

# Chaos and language

W. Garrett Mitchener<sup>1</sup> and Martin A. Nowak<sup>2</sup>

This research took place at the Institute for Advanced Study and Princeton University in Princeton, New Jersey, USA.

## Abstract

Human language is a complex communication system with unlimited expressibility. Children spontaneously develop a native language by exposure to linguistic data from their speech community. Over historical time, languages change dramatically and unpredictably by accumulation of small changes and by interaction with other languages. We have previously developed a mathematical model for the acquisition and evolution of language in heterogeneous populations of speakers. This model is based on game dynamical equations with learning. Here we show that simple examples of such equations can display complex limit cycles and chaos. Hence, language dynamical equations mimic complicated and unpredictable changes of languages over time. In terms of evolutionary game theory, we note that imperfect learning can induce chaotic switching among strict Nash equilibria.

Keywords: language; evolution; game dynamics; learning; dynamical systems; chaos.

## 1 Introduction

Human language consists of a stream of sounds that give rise to phonemes, words, phrases, sentences, poetry, and publications. Grammar is the internal computational system of language (Jackendoff, 2002; Pinker, 1990). Native speakers have a subconscious, internal representation of a grammar, which enables them to generate and understand sentences of their language. Children build this internal grammar by generalizing from linguistic data they receive from their speech community (Lightfoot, 1991, 1999). Languages change over time. The purpose of this research is to model language change using tools from evolutionary game dynamics, and to use such models to gain insight into the evolution of the language faculty.

Given an unrestricted set of languages and the problem of identifying a target language based on a sequence of example sentences, no learning algorithm can guarantee that it will correctly identify every possible target in finite time (Angluin, 1987; Gold, 1967; Nowak *et al.*, 2002; Osherson *et al.*, 1986; Valiant, 1984; Vapnik, 1995). *Universal grammar* or *UG* specifies the restricted set of languages that is learnable by the human brain (Chomsky, 1988). UG is

---

<sup>1</sup>Duke University Math Department, Box 90320, Durham, NC, 27708, USA. E-mail: [wgm@math.duke.edu](mailto:wgm@math.duke.edu).

<sup>2</sup>Center for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA. E-mail: [nowak@math.harvard.edu](mailto:nowak@math.harvard.edu)

influenced by evolution; it is a consequence of the architecture of our brain, which in turn is genetically determined (Hauser, 1996; Lai *et al.*, 2001; Nowak *et al.*, 2001).

Languages are not static. Phonological systems tend to change systematically yet unpredictably, as in Grimm’s Law, which describes consonant changes from Proto-Indo-European to Germanic languages, and vowel shifts such as those taking place in English (Trask, 1996). Contact with other languages can bring about catastrophic changes in syntax. For example, the loss of case endings on nouns in Old English is thought to be due to contact with Old Norse (Lightfoot, 1999). Language change also has the potential for oscillations, such as the morphology type cycle. Languages tend to use either isolating morphology, with many small words each carrying a single piece of meaning, or agglutinating morphology, in which words consist of a stem plus many affixes carrying a single piece of meaning, or inflecting morphology, in which each affix carries many pieces of meaning. Roughly, languages tend to change from isolating to agglutinating to inflecting and back to isolating (Crowley, 1998). English, for example, has lost case endings and other forms of inflection and is changing from inflecting to isolating morphology.

Changes such as these arise through learning error, meaning that a child has acquired a grammar different from the parents’. Errors arise when the primary linguistic data fails to trigger the acquisition of the expected grammar, as may happen when the data underspecifies the grammar or is inconsistent with every available hypothesis. Such data can arise from natural variation in speech and when sample sentences are drawn from multiple grammars as in language contact. Acquisition is often robust enough to ignore meaningless variation and detect the presence of multiple languages, resulting in bilingualism. However, this is not always the case, and history shows that learning errors are possible.

We would like to use mathematical models to study language change. Several different models have been proposed (Briscoe, 2000; Cangelosi & Parisi, 2001; Hurford *et al.*, 1998; Kirby, 2001; Niyogi & Berwick, 1996; Steels, 1996). The goal of this paper is to display instances of a model developed by Nowak *et al.* (2001) that are as simple as possible and also exhibit change and sensitivity analogous to that observed in real languages.

## 2 The Model

Let us consider a group of individuals whose UG admits a finite set of grammars  $G_1, G_2, \dots, G_n$ . Denote by  $Q_{ij}$  the probability that a child learner will acquire grammar  $G_j$  when exposed to sample sentences generated by a parent speaking  $G_i$ . The linguistic data available to the child and the acquisition algorithm determine  $Q$ . Imperfect learning means that  $Q_{ii} < 1$  for at least some  $i$ , which implies that sometimes the learner will end up with a different grammar. As a simplifying assumption, the  $Q$  matrix is taken to be constant in time. Hence, it most accurately reflects scenarios such an isolated population or one subject to a constant level of contact, and is of limited use when learning probabilities

are fluctuating.

Consider a large, well-mixed population where each individual speaks one of the grammars  $G_1, G_2, \dots, G_n$ . The fraction of the population that speaks  $G_i$  is denoted  $x_i$ . We have  $\sum_i x_i = 1$ . The population is subject to natural selection: Individuals derive a benefit from communicating successfully with each other. Let us define a payoff matrix  $B$ , where  $B_{ij}$  is the benefit to a speaker of  $G_i$  from an encounter with a speaker of  $G_j$ . The entries of this matrix may include effects such as the benefit of correct communication, cost of ambiguity, and so forth. A natural assumption is that people communicate best with others who have the same grammar. In this case,  $B$  is diagonally dominant, which implies that each grammar is a strict Nash equilibrium. With perfect learning, each language would then be an evolutionarily stable equilibrium.

The fitness associated with grammar  $G_j$  is the weighted average payoff  $F_j = \sum_k B_{jk} x_k$ . Here we make the simplifying assumption that communication is the dominant source of fitness, thereby incorporating selection in favor of individuals who communicate well.

The average fitness of the population is given by  $\phi = \sum_j F_j x_j$ . The language dynamical equation is given by

$$\dot{x}_j = \sum_{i=1}^n F_i x_i Q_{ij} - \phi x_j, \quad j = 1 \dots n. \quad (1)$$

This system can be interpreted as a replicator (or game dynamical) equation (Hofbauer & Sigmund, 1998) with learning or mutation (Stadler & Schuster, 1992). The standard replicator equation is obtained in the limit of perfect learning. Furthermore, there is a correspondence between (1) and the Price equation, which is a general description of evolutionary dynamics (Page & Nowak, 2002; Price, 1972). Equation (1) can also be formulated to describe competition among multiple UGs (Mitchener & Nowak, 2003).

### 3 Limit cycles and chaos

As a specific example, let us consider the following payoff matrix for a case of three grammars.

$$B = \begin{pmatrix} 0.88 & 0.2 & 0.2 \\ 0.2 & 0.88 & 0.2 \\ 0.2 & 0.2 & 0.88 \end{pmatrix}. \quad (2)$$

All grammars are equally good. For perfect learning, this payoff matrix leads to very simple dynamics: All trajectories converge to one of three stable equilibria where the whole population speaks the same language.

Imperfect learning, however, can induce very different behavior. For example, let us consider the following learning matrix in conjunction with (2):

$$Q = \begin{pmatrix} 0.79 & 0.2 & 0.01 \\ 0.01 & 0.79 & 0.2 \\ 0.2 & 0.01 & 0.79 \end{pmatrix}.$$

For each grammar, the most likely outcome of the learning process is the correct grammar, and there is one second-most-likely outcome. As shown in Figure 1a, these parameters produce stable oscillations, as learning errors in the subpopulation speaking  $G_1$  feed into  $G_2$ , and  $G_2$  feeds into  $G_3$ , and  $G_3$  feeds back into  $G_1$ . This cycle is suggestive of the morphology type cycle. Note that stable limit cycles such as this are impossible for 3-dimensional replicator equations (Hofbauer & Sigmund, 1998, p. 78).

If learning becomes less accurate, then the limit cycle breaks down, resulting in a spiral sink, as shown in Figure 1b.

This spiral opens the way for more complex behavior, as it can be used to construct a period doubling cascade similar to Šilnikov’s mechanism (Guckenheimer & Holmes, 1990): We add two more grammars, and set up the learning matrix so that the spiral sink is unstable in the new dimensions. We fix  $B$  as follows:

$$B = \begin{pmatrix} 0.88 & 0.2 & 0.2 & 0 & 0.3 \\ 0.2 & 0.88 & 0.2 & 0 & 0.3 \\ 0.2 & 0.2 & 0.88 & 0 & 0.3 \\ 0.3 & 0.3 & 0.3 & 0.88 & 0 \\ 0 & 0 & 0 & 0.3 & 0.88 \end{pmatrix}. \quad (3)$$

Thus, we consider 5 languages, each of which is a strict Nash equilibrium. For perfect learning, there would again be stable equilibria where all individuals speak the same language. Instead of perfect learning, let us consider a one parameter family of  $Q$  matrices:

$$Q = \begin{pmatrix} 0.75 & 0.2 & 0.01 & 0.04 & 0 \\ 0.01 & 0.75 & 0.2 & 0.04 & 0 \\ 0.2 & 0.01 & 0.75 & 0.04 & 0 \\ 0 & 0 & 0 & \mu & 1 - \mu \\ 1 - \mu & 0 & 0 & 0 & \mu \end{pmatrix}. \quad (4)$$

The parameter  $\mu$  denotes the learning accuracy of grammars  $G_4$  and  $G_5$ . For appropriate choices of  $\mu$ , trajectories can escape from the middle of the  $G_1, G_2, G_3$  spiral. The population slowly leaks into  $G_4$ , then into  $G_5$ , and back into  $G_1$  to return to the spiral. Varying  $\mu$  alters how trajectories escape the spiral. The result is a limit cycle in four dimensions that can undergo period-doubling bifurcations that lead to chaos, as shown in Figure 2.

## 4 Discussion and Conclusion

Much of the existing literature on language models focuses on equilibrium behavior, modeling the fact that languages are generally stable on time scales of about a century. In contrast, the behaviors displayed here attempt to capture two important features observed in changing languages. First, some language change arises from reanalysis and variation among speakers, and follows regular patterns; examples are lenition, vowel shifts, and changes of morphology type. Thus, for time scales on the order of several centuries, the oscillations discussed

here appear to be more realistic than stable equilibria as in (Mitchener, 2003). Second, language change is unpredictable and highly sensitive to perturbations. Many changes, particularly those associated with borrowed vocabulary, are triggered by language contact. The same kind of unpredictability and sensitivity is exhibited by chaotic dynamical systems.

It is known that the game dynamical equation for four or more dimensions can generate limit cycles and chaos (Nowak & Sigmund, 1993; Stadler & Schuster, 1992). In contrast, the present example of chaos is not caused by game dynamics, but rather by learning errors.

Briscoe (2000) has proposed an agent-based model for language acquisition and evolution. He uses a specific, detailed learning algorithm and UG motivated directly by linguistic research. With the intent of producing a more mathematically tractable model, the formulation of (1) assumes a larger population where individual-level details influence the dynamics through aggregate effects. It would be a fascinating task to express Briscoe's model in terms of (1) and study the possible range of evolutionary dynamics.

In summary, the language dynamical equation is a game dynamical equation with learning. Here, we show that complex limit cycles and chaos can arise even for very simple choices of the payoff and the learning matrices. In our example, we considered five languages (strategies), each of which is a strict Nash equilibrium. Pure game dynamics would have five stable equilibria corresponding to linguistically homogeneous populations. However, a carefully structured learning matrix is sufficient to induce chaos. Thus, very conservative, natural choices of payoff and learning matrices lead to deterministic chaos.

Our analysis has implications for historical linguistics, language evolution and evolutionary game theory. Simple learning errors can lead to complex, unpredictable and seemingly stochastic changes in languages over time. For game dynamics, we note that imperfect learning can lead to chaotic switching among strict Nash equilibria.

## References

- Angluin, D. 1987 Learning regular sets from queries and counterexamples. *Information and Computation*, **75**, 87–106.
- Briscoe, E. J. 2000 Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language*, **76**, 245–296.
- Cangelosi, A. & Parisi, D., eds. 2001 *Simulating the Evolution of Language*. Springer-Verlag.
- Chomsky, N. 1988 *Language and Problems of Knowledge*. MIT Press.
- Crowley, T. 1998 *An Introduction to Historical Linguistics*. Oxford University Press, 3rd edition.

- Gold, E. M. 1967 Language identification in the limit. *Information and Control*, **10**, 447–474.
- Guckenheimer, J. & Holmes, P. 1990 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag.
- Hauser, M. D. 1996 *The Evolution of Communication*. Cambridge, MA: Harvard University Press.
- Hofbauer, J. & Sigmund, K. 1998 *Evolutionary Games and Population Dynamics*. Cambridge University Press.
- Hurford, J. R., Studdert-Kennedy, M. & Knight, C., eds. 1998 *Approaches to the Evolution of Language*. Cambridge University Press.
- Jackendoff, R. 2002 *Foundations of Language*. Oxford: Oxford University Press.
- Kirby, S. 2001 Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, **5**, 102–110.
- Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. 2001 A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, **413**, 519–523.
- Lightfoot, D. 1991 *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.
- Lightfoot, D. 1999 *The Development of Language: Acquisition, Changes and Evolution*. Blackwell Publishers.
- Mitchener, W. G. 2003 Bifurcation analysis of the fully symmetric language dynamical equation. *Journal of Mathematical Biology*, **46**, 265–285.
- Mitchener, W. G. & Nowak, M. A. 2003 Competitive exclusion and coexistence of universal grammars. *Bulletin of Mathematical Biology*, **65**, 67–93.
- Niyogi, P. & Berwick, R. C. 1996 A language learning model for finite parameter spaces. *Cognition*, **61**, 161–193.
- Nowak, M. & Sigmund, K. 1993 Chaos and the evolution of cooperation. *Proceedings of the National Academy of Sciences, USA*, **90**, 5091–5094.
- Nowak, M. A., Komarova, N. L. & Niyogi, P. 2001 Evolution of universal grammar. *Science*, **291**, 114–118.
- Nowak, M. A., Komarova, N. L. & Niyogi, P. 2002 Computational and evolutionary aspects of language. *Nature*, **417**, 611–617.
- Osherson, D., Stob, M. & Weinstein, S. 1986 *Systems That Learn*. Cambridge, MA: MIT Press.

- Page, K. M. & Nowak, M. A. 2002 Unifying evolutionary dynamics. *Journal of Theoretical Biology*, **219**, 93–98.
- Pinker, S. 1990 *The Language Instinct*. New York: W. Morrow and Company.
- Price, G. R. 1972 Extension of covariance mathematics. *Annals of Human Genetics*, **35**, 485–490.
- Stadler, P. F. & Schuster, P. 1992 Mutation in autocatalytic reaction networks. *Journal of Mathematical Biology*, **30**, 597–632.
- Steels, L. 1996 Self-organizing vocabularies. In *Proceedings of the Artificial Life Conference*, volume 5. MIT Press.
- Trask, R. L. 1996 *Historical Linguistics*. London: Arnold.
- Valiant, L. G. 1984 A theory of the learnable. *Communications of the ACM*, **27**, 436–445.
- Vapnik, V. 1995 *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

## A Figures

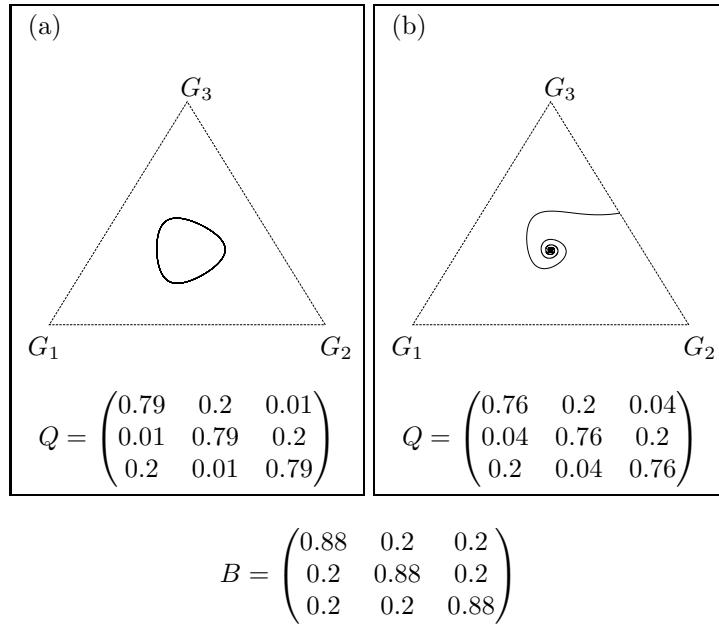


Figure 1: **(a)** A stable limit cycle. The oscillations are caused by the learning algorithm. Errors from children learning  $G_1$  feed into  $G_2$ , and  $G_2$  feeds into  $G_3$ , and  $G_3$  feeds back into  $G_1$ . **(b)** A spiral sink that results from the limit cycle in (a) when the  $Q$  matrix is changed as shown. The  $B$  matrix is common to both figures.



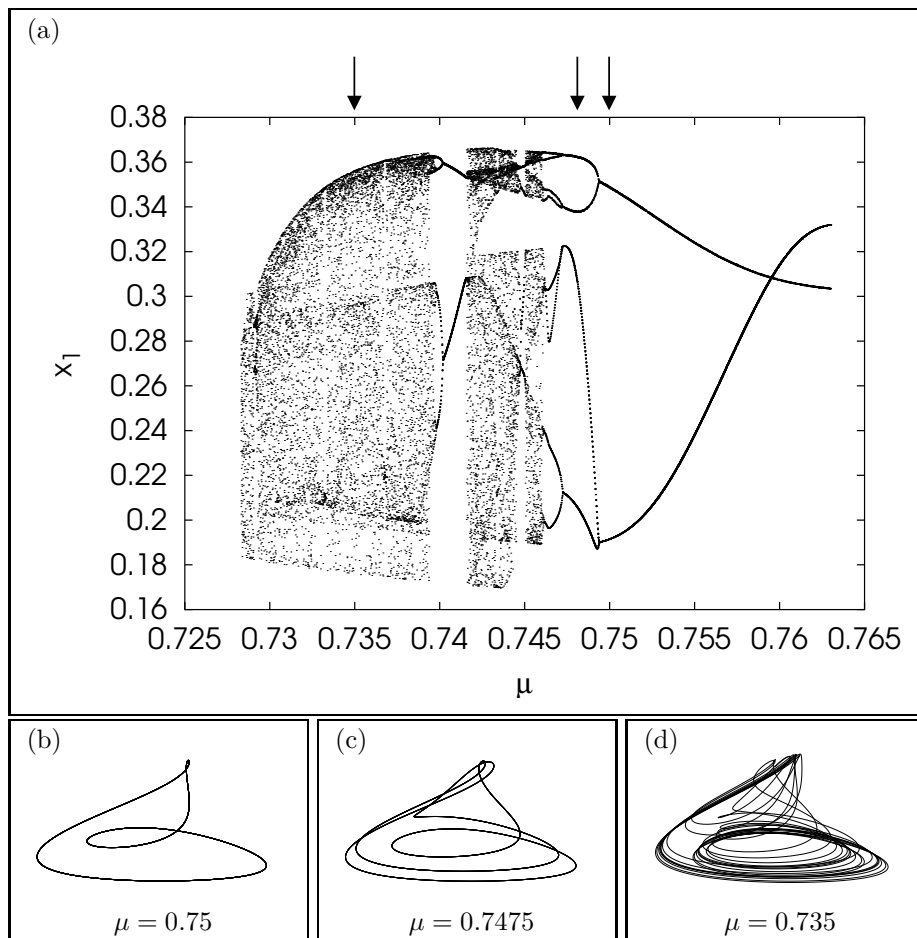


Figure 2: (a) Cascade diagram. The horizontal axis shows a range of values of  $\mu$ . For each value of  $\mu$ , an orbit is traced. Each time  $x_4$  crosses 0.19, a dot is plotted at  $(\mu, x_1)$ . Sample orbits are drawn for the three values of  $\mu$  indicated by arrows. (b) When  $\mu = 0.75$ , there are two dots representing two extremes of this stable limit cycle. (c) When  $\mu = 0.7475$ , there are twice as many dots because the limit cycle has undergone a period doubling bifurcation. (d) For  $\mu = 0.735$ , the orbit appears to be chaotic.