulations, each interaction occurs between two randomly chosen agents. Differential variant prestige is introduced via p_i. For conditions with variant prestige, P={1, 2}; for those without, P={1, 1}. If p_i=2, two tokens of the same type (instead of one) are added to the listener's urn, modeling the enhanced adoption of the high-prestige variant. Transmission error is introduced via mutation; c=0.02 is the probability that an added token becomes a mutant (of the other type). Figure 1 shows the simulation results in these conditions.

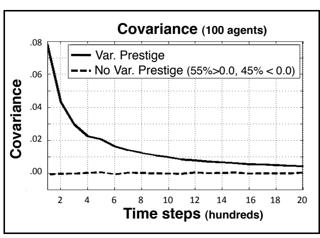

Figure 1(a)

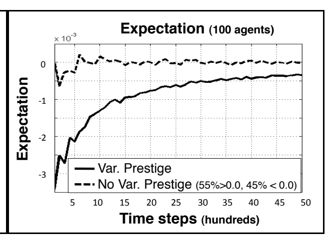

Figure 1 (b)

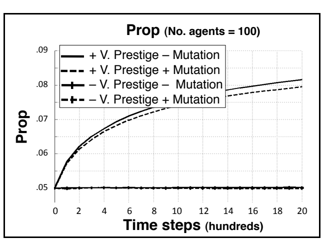

Figure 1 (c)

Figures 1(a) and 1(b) respectively show the covariance without transmission error and the expectation with transmission error. For covariance, with variant prestige, it becomes consistently positive; otherwise, it fluctuates around 0.0 (the proportions of the values above, below, and equal 0.0 are shown in the legend of Figure 1(a)). The gradual decrease in the absolute value of covariance is due to the increase in the total number of variants, which reduces the effect of a small number of changed variants in each interaction. The positivity of covariance clearly indicates that variant prestige is a selective pressure on diffusion. For expectation, with variant prestige, it becomes consistently negative; otherwise, it fluctuates around 0.0. This result suggests that transmission error can reduce the selective pressure of variant prestige. However, due to the low mutation rate, this effect is smaller than that of variant prestige.

Figure 1(c) shows the *Prop* in these conditions. With variant prestige, v_2, with the higher prestige value, becomes the majority type and its *Prop* gradually reaches a high level (above 0.8)[1]; without variant prestige, either type can become the majority type, but the small bias towards either type (due to random factors in early interactions) cannot be further amplified in later interactions, so *Prop* remains around 0.5. This result confirms the selective pressure of variant prestige, consistent with the conclusion drawn from the Price equation. Figure 1(c) also shows the *Prop* in the conditions with transmission error (the dotted lines). With variant prestige, the *Prop* with transmission error is lower than that without, showing that transmission error reduces the selective pressure of variant prestige; without variant prestige, the *Prop* with and without transmission error are similarly low, around 0.5, showing that transmission error alone has no significant effect on diffusion. This result also confirms the conclusion drawn from the Price equation.

The mathematical analyses based on the Price equation applied to the simulation results using the diffusion model formally show that variant prestige is indeed a selective pressure on diffusion and transmission error can reduce such pressure, but transmission error alone fails to consistently drive the diffusion.

The consistent positivity (or negativity) of the covariance based on variant feature identifies selective pressures on diffusion. In the following sections, therefore, we focus on covariance, and leave aside conditions without transmission.

Model Bias with and without Variant Prestige

Model bias reflects the phenomenon that members in a community tend to copy the variants, regardless of the actual forms, from certain individuals. It has been claimed that such bias could enhance the benefits of cultural transmission (Henrich & Gil-White, 2001). We analyze two types of model bias.

The first type involves *a single high-status agent*. Here, a single agent has a bias value of 2, and the other 99 agents' bias value is 1. Variant prestige and model bias take effect jointly during interactions. Without variant prestige, when the high-status agent speaks, the hearer adds 2 tokens of the

[1] *Prop* never reaches 100%, because the tokens of the other type are not removed.

produced type; when another agent speaks, the hearer adds 1 token of the produced type. With variant prestige, when the high-status agent produces a token of the prestigious type, the hearer adds 4 tokens of that type; when it produces a token of the other type, the hearer adds 2 tokens of that type; and when another agent speaks, the update is the same as in the condition with only variant prestige.

Figure 2 shows the results under this type of model bias. Without variant prestige, the *Covariance* fluctuates around 0.0; otherwise, it is consistently positive (see Figure 2(a)). These results show that the first type of model bias alone fails to exert a selective pressure; it has to take effect together with variant prestige. This conclusion can be confirmed by the *Prop* in Figure 2(b): without variant prestige, the *Prop* remains slightly higher than 0.5; otherwise, it approaches 1.0.

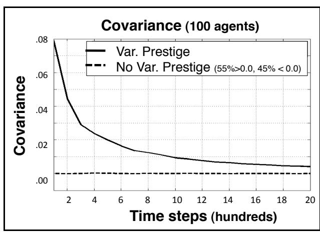

Figure 2 (a)

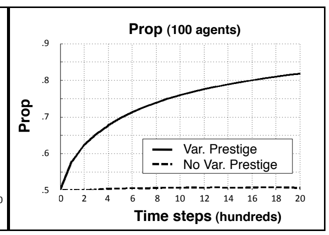

Figure 2 (b)

The second type of model bias concerns multiple individuals. The bias towards an individual is defined as the probability for this individual to participate in interactions, and all probabilities follow a normalized, power-law distribution (Newman, 2005). This implementation is inspired on empirical data (Newman, 2003), and has been adopted in previous work (e.g., Gong *et al.*, 2008). In this paper, we only consider the power-law distributions whose λ values are 0.0, 1.0, 1.5, 2.0, 2.5, and 3.0[2].

Figure 3 shows the results under the second type of model bias. Similar to the first type, the second type of model bias alone fails to exert a selective pressure; it has to take effect together with variant prestige. This conclusion is shown by comparing the *Covariance* in Figures 3(a) and 3(b), and confirmed by the *Prop* in Figures 3(c) and 3(d).
In addition, the *Prop* in the conditions with variant prestige seems correlated with the λ values (see Figure 3(d)). To illustrate such correlation, we define *MaxRange* (see Equation 3) as the maximum changing range of *Prop*:

$$MaxRange = \max_{t \in [1,2000]} (\mathrm{Prop}(t) - \mathrm{Prop}(0)) \qquad (3)$$

Figure 3(e) compares the *MaxRange* in the conditions with and without variant prestige. With variant prestige, the *MaxRange* increases along with the λ. With the increase in λ, some agents are more biased, and they can take part in more interactions than others. Due to variant prestige, the variant bias towards the prestigious variants in these biased agents will increase, quickly spread to others, and get further amplified after many interactions. However, without variant prestige, the variant bias in the biased agents remains small and cannot be amplified enough to increase substantially the proportion of one variant type. Therefore, the correlation between the *MaxRange* and the λ becomes less explicit.

[2] If λ=0.0, all agents have the same degree of bias, which resembles the case of random interaction. If $0.0 < \lambda < 1.0$, the bias values are sensitive to the population size and are, therefore, excluded.

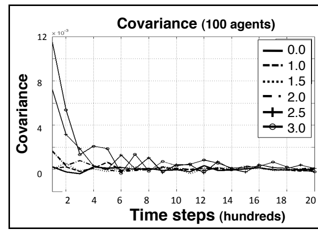

Figure 3 (a)

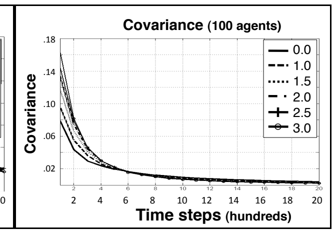

Figure 3 (b)

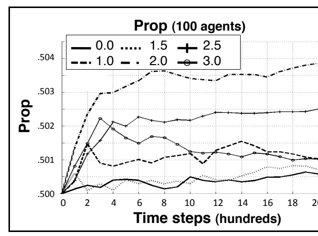

Figure 3 (c)

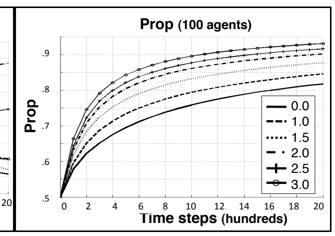

Figure 3 (d)

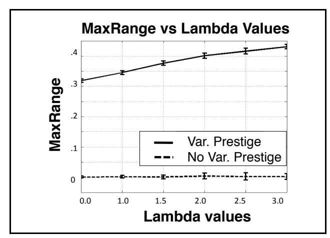

Figure 3 (e)

The power-law distributed model bias reflects the omnipresent scaling law in social and cognitive domains (Kello *et al.*, 2010). Our simulations show that in order for such a scaling characteristic to significantly affect diffusion, variant prestige is necessary. In addition, the correlation between the degree of diffusion and λ values in our study is different from that shown in other studies. For example, Gong *et al.* (2008) observe a threshold λ value (around 1.0), below or above which the spread of linguistic knowledge is less efficient – but our diffusion model differs from that of Gong et al. (2008) in that they explicitly modeled lexical and syntactic information. This different performance indicates that different types of linguistic knowledge may follow different diffusion trajectories in the population.

Discussions, Conclusions and Future work

Our study demonstrates that, of the factors studied, only variant prestige explicitly driving the spread of variants with higher prestige values in the population, whereas other individual learning or socio-cultural factors, such as transmission error or model bias, can take effect only if variant prestige is involved. As shown in our study, transmission error simply introduces noise in the effect of variant prestige, and model bias does not pose a selection pressure. However, if variant prestige is also present, the strength of selection for the high-prestige variant can be modulated by the distribution of individual status in the population.

Our findings indicate that external, domain-general factors, such as individual status, must take effect via intrinsic, domain-specific factors, such as variant prestige. In linguistics, this finding also alerts us not to exaggerate the effect of language-external factors and inspires us to re-evaluate conclusions in previous studies (e.g., Henrich & Gil-White, 2001). Meanwhile, the Pólya urn dynamics is a general transmission framework not specific to linguistic communication, and the simulation results are less dependent on population size, variant number, or interaction number. Additionally, the Price equation provides a concise description of evolutionary processes that abstracts away from the specific properties of biological evolution (Jäger, 2008; Gardner, 2008). These aspects make this finding also instructive to other phenomena that involve socio-cultural transmission.

Computer simulation and mathematical analysis jointly establish a theoretical platform for linguistic research (Loreto & Steels, 2007). Our work exemplifies how these two approaches assist each other to explore the target question. The conclusions drawn from the Price equation are difficult to prove purely mathematically, but they are nicely assessed by the proportions of the majority variant type in the simulations. The simulations can further examine the complementary roles of individual learning and socio-cultural factors in diffusion. When variants have differential variant prestige, transmission error delays the diffusion process and helps to preserve the tokens of less prestigious type; model bias accelerates the diffusion by spreading and amplifying the bias towards prestigious variants, and there is a correlation between the degree of diffusion and that of model bias.

Finally, some aspects of this study can be modified in order to explore further questions on the diffusion of linguistic and other cultural variants. Among possible manipulations we find: changing the structure of the social network, the population structure over time – for instance adding generation turnover involving death of agents and birth of new agents or implementing frequency-dependent prestige. These modifications and others are easily feasible within the combination of computer simulation based on the Pólya urn model and mathematical analysis using the Price equation presented here.

Acknowledgments

Tamariz is supported by the UK AHRC Grant AH/F017677/1. Gong is supported by the Society of the Scholars in the Humanities in the University of Hong Kong. The authors thank Bernard Comrie, Jelle Zuidema, Andy Gardner and two anonymous reviewers for the comments on this work.

References

Abram, D.M. & Strogatz, S.H. (2003). Modeling the dynamics of language death. *Nature,* 424, 900.

Beckner, C., Blythe, R.' Bybee' J., Christiansen' M.H., Croft' W., Ellis' N.C., Holland, J., Ke, J., Larsen-Freeman, D. & Schoenemann, T. (2009) Language is a Complex Adaptive System: Position paper. *Language Learning*, 59, S1 .

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Boyd, R. & Richerson, P.J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Cheng, R.L. (1987). Borrowing and internal development in lexical change. *Journal of Chinese Linguistics,* 15, 105-131.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow, UK: Longman.

Dall'Asta, L., Baronchelli, A., Barrat, A. & Loreto, V. (2006). Non-equilibrium dynamics of language games on complex networks. *Physics Review E,* 74, 036105.

Fisher, R.A. (1930) *The genetical theory of natural selection*. Clarendon.

Fisiak, J. (Ed.) (1995). *Linguistic change under contact conditions*. New York: M. de Gruyter.

Gardner, A. (2008). The Price equation. *Current Biology,* 18, R198-R202.

Gong, T., Minett, J.W. & Wang, W.S-Y. (2008). Exploring social structure effect on language evolution based on a computational model. *Connection Science,* 20(2), 135-153.

Gong, T., Minett, J.W. & Wang, W.S-Y. (2009). A simulation study on word order bias. *Interaction Studies,* 10(1), 51-75.

Hauser, M.D., Chomsky, N. & Fitch, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science,* 298, 1569-1579.

Henrich, J. & Gil-White, F. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior,* 22, 165-196.

Henrich, J. & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology, 12*, 123-135.

Jäger, G. (2008). Language evolution and George Price's "General theory of selection". In R. Cooper & R. Kempson (Eds.), *Language in flux: Dialogue coordination, language variation, change and evolution* (pp. 53-82). London: College Publications.

Johnson, N.L. & Kotz, S. (1977). *Urn models and their applications*. New York: Wiley.

Johnstone, J. (2010). Locating language in identity. In C. Llamas & D. Watt (Eds.), *Language and identity* (pp. 18-28). Edinburgh, UK: Edinburgh University Press.

Kalampokis, A., Kosmidis, K. & Argyrakis, P. (2007). Evolution of vocabulary on scale-free and random networks. *Physica A,* 379, 665-671.

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

Ke, J-Y., Gong, T. & Wang, W.S-Y. (2008). Language change and social networks. *Communications in Computational Physics,* 3, 935-949.

Kello, C.T., Brown, G.D.A., Ferrer-I-Cancho, R., Holden, J.G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G.C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences,* 14(5), 223-232.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129), 624-626.

Kirby, S. (2002). Learning, bottleneck and the evolution of recursive syntax. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 173-205). Cambridge, MA: Cambridge University Press.

Labov, W. (1963). The social motivation of a sound change. *Word,* 19, 237-309.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford, UK: Basil Blackwell.

Labov, W. (2001). *Principles of linguistic change: Social factors*. Oxford, UK: Blackwell.

Loreto, V. & Steels, L. (2007). Social dynamics: Emergence of language. *Nature Physics,* 3, 758-760.

Marshall, A.W. & Olkin, I. (1993). Bivariate life distributions from Pólya urn model for contagion. *Journal of Applied Probability,* 30, 497-508.

McElreath, R. & Henrich, J. (2007). Dual inheritance theory: the evolution of human cultural capacities and cultural evolution. In R. Dunbar and L. Barrett (Eds.) *Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Minett, J.W. & Wang, W.S-Y. (2008). Modeling endangered languages: The effects of bilingualism and social structure. *Lingua,* 118, 19-45.

Nettle, D. (1999). Using social impact theory to simulate language change. *Lingua,* 108, 95-117.

Newman, M.E.J. (2003). The structure and function of complex networks. *SIAM Review,* 45, 167-256.

Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics,* 46, 323–351.

Nowak, M.A., Komarova, N.L. & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature,* 417, 611-617.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences,* 13, 707-784.

Price, G.R. (1970). Selection and covariance. *Nature,* 227, 520-521.

Price, G.R. (1972), Ann Hum Genet 35: 485-490.

Shen, Z-W. (1997). *Exploring the Dynamic Aspect of Sound Change*. Journal of Chinese Linguistics Monograph, 11: Project on Linguistic Analysis. University of California Press: Berkely.

Shore, C.M. (1995). *Individual differences in language development*. Thousand Oaks, CA: Sage Publishing Co.

Steels, L. (1995). A self-organizing spatial vocabulary. *Artificial Life,* 2, 319-332.

Tomasello, M. (2008). *The origins of human communication*. New York: MIT Press.

Wang, W.S-Y., Ke, J-Y. & Minett, J.W. (2004). Computational studies of language evolution. In C-R. Huang & W. Lenders (Eds.), *Computational linguistics and beyond* (pp. 65-108). Taipei: Institute of Linguistics, Academia Sinica.

Wright, S. (1984). Evolution and the Genetics of Populations: Genetics and Biometric Foundations (4 vol), Chicago: University of Chicago Press.

Wolfe, P.M. (1972). *Linguistic Change and the Great Vowel Shift in English*. Berkeley: University of California Press.