International Journal of
Mathematical Modelling & Computations
Vol. 01, No. 02, April 2011, 87- 99

A Model for Evolutionary Dynamics of Words in a Language

A. Yadav[a], J. Dash[b] , M. Padhee[b], and S. Bhattacharya[c*]

[a]*Department of Computer Science & Engineering, Faculty of Science & Technology, ICFAI University, Rajawala Road, Selaqui, Dehradun 248 197 Uttarakhand India .*
[b]*Department of Biotechnology, Faculty of Science & Technology, ICFAI University, Rajawala Road, Selaqui, Dehradun 248 197 Uttarakhand India.*
[c]*Department of Mathematics, Faculty of Science & Technology, ICFAI University, Rajawala Road, Selaqui, Dehradun 248 197 Uttarakhand India .*

Abstract. Human language, over its evolutionary history, has emerged as one of the fundamental defining characteristic of the modern man. However, this milestone evolutionary process through natural selection has not left any 'linguistic fossils' that may enable us to trace back the actual course of development of language and its establishment in human societies. Lacking analytical tools to fathom the critical essentials of evolutionary mechanism of cultural transmission, we seek the recourse of simulation study as another useful method of enquiry into the evolutionary trajectory of language.

In this paper we use a toy model to understand an interesting feature of language evolution, namely, the scenario in which words gets fixed in a population of language users. We obtain simulation for the replicator dynamics that characterise the time rate of change of various words in the given language, using genetic algorithm to simulate the dynamics. We infer that two of the prime determinants for the establishment of a word within a linguistic population are its consonance with the grammar and its communicative efficiency.

Keywords: Language, Complex adaptive system, Evolution, Evolutionary game theory Genetic algorithm, Simulation.

Index to information contained in this paper

1. Introduction

The evolution of language obviously ranks as one of the most significant events in the course of natural history of human evolution. Being a defining moment in the evolutionary history of modern humans, language facilitates a multitude of possibility of transmitting non-genetic, cultural material from one generation to the next in the course of Darwinian evolution [1, 2, 3, 4, 5, 6, 44].

Communication through vocalization or symbolization of phonemes that make

*Corresponding author. Email: sudeeptobhattacharya@gmail.com

component-wise sense too, established itself as an evolutionary landmark in the making of modern humans. Such an ability to communicate syntactically within as well as outside of one's social group thus emerged as a definition of modern human's quintessential social identity. This feature also marked a new form of nongenetic transmission of information from one generation to the next [7, 8, 9, 10, 11, 12, 13, 14, 15].

Language, together with coding cultural information, codes information on its own structure. Thus, during the evolutionary history, a given language provides coded information that influences its own survival via feedback loops during the various evolutionary phases it passes through. Language therefore could be viewed as a consequence of interaction between three complex adaptive systems: biological evolution, learning and culture. Under this scenario, language itself gets characterized as a complex adaptive system [16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27].

Evolutionary emergence of human language has hardly left any tangible, data-intensive archaeological record. Also, since language should be viewed as a complex system following a dynamical evolutionary trajectory rather than as a single, monolithic human behaviour, any insight into the evolutionary history of language requires the behaviour of this complex dynamical system to be broken down into its constituent components, with each component possibly following an independent, distinct evolutionary path [28]. This fact has therefore put a constraint on the kind of studies that may be deemed plausible in the domain of evolution of language, and consequently there does not exist any exact analytical method as a roadmap to pursue problems relating to evolutionary emergence in this domain.

In absence of a general and structured analytical framework, any attempt to understand and analyze the trajectory of evolution and emergence of a language in a linguistic community of users of that language must take recourse to certain broadly speculative and conjectural theories. One of the potent tools for furthering explorations and studies in the above kind of domains could be the approach through simulations, which are essentially hypotheses and theories expressed as computer programs. Although such studies using simulations often tend to simplify (even oversimplify) the essential complexity involved in the evolutionary system, could nevertheless prove to be very valuable in order to obtain an insight into the actual phenomena. The results of simulations, often depicted graphically, are the conclusions as well as predictions about the system of interest that has been described by the program and has been simulated by the computer [29, 30, 31, 32]. In the present paper, we undertake an exploration of self replicating rules and structures of language, viewed as a complex system, using genetic algorithm.

For our modelling in the present work, we assume a population of individual agents who communicate with each other randomly, and thus are users of a language. Following the Five Graces group, we assume that: (i) the linguistic system consists of multiple agents interacting with each other as well as with the environment; (ii) the system is adaptive - individual agent's behaviour is based on their past interactions and current and past interactions together feed back into their future behaviour; (iii) an agent's behaviour is the consequence of competing factors from a wide range of environmental constraints [25]. These interacting agents in our model comprise the speech community, in which the agents themselves control all their interactions without the influence of any controller, and hence comprise an autonomous community [26].

In the present paper we propose a toy model for language evolution, and choose the methodology of computational simulation for evolution of language, assuming that as a complex adaptive system, language is continually evolving in cultural time and social space. We assume that a linguistic system is described by the linguistic structures used in a speech community, represented by words that belong to the language defined by its grammar , which essentially supplies the rewrite rules for construction of linguistic structures [26, 27]. Words, essentially constructed as strings of symbols from the alphabet used for describing the language, are taken as proxy for language users and are the agents in our modelling.

In a typical game theoretic setting, the agents play an evolutionary game iterated through generations where each agent, being rational, has the objective of optimizing its payoff [28, 29, 30, 31, 32].The interaction among the agents as well as those of the agents and the environment capture the evolutionary game. In the above perspective, the distinction between the agents (words) and the strategies played by the words could be blurred without loss of generality, and the agents are themselves the strategies in the game. In every generation, the above interactions result in the linguistic structures - words - getting replicated, either in totality or with some alterations. Hence, in the model, the words are the agents, the 'lingueme', and thus are the replicators [33, 34].

An evolutionary dynamics for a given system requires replication, variation and selection mechanisms for the agents to be identified and defined, in order to be modelled. These prerequisites for the agent dynamics in the present work are described in the following paragraphs.

The emergence of collections of replicators of various types during the evolutionary game played by the agents of the linguistic system causes a high degree of variation in the replication of linguistic forms in the language [27, 33, 35, 36]. We assume that the replication of words occurs through linguistic creativity, reanalysis and language contact [34]. We further assume that selection takes place as a manifestation of social preference of one word over another. This preferential bias supplies the linguistic forms with selection for learnability and frequency of use, and selection against ambiguity in meaning. Hence there exists a selection pressure on the evolutionary dynamics of the replicators.

We model the variations necessary for the evolutionary dynamics of the language system by using a stochastic selection protocol of the simple genetic algorithm, guided by frequency - manifested by a metric called fitness - dependent mechanism. The protocol used in the paper thus contains the updating mechanism for the population, and is a cultural procedure that mimics biological evolution through natural selection. In the paper, we show that such a fitness dependent selection of agents in the evolutionary game over iterative generations either reinforce and establishes it in the language as a component of the linguistic form, or deletes it from the language.

With an aim to make the paper self-contained, we present below a cursory overview of the essential basics of two key aspects of our work, namely, language and genetic algorithm, even at the cost of redundancy. The presentations draw heavily from the scholarly works cited in the corresponding references in the text.

1.1 *Language Preliminaries*

Following is a synopsis of the standard concept of formal language that we shall adhere to in the paper [44, 45, 46].

For a given alphabet Σ, a string (or a word/ sentence) is a finite sequence of symbols chosen from Σ. If $\Sigma = \{0, 1\}$, then we define the powers of Σ as $\Sigma^1 = \{0, 1\}$, $\Sigma^2 = \{00, 01, 10, 11\}$ and so on. Denoting the union of all such powers of Σ by Σ^*, we have $\Sigma^* = \{0, 1, 00, 01, 10, 11, 000, 001, \dots\}$, which is the collection of all possible sentences that can be generated from the alphabet Σ. Σ^* is a countably infinite set, where all the strings belonging to Σ^* can be enumerated using a lexicographical order.

A set of words (and their concatenations), all of which are chosen from some Σ^* is called a language L. Therefore, if Σ is an alphabet and $L \subseteq \Sigma^*$, then L is a language over Σ.

A context free grammar $G(V, \Sigma, R, S)$ is a 4-tuple where V is the set of variables and forms a finite set, Σ is the given alphabet whose elements are the letters or terminals (every element of the set V/Σ is a nonterminal), R is a finite set of rules representing the recursive definition of the language in the context (each rule is a variable or a string of variables and terminals taken from V and Σ) and $s \in V$ is the start variable that represents the language being defined. A context free grammar G specifying a language $L \subseteq \Sigma^*$ is therefore essentially a rewrite system for strings comprising a finite list of rules over an alphabet Σ.

1.2 *Genetic Algorithm Preliminaries*

Genetic Algorithms (GA) comprise a class of optimization routines that are essentially population-based search mtaheuristics. Initially developed by Holland for addressing problems of adaptive systems in domains that are characterized by both enormous search spaces and objective functions with nonlinearities (multiple local optima), discontinuities, high dimensionality and noise, GA provide a highly efficient search procedure to effectively address the problem of optimization in such 'difficult' domains [47, 48, 49, 50, 51].

GA are a large class of routines, where each member of the class is characterized by the following criteria: (i) a nonzero population of well defined structures agents (structures), (ii) action of the agents in an environment, (iii) evaluation of the performance of each of the agents using a fitness score, (iv) creating new populations using this score as the input and ranking the agents according to their respective scores, (v) selecting the agents with better performance scores, (vi) modifying the selected agents through stochastic genetic operators.

Of a multitude of evolutionary search routines that are current in scholarly literature and get labelled as genetic algorithm, we shall remain confined attention to the application of Simple Genetic Algorithm (SGA) as a tool to study the problem set in this paper [52].

SGA is a triple $\Gamma(\Omega, \Im, g)$, where the components are as below: Ω is the search space comprising agents represented as binary strings which are the candidate solutions for a given optimization problem. $\Im$ is an exogenously defined fitness function and g is the search heuristic acting on a nonzero population of the candidate strings.

The fitness function $\Im$ is an injective map from Ω to $\Re$, and defines the environment for the evolutionary scheme. It evaluates each string x_i, $i = 1, 2, \ldots, |\Omega|$ and declares a fitness score. The heuristic g comprises three stochastic operators: the selection operator ψ , the crossover operator χ and the mutation operator μ.

The operator ψ maps the simplex Λ representing the population of agents at a given generation P_t to the search space Ω:

$$\psi : \Lambda \rightarrow \Omega$$

Being a non explorative operator, the selection operator does not generate any new string in the population.

The crossover operator χ acts on a pair of elements of the search space,

$$\begin{aligned} \chi :& \Omega \times \Omega \rightarrow \Omega \\ & (x,y) \longmapsto \chi(x,y) \quad x,y \in \Omega \end{aligned}$$

The two elements $x, y \in \Omega$ on which χ acts are the parent strings, yielding two offspring strings as a result of the crossover operation. The crossover point is a randomly selected bit position from the interval $[o, l-1]$. The offspring thereby 'inherits' blocks of loci from both the parents, giving rise to the exchange of information between trial solutions. Traditionally, the value of probability of action of the operator χ is significant, ranging between 0.6-0.8.

The mutation operator μ acts on one single string x and changes the binary character at a locus on the string to obtain a different string:

$$\begin{aligned} \mu :& \Omega \rightarrow \Omega \\ & x_i \rightarrow x_j \end{aligned}$$

The probability of action of μ is generally taken to be small, of the order of 10^{-3}.

The two operators χ and μ act independently of one another on the population of strings, producing the mixing Θ of the strings. The heuristic g effectively is then the composition of the selection Ψ and the mixing Θ: $g = \Theta o \Psi$.

The net effect of the heuristic is the creation of new strings in the search space with a spectrum of fitness values, resulting in a very efficient sampling plan. One may thus expect subsets of the search space Ω containing strings of similar profile and sharing a particular set of fitness score (schema) to emerge and evolve in time. The behaviour of theses schema will remain largely conformal with the selection pressures acting on the dynamical, evolutionary language system within the environment [53].

2. Modelling

We consider a language model to study the persistence and fixation of the words in a language on an evolutionary scale, effected by communicative interactions and transmission of 'language genes' through generations of members of the linguistic community, using the language.

In order to address the problem as outlined in Introduction, we have performed a modelling based on evolutionary game theory for the above community using a SGA Γ mentioned above, as the heuristic for obtaining the necessary variation and selection in the interacting agent community. The inputs and outputs to the functions in the following are the binary strings of finite length, that is elements of the set Σ^*.

The model presented in the paper comprises three major components:

(i) allowable agents (words) which are strings of a finite length l, $l \in Z^*$ defined over the binary alphabet $\Sigma = \{0, 1\}$, with each string representing a sentence. Let Z_2 be the set of integers modulo 2 : $Z_2 = \{0, 1\}$. The search space Ω for the SGA in our model may then be represented as

$$\Omega = Z_2 \times Z_2 \times \cdots \times Z_2 \quad \text{(l factors)}$$

It may be noted in passing that $|\Omega| = 2^l$ and the space Ω may be established as an l-manifold through usual chart with a suitable metric defined on it via the fitness function $\Im$. Every locus on any given string x_i, $i = 1, 2, \ldots, \{0,1\}^l$ specifies and controls a particular syntactic characteristic of the language $L(G)$;

(ii) An environment described and defined by the fitness function $\Im$ that evaluates the performance of each string from the search space in terms of a fitness measure which is a real number, while the fitness function itself maps each string to a numerical score $\in \Re$, and is a measure of success of each string in the run of the SGA;

(iii) The evolutionary roadmap described by the SGA, that essentially comprises two operational mechanisms: selective reproduction and continuously adding variability in the populations.

For the purpose of developing a minimalist model, we make some simplifying assumptions (even at the cost of oversimplifications at some points) listed below:

a) The language in discussion is compositional. This means that the meaning of a signal is a function of its parts. Thus, each element of meaning will map onto a particular part of the signal, and therefore each subpart of the entire sentence (string) will be individually mapped onto by the given signal [21].
b) Each individual user (agent) that comprises the linguistic population is equipped with a well-formed syntax (through the grammar of that language) that does not get updated or revised regularly and thus serves as a backdrop for language use.
c) The system comprises an infinite population size of the agents.
d) Words (strings) proxy for the language user agents.
e) The membership of the words in the language is a bivalent function; either a word belongs to the language or does not.
f) The population of agents comprises a dynamical system with the SGA Γ. The heuristic mapping g together with the fitness function $\Im$ provide the update rule that describes the change of the system from one time step to the next. The search space Ω represents the state space that contains all possible linguistic compositions of the well-mixed population of agents The smooth dynamical system may be described by the following one-parameter family of mappings:

$$\phi^t : \Omega \to \Omega, \quad t \in Z$$

thus forming a one-parameter group through the rules

$$\phi^{t+s} = \phi^t \circ \phi^s$$
$$\phi^0 = Id$$

where Id is the identity map [54, 55].

We label the language as $L(G)$ and seek to find out which are the agents (words) that converge in (get fixed in) $L(G)$ after a period of evolution that mimics natural

selection. Thus, the problem for our present study may be written as:
What are the factors that determine if a proportion of population of agents will converge to the target language over discrete, finite steps of time while other competing proportions of the agent community die out? That is, what are the factors that determine if only one of the competing fractions of words in a language will exhibit dynamical stability after a finite number of iterations over discrete cultural time?

To attain this objective, we begin with an initially given non-zero population of language user agents using $L(G)$. Each individual agent is a binary-coded string of length 20 bits ($l = 20$) that is a word formed by catenation on the alphabet Σ and thus essentially represent a context free grammar. Next we evolve this set using Γ, maintaining a randomly selected population of 200 agents in the search space Ω at every generation.

Let this population of agents in the linguistic community play a certain 'language game', repeated over generations, defined by iteration of the heuristic Γ. Let the game be described by $\Theta(P, E, \Pi)$ where $P = \{P_1, P_2, \ldots, P_n\}$ is the set of interacting 'player' agents such that cardinal number of P is n, E the strategy space given by $E = \hat{e}_1 \times \hat{e}_2 \times \cdots \times \hat{e}_n$, $\hat{e}_i$ being the ith pure strategy, and $\Pi = \{\Pi_1, \Pi_2, \ldots, \Pi_n\}$ is the set of payoffs, Π_i being the payoff associated with $\hat{e}_i$. Let the game be repeated in periods of time $t \in \Re$. Assume that the agents are somehow 'hardwired' to play only pure strategies in the game Θ. Thus each strategy $\hat{e}_i$ in this game corresponds to a corner point, that is, a member of the standard basis where the ith coordinate is 1 and the rest are zeros, in the simplex

$$\Lambda = \left\{ \hat{p} = (p_1, p_2, \ldots, p_n)^T \in \Re : p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1 \right\}$$

The fitness function in our model is the mapping $\Im : \Lambda \to Z_+$ which assigns integer scores to the strings, and is a measure of the success of each agent in the game Θ, played by the agents iteratively during the run of Γ, and could be represented as a convex combination of a positive base fitness common to all strategy and the payoff obtained by a particular strategy per play of Θ.

As mentioned earlier, language, being a complex system, evolves in cultural time through availability of variations in the replicators of linguistic forms, that is, the agents in our model. The evolutionary process of such a variation dependent mechanism in agent's selection is captured by the replicator dynamics, which provides the equation of motion for the replicators in this time evolution in the language system. Being representative of general evolutionary dynamics scheme, these equations assume that the proportion of agents using a particular strategy increases at a rate proportional to the fraction of agents using that strategy and the difference between the current fitness score of these agents and the average fitness score of all the agents comprising the population.

Assume that the agents (strategies, words) in our linguistic evolutionary game model comprise a well-mixed population. Let this population be divided into n types x_i that use strategies $\hat{e}_i$ with frequencies σ_i, $i = 1, \ldots, n$. Starting with $t = 0$, game is played iteratively in positive integer periods $t = 1, 2, \ldots$. The state of the population of words at any given t is thus given by the state vector $\hat{\sigma}^t = \left(\sigma_1^t, \sigma_2^t, \ldots, \sigma_n^t\right)^T \in \Lambda$. Let $\Im_i^t$ be the fitness score of the proportion of agents using $\hat{e}_i$ with frequency σ_i^t and $\bar{\Im}^t$ be the average fitness score of the entire population at t. The replicator dynamics then describes the time evolution of this state of population of strategies represented by the corresponding frequencies, and is given by the deterministic ODE on the simplex Λ as

$$\dot{\sigma}_i^t = \sigma_i^t \lfloor \Im_i^t(\hat{\sigma}^t) - \bar{\Im}^t(\hat{\sigma}^t) \rfloor, \quad i = 1, \ldots, n$$

with a dot over denoting differentiation with respect to t [56, 57, 38].

The above equations drive the frequency of a strategy $\hat{e}_i$ to increase, provided its fitness (and therefore payoff) is higher than the average fitness of the population, and hence spread in the population through replicating itself to the next generation, and dwindles otherwise. This fitness dependence provides the selection mechanism in the evolutionary game played by the agents. It must, however, be noted that if a strategy (word) is not represented in the population at one particular generation t, then it cannot appear at any generation after that, and hence, entirely new words may not be innovated in the language under replicator dynamics [57, 58].

For the present work, we define fitness $\Im_i^t$ of an agent as the number of 1's in the agent's odd-numbered loci $\forall i$ at t. Qualitatively in our model, the fitness score would inform how well does a particular word conform to the grammar of the language $L(G)$, and hence, is indicative of the 'success' of the word at generation t to get fixed in $L(G)$. We stipulate a fitness score of 8 as a qualifying threshold for an agent to be a member of $L(G)$, and call such agents 'fit' strategies (strings) in $\Lambda \subset \Omega$ at a period t.

To form the subsequent pools of agents, we apply a selection pressure on the population of fit strings. A concatenation of word may adhere to the syntax prescribed by G, yet may not be effective in communicating the intended signal-object association, that is, may lack in communicative efficiency. A selection pressure will monitor the communicative efficiency of the sentences (words). This setting will also help to additionally bias the selection in favour of fit strings from Λ that belong to the target language $L(G)$. The process is isomorphic to selecting preferentially only those words for usage which are the best (at least near-best) representatives of the target language in terms of serving the dual objective of being faithful in associating the corresponding signal correctly, as well as being consonant with the syntactic demand of the grammar of $L(G)$. New strings (arising out of crossover and mutation from parent strings, hence not entirely new) are created at every generation on running the genetic algorithm, through its heuristic search. Emergence of new agents (offspring) adds to the heterogeneity of the search space (and thus of the selection pool) which in its turn leads to drive the population in directions so as to try and avoid fitness stasis or fitness traps during the evolutionary course [59, 60, 61, 62, 63].

We run the genetic algorithm for 30 generations, and obtain the simulation of time rate of evolution of $\hat{\sigma}^t$ through these generations.

The following routine describes the Γ used in our work, with Ω_i denoting the search space for Γ at the generation t:

1. Initialize Ω_t by choosing 200 strings at random.
2. Set $t = 0$, $t \in N$.
3. Calculate the fitness score for each string $x_i \in \Lambda \subset \Omega_t$, $i \in N$.
   (i) Let $V[x_i] = \Im_i^t$.
   (ii) If $V[x_i] \geqslant 8$ then declare x_i fit string $\bar{x}_i$; else declare x_i unfit.
4. Let σ_i^t be the proportion of fit strings in Ω_t. Let $n(\sigma_t)$ be the number of fit strings.
5. Calculate:
   (i) $\bar{\Im}^t = \frac{\sum V[x_i]}{200}$ average fitness of the entire population.
   (ii) $\dot{\sigma}_i^t = \sigma_i^t \lfloor \Im_i^t(\hat{\sigma}^t) - \bar{\Im}^t(\hat{\sigma}^t) \rfloor$.

6. Reproduce strings in Ω_t:
   (i) Select the best scoring strings from proportion σ_t of Ω_t.
   (ii) Apply a pressure.
   (iii) Allow $\frac{V[\bar{x}_i]}{\Im_i^t}$ offspring.
   (iv) Adjust the population if the number of offspring $\neq 200$:
      a) If number of offspring ¡ 200, then allow highest scoring string in the new population to produce additional offspring.
      b) If number of offspring ¿ 200, then eliminate the lowest scoring strings from the new population.
7. Apply crossover operator χ to the new population with probability P_χ:
   a) Randomly select 2 parent strings.
   b) Randomly select a crossover point $c \in [0, 19]$.
   c) Form 2 offspring by applying χ: crossover the parents immediately after the c$^t h$ locus.
8. Apply mutation operator μ to the new population with probability P_μ:
   a) $\forall$ locus $\forall x_i$, change the symbol to its alternative in $\Sigma = \{0, 1\}$..
9. Let $t = t + 1$
10. If $t < 30$, go to step 3.
11. End.

3. Results and Discussion

The following simulations for the replicator dynamics were obtained by fixing selection pressures at 0.8 (high selection pressure, a strict linguistic community that would admit only those replicators that have high learnability and occur with a high frequency, but have low ambiguity), 0.6 (moderate selection pressure, a slightly more accommodative community) and 0.4 (low selection pressure, a relaxed linguistic community, tends to admit words even with marked degree of ambiguity in meaning, low frequency of occurrence and low learnability) respectively on the population of the agents, with the horizontal axis representing number of generations t against the vertical axis representing $\dot{\sigma}$:

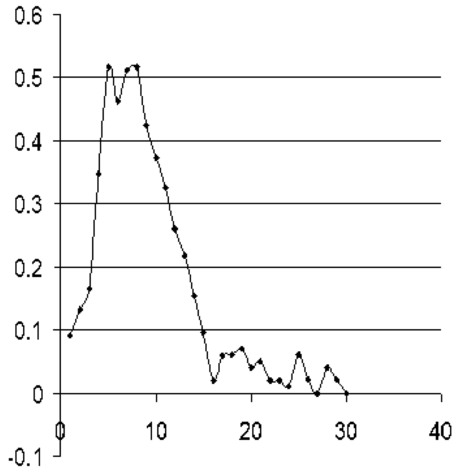

Figure 1. Graph for selection pressure 0.8.

In such iteration as above, the dynamical stability of words in a language over (cultural) time, expressed through a user population converging to it relates to the

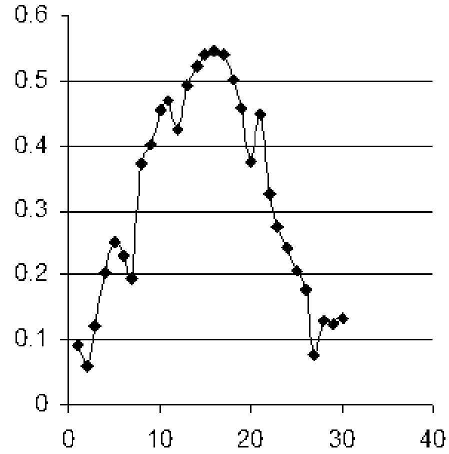

Figure 2. Graph for selection pressure 0.6.

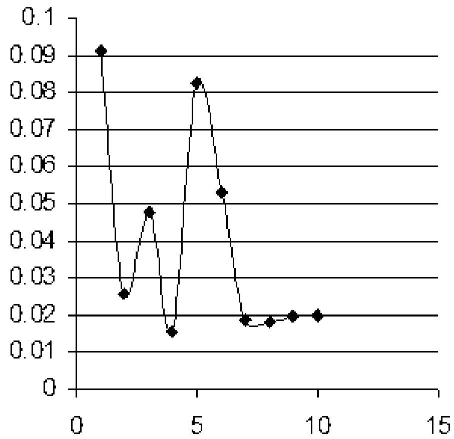

Figure 3. Graph for selection pressure 0.4.

expressivity and thus the communicative efficiency of the language. The emergent scenario as a result of the interactions among the agents both with themselves as well as with the environment through Γ,could be summarized in the following paragraph:

From Figure 1 we observe that for a selection pressure of 0.8, the population of replicators settles on the target language $L(G)$ with time rate of change in the parameter σ becoming zero after approximately 30 generations. We may infer from the scenario, that after 30 generations, the words get fixed in the language, and do not change thereafter. Figure 2 shows that with a selection pressure of 0.6, the population fails to converge to the language even after 30 generations, and in fact shows an increase in rate of change in σ. Thus, at least till 30 generations of communicative interaction, the linguistic ambiguities persist in the population, though with a change in user proportion. Figure 3 graphs the iterations only up to 10 generations, as during the simulation it was observed that even in this instance with a selection pressure of 0.4, the population of agents does not settle on $L(G)$ even up to 30 generations, and in this case too, therefore, linguistic ambiguity persists in the population.

The above observations lead us to infer that for a proportion of linguistic replicators - words or 'linguemes' - to evolve through the process of an adaptive evolutionary optimization mechanism over a period of cultural time and exhibit dynamical stability in the language system, the following two factors become vitally important:

(i) The language to which we seek the population to evolve, the target language $L(G)$ should be rather steeply demanding in regards to its syntactic requirements, reflected through a high selection pressure and indicated in the model by setting a high threshold fitness score. It should be discriminate in accepting only those words as 'proper' words of the language that have a particular phenotypic configuration which endows them a fitness score of at least 8 out of a maximum possible 10. This means that the language should comprise sentences that are almost totally in consonance with its grammar G, and colloquial and derivatives or a marked tendency to admit all and sundry words in a language be best avoided to dynamically evolve and stabilize a population of words to that language;

(ii) It is useful to admit only those words for usage that not only have a fitness score more than or equal to the set threshold value, but also are at least near-best representatives of the target language indicated by setting a high selection pressure. Thus, in order to belong to the language, a word must have a relatively high "communicative efficiency", or a social acceptance. Therefore, a large amount of foreign words in a language, though may be grammatically admissible, does not guarantee the population evolving to use only that language unless the words (or the sentences made from the words) are also near-faithful representative of that language, in terms of being socially popular in usage.

It may also be inferred from the results that in absence of any one of the above two factors, the language under discussion may face a prospect of either never establishing itself in a population of users uniquely or even becoming extinct at least locally.

References

[1] Hurford, J.R. (2003) The language mosaic and its evolution. In Language Evolution: The States of the Art. Kirby, S., Christiansen, M. (Eds.); Oxford University Press.
[2] Hurford, J. R. (1989) Biological evolution of the Saussurean sign as a component of the language acquisition device. Lingua 77 (2): 187 - 222.
[3] Acerbi, A., Parisi, D. (2006) Cultural transmission between and within generations. J. Artificial Societies and Social Simulation 9.
[4] Nowak, M.A., Krakauer, D. (1999) The evolution of language. Proc. Of the National Academy of Sciences USA 96: 8028 - 8033.
[5] Maynard-Smith, J., Szathmaary, E. (1995) The Major Transitions in Evolution. Oxford University Press.
[6] Hofbauer, J., Sigmund K. (2003) Evolutionary game dynamics. Bulletin of Am. Math. Soc.40(4):479 - 519.
[7] Hurford, J. R. (1999) The evolution of language and languages. In The Evolution of Culture. Dunbar, R., Knight, C., Power, C (Eds.); Edinburgh University Press.
[8] Pinker, S., Jackendoff, R. (2005) What's special about the human language faculty? Cognition 95 (2): 201 - 236.
[9] Pinker, S. (1994). The Language Instinct. William Morrow and Company.
[10] Liberman, P. (1984) The Biology and Evolution of Language. Harvard University Press.
[11] Mirolli, M. Parisi, D. (2005) How can we explain the emergence of a language that benefits the hearer but not the speaker? Connection Science 17 (3 - 4): 325 - 341.
[12] Mirolli, M., Parisi, D. (2006) Talking to oneself as a selective pressure for the emergence of language. In The Evolution of Language. Cangelosi, A., Smith, A.D.M., Smith,K. (Eds.); World Scientific.
[13] Pinker, S., Bloom, A. (1990) Natural Language and natural selection. Behavioral and Brain Sciences 13: 707 - 784.
[14] Liberman, P. (1991) Uniquely Human: The Evolution of Speech, Thought and Selfless Behavior. Harvard University Press.
[15] Nowak, M.A., Plotkin, J. B., Jansen, V.A.A. (2000) The evolution of syntactic communication. Nature 404: 495 - 498.

[16] Brighton, H., Smith, K., Kirby, S. (2005) Language as an evolutionary system. Physics of Life Reviews 2: 177 - 226.
[17] Gell-Mann, M. (1994) Complex adaptive systems. In Complexity: metaphors, Models and Reality. Eds. Cowan, G., Pines, D., Meitzer, D. SFI Studies in the Sciences of Complexity. Proc. Vol. XIX, Addison-Weseley.
[18] Levin, S.A. (2002) Complex adaptive systems: exploring the known, the unknown and the unknowable. Bull Amer. Math. Soc. 40 (1): 3 - 19.
[19] Steels, L. (1997) The synthetic modeling of language origins. Evol. Of Communication Jour. 1(1): 1 - 35.
[20] Steels, L. (2000) Language as a Complex Adaptive System. Lecture Notes in Computer Science. Parallel Problem Solving in nature PPSN-VI. Ed. Schoenaur, A. Springer-Verlag.
[21] Brighton, H. (2002) Compositional syntax from cultural transmission. Artificial Life 8: 25 - 54.
[22] Kirby, S., Dowman, M., Griffiths, T.L. (2007) Innateness and culture in the evolution of language. PNAS 104: 5241 - 5245.
[23] Demichelis, S., Weibull, J.W. (2008) Language, meaning and games: a model of communication, coordination and evolution. Amer. Econ. Rev. 98: 1292 - 1311.
[24] Fromm, J. (2004) The Emergence of Complexity. Kassel University Press.
[25] 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 Suppl. 1:1 - 26.
[26] Blythe, R.A., Croft, W.A. (2009) The speech community in evolutionary language dynamics. Language Learning 59 Suppl. 1: 47 - 63.
[27] Croft, W. (2009) Towards a social cognitive linguistics. In Evans, V. & Pourcel, S. (Eds.) New Directions in Cognitive Linguistics. Benjamins. pp. 395 - 420.
[28] Binmore, K. (1999) Fun and Games. Houghton Mifflin Co.
[29] Osborne, M. J., Rubinstein, A. (1994) A Course in Game Theory. The MIT Press.
[30] Fudenberg, D., Tirole, J. (1991) Game Theory. The MIT Press.
[31] Webb, J. N. (2007) Game Theory, Decisions, Interaction and Evolution. Springer-Verlag.
[32] Peters, H. (2008) Game Theory, A Multi-Leveled Approach. Springer-Verlag.
[33] Croft, W. (2000) Explaining Language Change: An evolutionary Approach. Longman.
[34] Jager, G. (2006) Convex meanings and evolutionary stability. In A. Cangelosi, A.D.M. Smith, K. Smith (Eds.) The Evolution of Language. Proceedings of 6th International Conference. pp. 139 - 144.
[35] Croft, W. (2010) The origins of grammaticalizations in the verbalization of experience. Linguistics 48
[36] Komarova, N.L., Nowak, M.A. (2003) Learning language: from individuals to populations. In Language Evolution: The States of the Art. Kirby, S., Christiansen, M. (Eds.); Oxford University Press.
[37] Hurford, J.R. (2006) Recent developments in the evolution of language. Cognitive Systems 7 (1): 23 - 32.
[38] Gintis, H (2009) Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction. Princeton University Press.
[39] Cangelosi, A., Parisi, D. (2001) Computer simulation: a new scientific approach to the study of language evolution. In Simulating the Evolution of Language. Cangelosi, A., Parisi, D. (Eds.); Springer.
[40] Parisi, D., Mirolli, M. (2006) The emergence of language: how to simulate it. In Emergence of Communication and Language. Lyon, C., Nehaniv, C., Cangelosi, A. (Eds.); Springer.
[41] Parisi, D. (2006) Simulating the evolutionary emergence of language: a research agenda. In The Evolution of Language. Cangelosi, A., Smith, A.D.M., Smith, K. (Eds.); World Scientific.
[42] Kirby, S. (2002) Natural language from artificial life. Artifiical Life 8: 185 - 215.
[43] Cangelosi, A., Parisi, D. (Eds) (2002) Simulating the Evolution of Language. Springer.
[44] Hopcroft, J.E., Motwani, R., Ullman, J.D. (2006) Introduction to Automata Theory, Languages and Computation. Addison Wesley.
[45] Nowak, M.A., Komarova, N.L., Niyogi, P. (2002) Computational and evolutionary aspects of language. Nature 417: 611 - 617.
[46] Kolman, B., Busby, R.C., Ross, S.C. (2009) Discrete Mathematical Structures. 6th Ed. Prentice-Hall India.
[47] Holland, J. H. (1975) Adaptation in Natural and Artificial Systems. University of Michigan Press.
[48] Goldberg, D.E., Holland, J. H. (Eds.) (1988) Special issue on genetic algorithms. Machine Learning 3 (2 - 3).
[49] Goldberg, D. E. (1989) Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley.
[50] Mitchell, M. (2004) An Introduction to Genetic Algorithms. Prentice-Hall of India.
[51] Reeves, C.R., Rowe, J.E. (2002) Genetic Algorithms - Principles and Perspectives: A Guide to GA Theory. Kluwer Academic Publishers.
[52] Vose, M. (2007) The Simple Genetic Algorithm. Prentice-Hall of India.
[53] Niyogi, P., Berwick, R.C. (1997) A dynamical systems model for language change. Complex Systems 11: 161 - 204.
[54] Hirsch, M.W., Smale, S., Devaney, R.L. (2004) Differential Equations, Dynamical Systems & An Introduction to Chaos. 2nd ed. Academic Press.
[55] Perko, L. (2004) Differential Equations and Dynamical Systems. Springer.
[56] Hofbauer, J., Sigmund, K. (1998) Evolutionary Games and Population Dynamics. Cambridge University Press.
[57] Taylor, C., Fudenberg, D., Sasaki, A., Nowak, M. A. (2004) Evolutionary game dynamics in finite populations. J. Math. Biol. 66: 1621 - 1644.
[58] Page, K. M., Nowak, M. A., (2002) Unifying evolutionary dynamics. J. Theor. Biol. 219: 93 - 98.
[59] Prabhu, A., Dhakan, P., Bhattacharya, S. (2008) A genetic algorithm for the replicator dynamics of a single-species population. J. Computational Mathematics 1 (1): 60 - 68.
[60] Nimwegen, E., Crutchfield, J. P., Mitchell, M. (1997) Finite populations induce metastability in

evolutionary search. Phys. Lett. A 229 (2): 144 - 150.

[61] Forrest, S., Mitchell, M. (1993) What makes a problem hard for a genetic algorithm? Some anomalous results and their explanation. Machine Learning 13: 285 - 319.

[62] Wright, A. H., Rowe, J. E., Poli, R., Stephens, R. (2003) Bistability in a gene pool genetic algorithm with mutation. In Foundations of Genetic Algorithms, Vol 7. De Jong, E., Poli, R., Stephens, R. (Eds.) Morgan Kauffman.

[63] Brighton, H., Kirby, S. (2006) Understanding linguistic evolution by visualizing the emergence of topographic mappings. Artificial Life 12: 229 - 242.