Bulletin of Mathematical Biology (2003) **00**, 1–30

Competitive Exclusion and Coexistence of Universal Grammars (preprint with corrections)

W. GARRETT MITCHENER
E-mail: wmitchen @ princeton.edu
Program in Applied and Computational Mathematics
Fine Hall
Washington Road
Princeton, NJ 08544-1000, USA

MARTIN A. NOWAK
E-mail: nowak @ ias.edu
Institute for Advanced Study
Einstein Drive
Princeton, NJ 08540, USA

Universal grammar(UG) is a list of innate constraints that specify the set of grammars that can be learned by the child during primary language acquisition. UG of the human brain has been shaped by evolution. Evolution requires variation. Hence, we have to postulate and study variation of UG. We investigate evolutionary dynamics and language acquisition in the context of multiple universal grammars. We provide examples for competitive exclusion and stable coexistence of different UGs. More specific UGs admit fewer candidate grammars, and less specific UGs admit more candidate grammars. We will analyze conditions for more specific UGs to outcompete less specific UGs and vice versa. An interesting finding is that less specific UGs can resist invasion by more specific UGs if learning is more accurate. In other words, accurate learning stabilizes UGs that admit large numbers of candidate grammars.

1. Introduction

Human languages are composed of a lexicon, which is a set of words and their meanings, and a grammar, which is a set of rules for building and interpreting sentences (Pinker, 1990). Children learn both parts inductively, based on the linguistic input they receive. The task of acquiring grammar from example sentences is known to require some constraints on the set of possible grammars. Universal grammar is a set of constraints that guide primary language acquisition (Chomsky, 1965; Chomsky, 1972; Lightfoot, 1999; Lightfoot, 1991; Wexler & Culicover, 1980). In general, language

0092-8240/03/000001 + 30 $17.00/0mb03????

acquisition can be formulated as a process of choosing among a (finite) number of candidate grammars specified by UG (Gibson & Wexler, 1994).

UG is both a product of evolution and a consequence of mathematical or computational constraints that apply to any communication system (Uriagereka, 1998). Since evolution requires variation, we have to study natural selection among different UGs. Hence, this paper is an investigation into what happens when more than one UG is present in a given population.

While a genetically encoded (innate) UG is a logical requirement for the process of language acquisition (see (Nowak *et al.*, 2002)), there is considerable debate about the nature of the genetically encoded constraints. Interestingly, in a recent study, a mutation in a gene was linked to a language disorder in humans (Lai *et al.*, 2001) providing a specific example of a genetic modification that affects linguistic performance. It is therefore natural to construct population models which incorporate genetic variation in the form of multiple universal grammars, and to explore the long term behavior of such models.

We explore three possibilities of selective dynamics. The first, *dominance,* means that one particular UG takes over the population from any initial state. The second, *competitive exclusion,* happens when some UG takes over the population, but the initial state influences which one. The third, *coexistence,* means that two or more UGs exist stably. We construct a dynamical system describing a population of individuals. Each individual has an innate UG and speaks one of the grammars generated by this UG. Individuals reproduce in proportion to their ability to communicate with the whole population, passing on their UG to their offspring genetically, and attempting to teach their grammar to their children. The children can make mistakes and learn a different grammar than their parents speak, but within the constraints of their UG.

Section 2 describes the mathematical details of the model, which is an extension of the language dynamical equation from (Komarova *et al.*, 2001), (Nowak *et al.*, 2001), (Mitchener, 2002) and (Nowak *et al.*, 2002). It assumes that there are a number of universal grammars, and people acquire one of the grammars specified by their UG based on sample sentences they hear from their parents.

Section 3 analyzes a one-dimensional case with one UG that specifies two candidate grammars. This simple case is used as a building block for subsequent analysis.

In Section 4, we study the selection between two universal grammars: U_1 admits grammar G_1 while U_2 admits grammars G_1 and G_2. This case is of interest because it illustrates the competition between a more specific UG, that is, one with more constraints and therefore fewer options, and a less specific UG. We never find coexistence between U_1 and U_2. For certain parameter values, U_1 dominates U_2, meaning that the only stable equilibrium consists entirely of individuals with U_1. For other parameter values, we find competitive exclusion: Both U_1 and U_2 can give rise to stable equilibria. In particular, U_2 is stable against invasion by U_1 if learning is sufficiently accurate and if most individuals use G_2.

In Section 5, we study two extensions. First, we consider what happens if a multi-grammar UG denoted by U_0, which allows grammars G_1 through G_n, competes with n single-grammar UGs denoted by U_1 through U_n, where U_j allows only G_j. To simplify the analysis, symmetry is imposed on the model. It turns out that U_0 is never able to take over the population, but that any one of the single-grammar UGs can. In a second extension, U_0 only competes against U_1. In this case, there can be a stable equilibrium that consists entirely of individuals with U_0, provided its learning algorithm is sufficiently reliable, and the population does not contain too many speakers of G_1.

In Section 6 we allow grammars to be ambiguous, and study the case where U_1 admits grammar G_1, while U_2 admits grammars G_2 and G_3. We provide an example where U_2 dominates U_1 and an example where U_1 and U_2 coexist in a stable equilibrium.

In Section 7, we draw some conclusions and discuss the next steps in this line of research.

The fascinating question of language evolution has generated an extensive literature (Aitchinson, 1987; Bickerton, 1990; Ghazanfar & Hauser, 1999; Grassly *et al.*, 2000; Hauser, 1996; Hauser *et al.*, 2001; Hurford *et al.*, 1998; Krakauer, 2001; Jackendoff, 1999; Lachmann *et al.*, 2001; Lieberman, 1984; Pinker, 1990; Pinker & Bloom, 1990; Ramus *et al.*, 2000; Studdert-Kennedy, 2000). The purpose of this paper is to contribute to the understanding of the evolution of grammar through mathematical models (Cangelosi & Parisi, 2001; Ferrer i Cancho & Solé, 2001b; Ferrer i Cancho & Solé, 2001a; Kirby, 2001; Nowak & Krakauer, 1999; Nowak *et al.*, 2000; Nowak *et al.*, 2001; Nowak *et al.*, 2002) that incorporate ideas from linguistics, as well as evolutionary game theory (Hofbauer & Sigmund, 1998) and different forms of learning theory (Gold, 1967; Niyogi, 1998; Niyogi & Berwick, 1996; Valiant, 1984; Vapnik, 1995).

2. Language dynamics with multiple universal grammars

Suppose we have a large population, each member of which is born with one of the N universal grammars $U_1, U_2, \dots U_N$ and speaks one of the n grammars $G_1, G_2, \dots G_n$. Each UG consists of a list of which grammars it allows, and has an associated language acquisition algorithm. The grammars are assumed to have an overlap given by the matrix A, where $A_{i,j}$ is the probability that a sentence spoken at random by a speaker of G_i can be parsed by a speaker of G_j. A grammar G_i is said to be *unambiguous* if

$A_{i,i} = 1$, because $A_{i,i} < 1$ implies that two people with the same grammar can misunderstand each other due to some sentence with multiple meanings.

Define $x_{j,K}$ to be the fraction of the population which speaks G_j and possesses universal grammar U_K. We have $\sum_K \sum_j x_{j,K} = 1$. Every population state can be represented as a point on a simplex. The population changes over time in that individuals reproduce at a rate determined by their ability to communicate with everyone else, passing their universal grammar to their offspring via genetic inheritance, and passing their language on through teaching and learning. As a simplifying assumption, we ignore genetic mutation, but include the possibility that children make mistakes learning their parents' language. The learning process is expressed by the three-axis matrix Q, where $Q_{i,j,K}$ is the probability that a parent speaking G_i produces a child speaking G_j given that both have universal grammar U_K. Since every child must speak some language, Q is row-stochastic, that is, $\sum_j Q_{i,j,K} = 1$ for all i and K. The reproductive rate F_j depends on which grammar an individual uses and the composition of the rest of the population, and is given by

$$F_j = \sum_{K=1}^{N} \sum_{i=1}^{n} B_{i,j} x_{i,K} \quad \text{where} \quad B_{i,j} = \frac{A_{i,j} + A_{j,i}}{2}. \tag{1}$$

To write the ordinary differential equation (ODE) governing the population dynamics, we also need the variable ϕ which represents the average reproductive rate of the population:

$$\phi = \sum_{K=1}^{N} \sum_{j=1}^{n} F_j x_{j,K}. \tag{2}$$

The language dynamical equation with multiple universal grammars is then

$$\dot{x}_{j,K} = \sum_{i=1}^{n} F_i x_{i,K} Q_{i,j,K} - \phi x_{j,K} \text{ where } j = 1 \ldots n, K = 1 \ldots N. \tag{3}$$

The first term says that the sub-population which has universal grammar U_K and speaks with grammar G_i will produce $F_i x_{i,K}$ offspring, of which $Q_{i,j,K}$ end up speaking G_j. The second term is to enforce the constraint $\sum_j \sum_K \dot{x}_{j,K} = 0$ so that $\sum_j \sum_K x_{j,K} = 1$ for all time. To see this, let

$$M_k = \sum_{j=1}^{n} \sum_{K=1}^{N} x_{j,K}^k \tag{4}$$

so that

$$\begin{aligned}
\dot{M}_1 &= \sum_{j=1}^{n}\sum_{K=1}^{N} \dot{x}_{j,K} \\
&= \sum_{K=1}^{N}\left(\sum_{i=1}^{n} F_i x_{i,K} \sum_{j=1}^{n} Q_{i,j,K}\right) - \phi \sum_{j=1}^{n} x_{j,K} \\
&= \sum_{K=1}^{N}\sum_{i=1}^{n} F_i \dot{x}_{i,K} - \phi M_1 \\
&= \phi(1 - M_1).
\end{aligned}$$

Since $\phi \geq 0$, there is a stable equilibrium at $M_1 = 1$. Hence, the population state is confined to the hyperplane $\sum_j \sum_K x_{j,K} = 1$.

Furthermore, the positive orthant, defined by the inequalities $x_{j,K} \geq 0$ for all j and K, is a trapping region. To see this, observe that for each bounding hyperplane given by $x_{j,K} = 0$, the value of $\dot{x}_{j,K}$ is a sum of terms each of which is $F_i x_{i,K} Q_{i,j,K} \geq 0$. Thus, the vector field points either into the bounding hyperplane or into the interior of the positive orthant. We will therefore restrict our attention to trajectories in the simplex $S_{(Nn)}$, which is the intersection of the hyperplane given by $\sum_j \sum_K x_{j,K} = 1$ and the positive orthant.

In some cases, such as the one in Section 4, we will further restrict our attention to a face of $S_{(Nn)}$, which is itself a lower-dimensional simplex. This restriction comes from assuming that some U_K disallows some G_j, so that $x_{j,K}$ is fixed at 0.

3. Two grammars and one universal grammar

The case to be examined here, that of a single universal grammar which generates two unambiguous grammars, takes place in S_2, a one-dimensional phase space. We use this case as an essential building block in later sections.

3.1. *Parameter values.* Since there is only one universal grammar, we will omit the K subscript from x and Q. There are three choices of real numbers which fill in all the parameters for this case of the language dynamical equation, which come from considering the possibilities for A and Q as follows. The most general form of the overlap matrix A for two unambiguous grammars is

$$A = \begin{pmatrix} 1 & a_{1,2} \\ a_{2,1} & 1 \end{pmatrix}.$$

However, the A matrix only enters the dynamical system through the B matrix, as in (1), and since B is a symmetric matrix,

$$B = \frac{A + A^T}{2} = \begin{pmatrix} 1 & (a_{1,2} + a_{2,1})/2 \\ (a_{1,2} + a_{2,1})/2 & 1 \end{pmatrix},$$

there is really only one degree of freedom in choosing A. So, we define

$$b = \frac{a_{1,2} + a_{2,1}}{2}, \tag{5}$$

and allow this to be the one free parameter determined by the overlap between G_1 and G_2. The most general form for the learning algorithm matrix Q is

$$Q = \begin{pmatrix} q_1 & 1 - q_1 \\ 1 - q_2 & q_2 \end{pmatrix}, \tag{6}$$

which has two degrees of freedom. The ranges of the parameters are $0 < b < 1$, $0 < q_1 < 1$, and $0 < q_2 < 1$. Although we can certainly consider the cases where q_1 and q_2 are less than $1/2$, these are somewhat pathological because they represent a situation where children are more likely to learn the grammar opposite to the one their parents speak. Furthermore, if $b = 0$ then G_1 and G_2 have nothing in common and when $b = 1$ they are identical. Both of these settings are degenerate and will not be analyzed here.

3.2. *Fixed point analysis.* In the present case, everything takes place on a unit interval $0 \leq x_1 \leq 1$, and the dynamical system is one dimensional, as can be seen by expanding (3) and replacing x_2 with $1 - x_1$:

$$\begin{aligned} \dot{x}_1 = &(1 - q_2) \\ &+ (-3 + b(1 + q_1 - q_2) + 2q_2)x_1 \\ &+ (1 - b)(3 + q_1 - q_2)x_1^2 \\ &- 2(1 - b)x_1^3. \end{aligned} \tag{7}$$

It is useful to change coordinates to $x_1 = 1 - 2r$ so that the dynamical system inhabits an interval $-1 \leq r \leq 1$ that is symmetric about 0. The vector field now takes on the form

$$\begin{aligned} \dot{r} = -\frac{1}{2}\Big(&(1 + b)(q_1 - q_2) \\ &+ (3 + b - 2(q_1 + q_2)))r \\ &+ (1 - b)(q_1 - q_2)r^2 \\ &+ (1 - b)r^3\Big). \end{aligned} \tag{8}$$

By straightforward calculation, if $r = -1$ then $\dot{r} = 2(1 - q_1) > 0$, and if $r = 1$ then $\dot{r} = 2(-1 + q_2) < 0$. By the intermediate value theorem, there must be at least one fixed point in the interval. Since $\dot{r}$ is a cubic polynomial in r, there can be either one, two, or three fixed points, depending on the choice of parameters. Keeping in mind that the vector field points inward at both ends of the interval, the dynamical system must follow one of the phase portraits in Figure 1. Two kinds of bifurcations are possible: saddle-node and pitchfork. The remainder of this section will be spent developing a partial answer to the question of which parameter values cause particular bifurcations, and where the fixed points are when they take place. Rather than solve $\dot{r} = 0$ directly, we will make use of the following variations of some well-known lemmas (see Chapter 1 of (Andronov *et al.*, 1971) or Chapter 4 of (Ahlfors, 1979)) and indirect methods to extract information about the bifurcations.

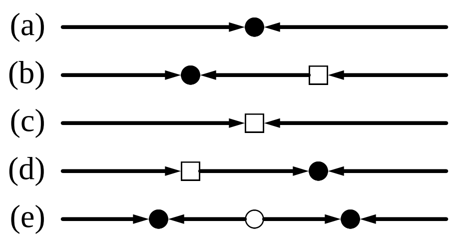


Figure 1. Possible phase portraits for the base line of the simplex. Key: $\bullet$ indicates a sink, $\circ$ indicates a source, $\square$ indicates a non-hyperbolic fixed point. Pictures (a) and (e) are structurally stable, (b) and (d) are saddle-node or transcritical bifurcations, and (c) is a pitchfork bifurcation.

Lemma 1. *Let $f(x)$ be a polynomial with a root z of multiplicity $n \geq 1$. Then z is a root of $f'(x)$ with multiplicity $n - 1$.*

Proof. Write $f(x) = (x - z)^n g(x)$ where $g(z) \neq 0$. Then

$$\begin{aligned} f'(x) &= n(x - z)^{n-1} g(x) + (x - z)^n g'(x) \\ &= (x - z)^{n-1}(ng(x) + (x - z)g'(x)). \end{aligned}$$

Observe from the first factor in the bottom line that z is a root of $f'(x)$ of multiplicity at least $n - 1$. At $x = z$, the second factor takes the value $ng(z)$ which is nonzero, so the multiplicity of z is exactly $n - 1$. $\square$

Lemma 2. *Let $f(x)$ be a polynomial with a root z such that $f'(z) = 0$. Then z is a root of multiplicity two or more.*

Proof. Let z be a root of f with multiplicity n. Since z is a root of f' of multiplicity $n - 1$ and $n - 1 \geq 1$, it follows that $n \geq 2$. $\square$

Lemma 3. *Given a real-valued polynomial dynamical system $\dot{x} = f(x)$, the non-hyperbolic fixed points are exactly the roots of f of multiplicity two or more.*

Proof. From Lemma 1, every root of f of multiplicity two or more is a non-hyperbolic fixed point. Conversely, if z is a non-hyperbolic fixed point, then $f(z) = 0$ and $f'(z) = 0$, and Lemma 2 guarantees that z is a root of f of multiplicity two or more. □

Lemma 3 is the most useful, as it allows us to find the bifurcation parameters of (8) without explicitly solving a cubic. In particular, for saddle-node and transcritical bifurcations there is a double root of the polynomial and for pitchfork bifurcations there is a triple root of the polynomial. Thus, the parameter settings which generate the non-hyperbolic fixed points in Figure 1 parts (b), (c), and (d) may be found by matching (8) against a general template polynomial with multiple roots, as will be illustrated below.

As a side note, the results of this section will be used to analyze higher dimensional dynamical systems in which both saddle-node and transcritical bifurcations will be possible, both of which are characterized by a double root. Saddle-node bifurcations are distinguished from transcritical bifurcations in that the double root comes into existence at the bifurcation rather than forming from the collision of two pre-existing fixed points. The template polynomial method does not distinguish between these two cases as it can only locate parameter settings that produce non-hyperbolic fixed points. The way in which the parameters change so as to pass through such settings determines which type of bifurcation takes place.

Proposition 4. *The unique parameter setting which produces the phase portrait given in Figure 1 (c) (the pitchfork bifurcation) is*

$$q_1 = q_2 = \frac{3+b}{4}.$$

The non-hyperbolic fixed point is at $r = 0$, corresponding to $x_1 = x_2 = 1/2$, the center of the phase space.

Proof. The technique is to set $\dot{r} = 0$ and seek parameters that generate a triple root. We divide the resulting cubic equation by the coefficient of r^3 to produce a monic polynomial, and set the resulting coefficients equal to the corresponding coefficients of $(r - p)^3$ where p is an unknown variable, corresponding to the non-hyperbolic fixed point. The resulting system of equations is

$$-p^3 = \frac{(1+b)(q_1 - q_2)}{1-b}, \tag{9a}$$

$$3p^2 = \frac{3+b-2q_1-2q_2}{1-b}, \tag{9b}$$

$$-3p = q_1 - q_2. \tag{9c}$$

It turns out that this system can be solved for q_1 and q_2 in terms of b. To begin, we use (9c) to eliminate q_2 in the (9a) which yields

$$p^3 + \frac{3(1+b)}{-1+b} p = 0.$$

This equation has three roots,

$$p = 0, \quad p = \pm\sqrt{3}\sqrt{\frac{1+b}{1-b}}.$$

The second and third roots lie outside the interval of interest $-1 \le p \le 1$, so the only possible solution is $p = 0$ from which it follows that $q_1 = q_2 = (3+b)/4$. □

The cases in which there are two fixed points and one is a double root is significantly more complicated because there is an additional unknown variable. This next result is a partial solution.

Proposition 5. *For the phase portraits shown in Figure 1 parts (b) and (d) (which are saddle-node or transcritical bifurcations), the sink and non-hyperbolic fixed point lie on opposite halves of phase space.*

Proof. We begin as in Proposition 4, but this time matching $\dot{r} = 0$ against the cubic template $(r - p_1)^2(r - p_2)$ where p_1 is the non-hyperbolic fixed point and p_2 is the sink. The initial system of equations is

$$-p_1^2 p_2 = \frac{(1+b)(q_1 - q_2)}{1-b}, \tag{10a}$$

$$p_1^2 + 2p_1 p_2 = \frac{3+b-2q_1-2q_2}{(1-b)}, \tag{10b}$$

$$-2p_1 - p_2 = q_1 - q_2. \tag{10c}$$

We proceed by solving for p_1 in terms of p_2. Substituting (10c) into (10a) results in the quadratic equation

$$\left(\frac{1+b}{1-b}\right)(2p_1 + p_2) = p_1^2 p_2,$$

whose roots are

$$p_1 = \frac{C}{p_2} \pm \sqrt{\frac{C^2}{p_2^2} + C} \quad \text{where} \quad C = \frac{1+b}{1-b} > 1.$$

From here, we demonstrate that $p_2 > 0$ implies $p_1 < 0$. Clearly

$$\sqrt{\frac{C^2}{p_2^2} + C} > 1,$$

which implies that the + root lies outside the phase space and is therefore extraneous. Hence the non-hyperbolic fixed point must be located at the − root. It is easy to see that

$$\sqrt{\frac{C^2}{p_2^2} + C} > \frac{C}{p_2},$$

from which it follows that

$$p_1 = \frac{C}{p_2} - \sqrt{\frac{C^2}{p_2^2} + C} < 0.$$

A similar argument shows that $p_2 < 0$ implies $p_1 > 0$. If $p_1 = p_2 = 0$, we have the case of Proposition 4 which is a different phase portrait. □

This next proposition is a constraint that is needed in Section 4.

Proposition 6. *There is no setting of the parameters for which three fixed points lie on the same side of the middle.*

Proof. Suppose that we start at parameter values for which there is only one fixed point, and change them smoothly so that there are three afterward. This means the system must undergo either a saddle-node or pitchfork bifurcation. In the case of a saddle-node bifurcation, Proposition 5 ensures that the two new fixed points lie on the other side of the middle from the original fixed point. If a pitchfork bifurcation happens, it must occur at the middle of the phase space according to Proposition 4, and the two new fixed points must lie to either side of it.

Now assume that three fixed points do exist, and the parameters change so that one of them crosses the middle, that is, at $r = 0$, we have $\dot{r} = 0$. Plugging this assumption into the dynamical system in (8) implies that $q_1 = q_2$. Thus in this circumstance,

$$\dot{r}\big|_{q_2=q_1} = \frac{1}{2} r(4q_1 - 3 - b - (1-b)r^2),$$

so the other two fixed points must be at

$$\pm\sqrt{\frac{4q_1 - 3 - b}{1 - b}}.$$

Therefore, the only fixed point which can cross the middle of the phase plane is the central one. □

The complete set of bifurcation parameters can be found implicitly by building from Lemma 3 and using the discriminant. By definition, the discriminant of a polynomial is the product of the squares of the pair-wise differences of its roots, so it will be zero when a polynomial has a multiple root. The discriminant can be expressed entirely in terms of the coefficients of the polynomial. For a general cubic $a_3z^3+a_2z^2+a_1z+a_0$, the discriminant is

$$\frac{a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 + 18a_0a_1a_2a_3 - 27a_0^2a_3^2}{a_3^4}. \tag{11}$$

For $\dot{r}$, the discriminant is a large expression in terms of q_1, q_2, and b obtained by filling in this general formula. The bifurcation parameters are the values of q_1, q_2, and b which make this expression zero, and that surface may be plotted implicitly, as shown in Figure 2. According to the picture, the surface consists of two curved surfaces which meet in a spine where $q_1 = q_2 = (3+b)/4$ (the pitchfork bifurcation). The bottom corner is at $q_1 = q_2 = 3/4$, $b = 0$, and the rest of the surface appears to lie in the region $q_1, q_2 > (3+b)/4$. The important thing to notice is that if q_1 and q_2 are both close to 1, that is, under the surface, then the dynamical system has three hyperbolic fixed points. Above the surface, there is a single hyperbolic fixed point, and on the surface, there are one or two fixed points with at least one non-hyperbolic. The closer b is to 1, the larger q_1 and q_2 must be to be under the surface.

4. Two grammars and two universal grammars

In this section, we analyze a two-dimensional, asymmetric instance of the language dynamical equation. It models the following scenario: Suppose the population has a universal grammar U_1 which generates exactly one grammar G_1; learning and communication are both perfect. Under what circumstances could the population shift in favor of a new universal grammar U_2 which generates G_1 plus an additional grammar G_2? That is, within this model, when is it advantageous to have a choice between two grammars? The analysis builds heavily on the results from Section 3.

4.1. ***Parameter settings.*** The dependent variables of interest are $x_{1,1}$, $x_{1,2}$, and $x_{2,2}$. The variable $x_{2,1}$ represents the part of the population which

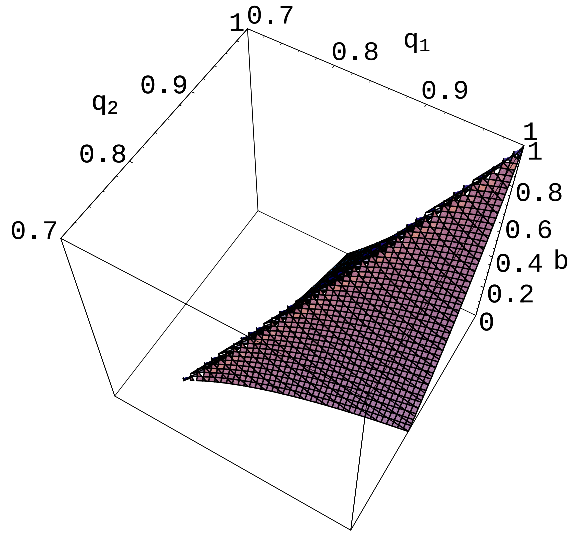


Figure 2. Bifurcation surface. Observe that in this picture, the q_1 and q_2 axes run only from 0.7 to 1. On the tent-shaped surface, there are one or two fixed points with at least one non-hyperbolic. Above the surface, the system has one hyperbolic fixed point, and below, it has three.

speaks G_2 but has universal grammar U_1, and by assumption, this is zero. Thus, the dynamical system in this case is in three variables with two degrees of freedom and can therefore be analyzed as a planar system.

As in Section 3, the A matrix only enters the dynamical system through the B matrix, as in (1), and since B is a symmetric matrix, there is really only one degree of freedom in choosing A. So, we define

$$b = \frac{a_{1,2} + a_{2,1}}{2}, \tag{12}$$

and allow this to be the one free parameter determined by the overlap between G_1 and G_2. The most general form for the learning algorithm matrix Q is

$$Q_{i,j,1} = \begin{pmatrix} 1 & 0 \\ * & * \end{pmatrix}, \qquad Q_{i,j,2} = \begin{pmatrix} q_1 & 1-q_1 \\ 1-q_2 & q_2 \end{pmatrix}, \tag{13}$$

which has two degrees of freedom. The entries filled with $*$ are always multiplied by $x_{2,1}$ which is assumed to be 0, so they do not matter. Thus this model has a total of three free parameters: b, q_1, and q_2, all of which are assumed to lie strictly between 0 and 1.

4.2. *Geometric analysis of the dynamics.* With these parameter settings, and the fact that $x_{1,2} = 1 - x_{1,1} - x_{2,2}$, the dynamical system (3) simplifies to

$$\begin{aligned}\dot{x}_{1,1} =& -(1-b)x_{1,1}x_{2,2}(2x_{2,2}-1),\\ \dot{x}_{2,2} =& 1 - q_1 + (-1+q_1)x_{1,1}\\ &+ (-3 + b + q_1 + (1-b)q_1 + bq_2 + (-1+b)(-1+q_1)x_{1,1})x_{2,2}\\ &+ (-2(-1+b) + (-1+b)(-1+q_1) + q_2 - bq_2)x_{2,2}^2\\ &+ 2(-1+b)x_{2,2}^3.\end{aligned} \tag{14}$$

It lives on the three-vertex simplex S_3, that is, a triangle. The vertices correspond to $x_{j,K} = 1$ and will be labeled $X_{j,K}$ in diagrams.

From here, a fairly complete understanding of the bifurcations of this system can be derived from some simple calculations and geometric considerations. To begin, we will find lines along which $\dot{x}_{1,1} = 0$, and the vector field is therefore parallel to the base of the simplex. These are called $x_{1,1}$ *null-clines*. From (14), it is clear that $\dot{x}_{1,1}$ is zero in three places: the lines $x_{1,1} = 0$, which is the base of the simplex, and $x_{2,2} = 0$, which is the left edge, and the line $x_{2,2} = 1/2$, which runs across the simplex. In particular, the base line $x_{1,1} = 0$ is invariant under this vector field. See Figure 3.

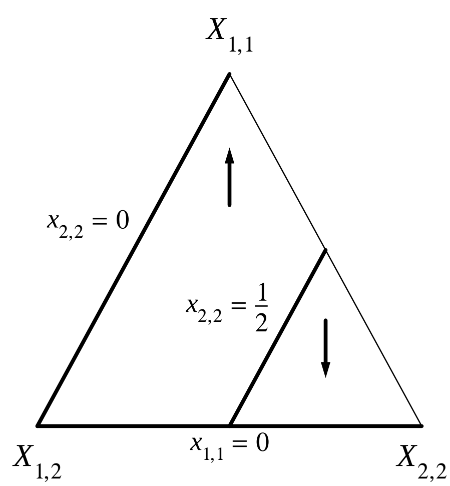


Figure 3. The simplex, with null-clines. The bold lines indicate where $\dot{x}_{1,1} = 0$. The arrows indicate the sign of $\dot{x}_{1,1}$ in the regions in between, up for positive, down for negative.

Several fixed points are easily located. Observe that if $x_{2,2} = 0$ then $\dot{x}_{1,1} = 0$ and $\dot{x}_{2,2} = (1 - q_1)(1 - x_{1,1})$. So the apex is the only fixed point on the left side of the simplex. Also, since the vector field always points upward toward it, it is stable. Another fixed point may be located on the cross line by substituting $x_{2,2} = 1/2$ into (14) yielding

$$\begin{aligned} \dot{x}_{1,1}|_{x_{2,2}=1/2} &= 0, \\ \dot{x}_{2,2}|_{x_{2,2}=1/2} &= \frac{1}{4}(1+b)\left(q_2 - q_1 - 2(1-q_1)x_{1,1}\right), \end{aligned} \tag{15}$$

from which we find that

$$(x_{1,1}, x_{2,2}) = \left(\frac{q_2 - q_1}{2(1-q_1)}, \frac{1}{2}\right)$$

is the unique fixed point on the line $x_{2,2} = 1/2$. It is located inside the simplex for $q_2 \geq q_1$ and outside otherwise. Observe that the vertical component of the vector field is upward above this fixed point, and downward below it, so it must be unstable. The horizontal component of the vector field to its right points leftward, and to its left it points rightward, indicating that locally, orbits flow toward the fixed point from either side. Thus, this fixed point is a saddle.

Consider the base line, which is invariant under this vector field and may therefore be partially analyzed in isolation. It is exactly the same as the general two-grammar problem from Section 3, and must look like one of the phase portraits in Figure 1, except that those pictures show only stability or instability in the horizontal direction. Stability of one of these fixed points in the vertical direction is determined by which side of the cross line it lies on: $\dot{x}_{1,1}$ is positive on the left side, indicating instability, and negative on the right side, indicating stability.

We must determine where the fixed points in Figure 1 may lie with respect to the point $(x_{1,1}, x_{2,2}) = (0, 1/2)$, which we do by examining the behavior of the saddle point on the cross line $x_{2,2} = 1/2$. The key fact is that the vector field on the cross line points leftward above the saddle point, and rightward below it, and changes direction only at that fixed point. Observe that the vector field at the upper right end of the cross line $(x_{1,1}, x_{2,2}) = (1/2, 1/2)$ is $(\dot{x}_{1,1}, \dot{x}_{2,2}) = (0, -(1/4)(1+b)(1-q_2))$, which points leftward. The direction of the vector field at $(x_{1,1}, x_{2,2}) = (0, 1/2)$ is either left or right, depending on the configuration of fixed points on the base line. If it points to the left, then the fixed point on the cross line must lie outside the simplex because the vector field must point left along the entire segment of the cross line within the simplex. Similarly, if the vector field points to the right at $(0, 1/2)$, then the the saddle point must lie inside the simplex. From previous analysis, the saddle point lies inside the simplex if and only if $q_2 \geq q_1$, so we have a link between the values of q_1 and q_2 and the phase portraits in Figure 1.

Now we must determine how the saddle point crosses the base line into the simplex. It must pass through the point $(x_{1,1}, x_{2,2}) = (0, 1/2)$. Substituting this point into (15), we see that the parameter values which cause this must satisfy $q_1 = q_2$. As it crosses the base line, it must coincide exactly with one of the fixed points there. Since the saddle point passes through the collision, the fixed points must cross in a transcritical bifurcation. To determine which fixed point is crossed, we substitute $q_2 = q_1 = q$ into the dynamical system in (14) and examine the base line. Note that $x_{1,1} = 0$ so $\dot{x}_{1,1} = 0$. Also:

$$\dot{x}_{2,2}\big|_{q_2=q_1=q, x_{1,1}=0} = (-1 + 2x_{2,2})\left(-1 + q + (1-b)x_{2,2} + (-1+b)x_{2,2}^2\right). \tag{16}$$

The roots of this cubic correspond to the fixed points on the base line; they are

$$\frac{1}{2} \quad \text{and} \quad \frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{4q-3-b}{1-b}}. \tag{17}$$

We now get three cases. If $q > (3+b)/4$, then there are three fixed points as in Figure 1 (e), one exactly in the middle and two to either side. If $q < (3+b)/4$, then there is one fixed point, exactly in the middle as in Figure 1 (a). If $q = (3+b)/4$, then there is one degenerate fixed point exactly in the middle as in Figure 1 (c), in which case the pitchfork and transcritical bifurcations happen simultaneously. At any rate, the saddle point can only enter the simplex by passing through the central fixed point on the base line.

The parameter space breaks up into four regions as shown in Figure 4. The tent-shaped surface is the same as the one in Figure 2. Above it, there is one fixed point on the base line. Below it, there are three fixed points on the base line. On the faces, there are two fixed points, one non-hyperbolic, and on the edge, there is one non-hyperbolic fixed point. The vertical plane separates the regions where $q_1 < q_2$ from the regions where $q_2 < q_1$. The complete bifurcation scenario is shown in Figure 5. The fixed points on the base line are constrained by Propositions 4, 5, and 6, so the cases shown are the only possibilities. Phase portraits in Figure 5 are labeled according to which part of the parameter space in Figure 4 they represent.

4.3. ***Competition between the universal grammars.*** The bifurcation scenario depicted in Figure 5 can be analyzed in terms of competition between the two universal grammars. The structurally stable pictures are (a), (c), (g), and (i); these are the ones that occur generically. Observe that in (a), there is only one stable fixed point, and it occurs at the apex of the triangular phase space. All interior orbits will approach this fixed point. Thus, in the case where $q_2 < q_1$ and both are fairly small, U_1 dominates. In (c), there are two stable fixed points, the one at the apex corresponding to a takeover by U_1 and the one on the base line corresponding to a takeover by U_2. Their basins of attraction are separated by the stable manifold of the saddle point on the cross line. Approximations to their basins of attraction

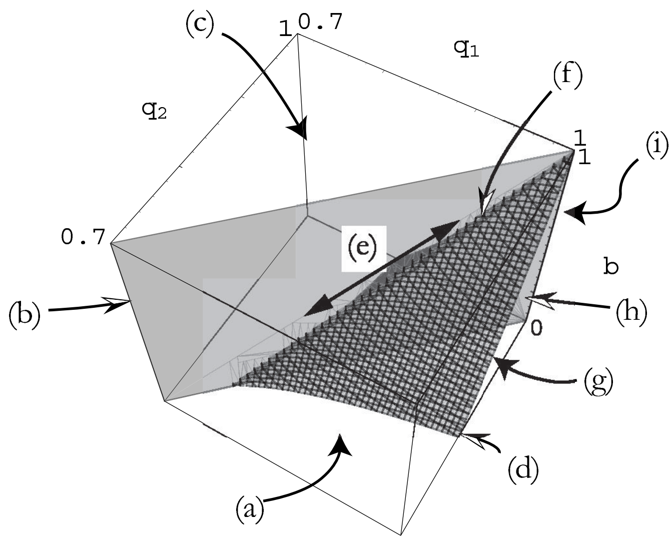


Figure 4. Parameter space with bifurcation surfaces. Areas (a), (c), (g) and (i) are regions in space, indicated by bold arrows. Areas (b), (d), (f), and (h) are the surfaces that separate those regions, indicated by black and white arrows. Areas (d) and (f) are the front and back surfaces of the tent. Area (b) is the part of the plane above the tent, and (h) is the part below it. Area (e) is a line where the two curved surfaces and the vertical plane intersect, indicated by thin arrows.

can be found by drawing a dashed horizontal line through the saddle. Orbits can only cross the left-hand segment of the dashed line by going upward, and the upper segment of the cross line by going leftward, which means these two segments bound a trapping region containing the apex. Similarly, orbits can only cross the right-hand segment of the dashed line by going downward, and the lower segment of the cross line by going rightward, which means there is another trapping region containing the sink on the base line. In this case, where $q_1 < q_2$ and both are fairly small, there is competitive exclusion between the two universal grammars. The most direct transition from (a) to (c) is a transcritical bifurcation passing through (b). Shortly after this bifurcation, the saddle point will be very close to the base line, so the trapping region for U_2 will be quite small. As q_2 increases, the saddle point moves upward and trapping region expands. A similar situation exists in (i), the difference being that the base line contains two other fixed points which affect a negligible fraction of the phase space. The situation is slightly different in (g). Again, there are two stable fixed points, but the saddle point whose stable manifold separates their basins of attraction is on the base line rather than on the cross line. There does not seem to be a simple trapping region that approximates the basins of attraction in this picture.

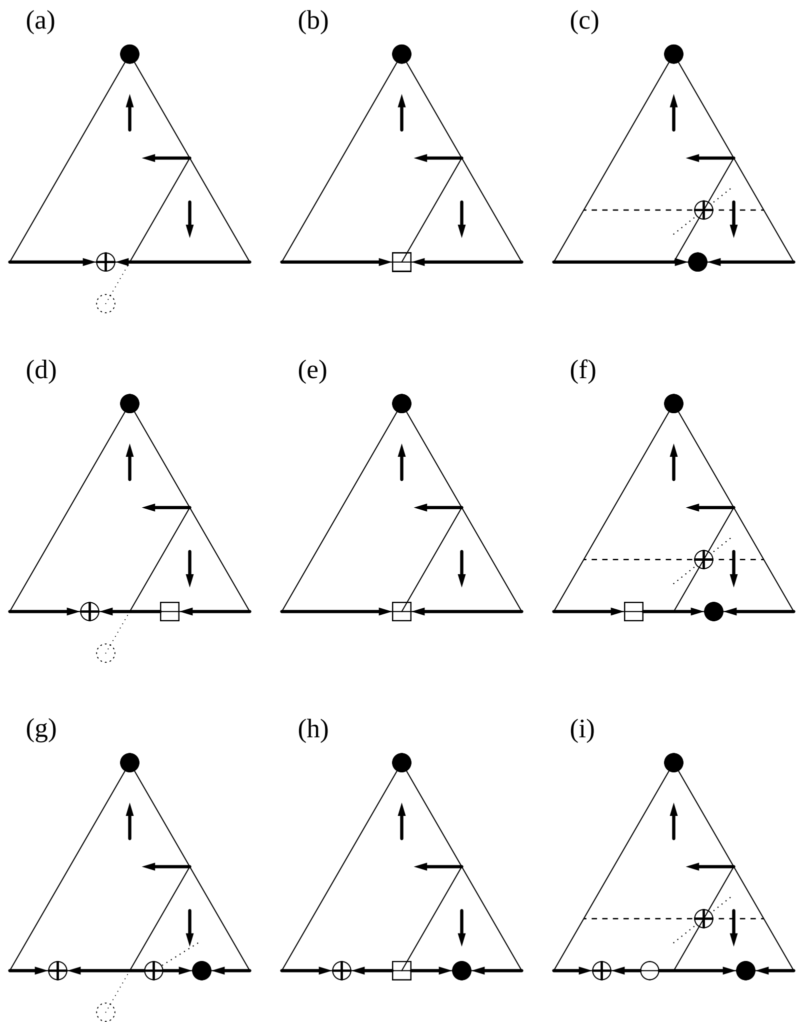


Figure 5. Phase portraits for the selection dynamics between U_1 (apex) and U_2 (base line). U_1 admits G_1, and U_2 admits G_1 and G_2. Either U_1 dominates as in (a), or there is bistability between U_1 and U_2. The parameters for each picture come from the region of the same label in Figure 4. Key: $\bullet$ indicates a sink, $\oplus$ indicates a saddle, $\circ$ indicates a source, $\Box$ indicates a non-hyperbolic fixed point. Arrows indicate (roughly) the direction of the vector field. In pictures (c), (f), and (i), the cross line and the horizontal dashed line through the saddle point define approximate upper and lower trapping regions for the two sinks. The actual boundary between their basins of attraction is the stable manifold of the saddle point, which is sketched as a dotted line. Picture (g) also contains such a boundary.

4.4. *Discussion of Section 4.* To summarize, the scenario examined in this section generically contains instances where U_1 dominates, and instances where there is competitive exclusion, but none where U_2 dominates or where both universal grammars coexist. Furthermore, U_2 can only take over if $q_2 > q_1$ as in pictures (c) and (i), or if q_1 and q_2 are both close to 1 as in picture (g). In the first case, G_2 is acquired more accurately than G_1, so it has an advantage and tends to increase in the population thereby putting U_1 at a disadvantage. In the second case, it appears that although G_1 may be learned more reliably than G_2, the learning reliability of G_2 is sufficiently high that it can maintain a large portion of the population through "market share" effects, again putting U_1 at a disadvantage. Observe that in any case, U_2 can only take over the population through G_2. A population of U_2 people speaking G_1 can be invaded by U_1. This is an illustration of a process by which a valuable acquired trait can become innate. This effect suggests that human universal grammar may have once allowed many more possible grammars than it does now, and that as portions of popular grammars became innate, UG became more restrictive.

5. A multi-grammar UG competing with single-grammar UGs

In this section, we will examine cases in which a UG with multiple grammars competes with a number of UGs that have only a single grammar each. We will begin by building on the results from Section 4 in two ways, extending that analysis to symmetric cases in an arbitrary number of dimensions.

5.1. *The case of full competition.* Let us extend the case from Section 4 by assuming that there are three universal grammars. The first, U_1, allows only G_1. The second, U_2, allows only G_2. The third, U_0, allows both G_1 and G_2. Since there is one single-grammar UG for each possible grammar, this case will be called *full competition.* We would like to determine whether one of these UGs can take over the population.

This situation contains two copies of the case from Section 4, one in which everyone uses U_0 or U_1, and a second in which everyone uses U_0 or U_2. From the former, there is generically no stable equilibrium in which U_0 takes over with a majority of people speaking G_1. From the latter, there is generically no stable equilibrium in which U_0 takes over with a majority of people speaking G_2. If U_0 is to take over, either G_1 or G_2 must be in the majority, so it follows that U_0 is unable to take over.

This result extends to an arbitrary number of grammars as follows. Let the grammars be G_1 to G_n, and assume there are universal grammars U_i which specify only the grammar G_i. Assume there is an additional UG U_0 which allows any of the n grammars. As a simplification, assume that the grammars are fully symmetric and unambiguous, that is, $A_{i,i} = 1$ and

$A_{i,j} = a$ for $i \neq j$. The parameter a is required to be strictly between 0 and 1. For reasons that will become clear in a moment, the learning matrix Q is allowed to be fully general except that no grammar is allowed to have perfect learning under U_0, that is, $Q_{i,i,0} < 1$ for all i.

We will need the following new notation. We are interested in determining if one universal grammar out of the U_K can take over the population, and if so, which one. We therefore define

$$y_K = \sum_{j=1}^{n} x_{j,K} \tag{18}$$

to be the total population with U_K. The dynamics for y_K can be expressed succinctly by using the fact that Q is row stochastic:

$$\begin{aligned} \dot{y}_K &= \sum_{j=1}^{n} \dot{x}_{j,K} \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} (F_i x_{i,K} Q_{i,j,K}) - \phi x_{j,K} \right) \\ &= \sum_{i=1}^{n} \left(F_i x_{i,K} \sum_{j=1}^{n} Q_{i,j,K} \right) - \phi \sum_{j=1}^{n} x_{j,K} \\ &= \sum_{i=1}^{n} F_i x_{i,K} - \phi y_K. \end{aligned} \tag{19}$$

We may further simplify the notation by introducing the variable

$$\phi_K = \sum_{i=1}^{n} F_i x_{i,K}, \tag{20}$$

from which it follows that $\phi = \sum_K \phi_K$ and

$$\dot{y}_K = \sum_{i=1}^{n} F_i x_{i,K} - \phi y_K = \phi_K - \phi y_K. \tag{21}$$

There is no explicit reference to Q in $\dot{y}_K$, although Q does influence the dynamics. It happens that the main result of this section does not depend on Q for exactly this reason.

Because of the symmetry imposed on A, the dynamics of the y_K simplify considerably. If we further define

$$v = (x_{1,0}, x_{2,0}, \ldots, x_{n,0}), \tag{22}$$

$$w = (x_{1,1}, x_{2,2}, \ldots, x_{n,n}), \tag{23}$$

then

$$\begin{aligned} \dot{y}_0 &= -(1-a)((v+w)\cdot w)y_0, \\ \dot{y}_K &= (1-a)(x_{K,0}+y_K-(v+w)\cdot(v+w))y_K \text{ where } K=1\ldots n. \end{aligned} \tag{24}$$

Note that the sum of the entries of v is equal to y_0, and the sum of the entries of w is equal to $1-y_0$.

Proposition 7. *The multi-grammar universal grammar, U_0, is always unstable, that is, if $y_0<1$, then $\dot{y}_0<0$. The single-grammar UGs are stable, meaning that for $K\geq 1$, if y_K is close to 1, then y_K is increasing.*

Proof. We will prove both statements by starting from a population that consists entirely of one UG, and perturbing it by converting ε of the population to another UG.

To prove the first statement, suppose that $y_0=1-\varepsilon$. All the entries of v and w are greater than or equal to zero, so $v\cdot w\geq 0$. Since w must be non-zero, it follows that $\dot{y}_0=-(1-a)(v\cdot w+w\cdot w)y_0<0$. In fact, in any population state where not everyone has U_0, the number of people with U_0 will decrease. Thus, U_0 is unstable and cannot take over the population.

To prove the second statement, fix $K\geq 1$ and assume that $y_K=1-\varepsilon$. Observe that

$$\begin{aligned} (v+w)\cdot(v+w) &= \sum_{i=1}^{n}(x_{i,0}+x_{i,i})^2 \\ &= (x_{K,0}+1-\varepsilon)^2+\sum_{\substack{i=1\ldots n,\\ i\neq K}}(x_{i,0}+x_{i,i})^2. \end{aligned}$$

The summation is over $n-1$ terms, each of which greater than or equal to zero, and their sum is fixed at $1-(1-\varepsilon)-x_{0,K}$. Therefore, the summation is at most $(\varepsilon-x_{0,K})^2$. (See Lemma 8.) It follows that

$$\begin{aligned} \dot{y}_K &\geq (1-a)(1-\varepsilon)\left(1-\varepsilon+x_{K,0}-(1-\varepsilon+x_{K,0})^2-(\varepsilon-x_{K,0})^2\right) \\ &= (1-a)(1-\varepsilon)(\varepsilon-x_{K,0})(1-2(\varepsilon-x_{K,0})). \end{aligned}$$

As long as $x_{K,0}<\varepsilon$, we have $\dot{y}_K>0$. This will continue to be true as $x_{K,0}\leq y_K$.

If $x_{K,0}=\varepsilon$, that is, $x_{K,0}$ accounts for the entire perturbation, then we need the assumption that under U_0, no language is learned perfectly. So, a short time later, $x_{K,0}$ will decrease as some children will have mistakenly learned another grammar, say G_h, so $x_{h,0}>0$. At this point, we will have a new perturbation with $y_K=\varepsilon'$ and $x_{K,0}<\varepsilon'$, and it follows that $\dot{y}_K>0$. □

The following lemma is used to make approximations in this and other proofs in this section.

Lemma 8. *Suppose that for $i = 1 \dots m$, we have numbers $\alpha_i \geq 0$ such that $\sum_i \alpha_i = \sigma$. Then*

$$\frac{\sigma^2}{m} \leq \sum_{i=1}^{m} \alpha_i^2 \leq \sigma^2.$$

Proof. Consider $\alpha = (\alpha_i)_{i=1}^m$ as a vector in $\mathbf{R}^m$. It is contained in a simplex because the sum of its entries is fixed. The point on the simplex closest to the origin is the center, corresponding to $\alpha_i = \sigma/m$ for all i, and this point yields the lower bound. The vertices of the simplex are the farthest points from the origin, corresponding to $\alpha_j = \sigma$ and $\alpha_i = 0$ for all $i \neq j$, and these points give the upper bound. □

Proposition 7 implies that UGs with many grammars are unable to compete directly with UGs that specify only one grammar.

5.2. ***The case of limited competition.*** The two-dimensional case from Section 4 illustrates a situation where a multi-grammar UG can have a stable equilibrium where a majority of the people use a grammar that does not occur as part of a single-grammar UG. We now turn our attention to a different extension of this case in which there are two UGs, U_0 which specifies $G_1, \dots, G_n$, and U_1 which specifies only G_1. As before, the A matrix is assumed to be fully symmetric, with all diagonal entries $A_{i,i} = 1$ and all off-diagonal entries $A_{i,j} = a$. The Q matrix disappears again, and we need only the assumption that no grammar is learned perfectly under U_0. By using the fact that $y_1 = x_{1,1} = 1 - y_0$, the model can be reduced to one differential equation of interest,

$$\begin{aligned} \dot{y}_0 &= (1-a)(-x_{1,1} - x_{1,0} + 2x_{1,1}x_{1,0} + M_2)x_{1,1} \\ &= (1-a)(-1 + y_0 - x_{1,0} + 2(1-y_0)x_{1,0} + M_2)(1-y_0). \end{aligned} \tag{25}$$

(Recall the definition $M_k = \sum_j \sum_K x_{j,K}^k$.) There is a fixed point at $y_0 = 0$, as can be seen by substituting this state into the differential equation. Furthermore, $\dot{y}_0 = 0$ when $y_0 = 1$, so the model can have trapping regions and stable fixed points in the subset of states which satisfy $y_0 = 1$. We are interested in determining when these various states are stable under perturbations.

Proposition 9. *The fixed point $y_0 = 0$, corresponding to a takeover by U_1, is stable.*

Proof. Consider a small perturbation, $y_0 = \varepsilon$. Then we must have $y_1 = x_{1,1} = 1 - \varepsilon$, and the differential equation satisfies

$$\begin{aligned}\dot{y}_0 &= (1-a)\left(-1+\varepsilon - x_{1,0} + 2(1-\varepsilon)x_{1,0} + (1-\varepsilon)^2 + \sum_{j=1}^{n} x_{j,0}^2\right)(1-\varepsilon) \\ &\le (1-a)(-1+\varepsilon + x_{1,0}(1-2\varepsilon) + 1 - 2\varepsilon + \varepsilon^2 + \varepsilon^2)(1-\varepsilon),\end{aligned}$$

where we have used Lemma 8 to bound the summation by ε^2. This expression factors into

$$\dot{y}_0 \le -(1-a)(\varepsilon - x_{1,0})(1-2\varepsilon).$$

If the perturbation is such that $x_{1,0} < \varepsilon$, then the right hand side is negative, and as $x_{1,0} \le y_0$, it will remain negative, so y_0 will shrink to 0.

If the perturbation is such that $x_{1,0} = \varepsilon$, then we must use the fact that under U_0 there is no perfect learning. After a short time, some other part of the population with U_0, say, $x_{h,0}$, will be non-zero due to learning error. This new perturbation will have $y_0 = \varepsilon'$ and $x_{1,0} < \varepsilon'$, and as before y_0 will shrink to 0. □

The following results show that U_0 can still take over, but not with G_1. It states that if $x_{1,0}$ is small enough, then a population consisting only of people with U_0 that is perturbed by adding a small number of people with U_1 will recover, at least in the short term.

Proposition 10. *Let $\varepsilon > 0$ be small and suppose $y_0 = 1 - \varepsilon$ and $x_{1,1} = \varepsilon$. Define $\kappa = 1/n - x_{1,0}$. If $\kappa > \varepsilon/(1-2\varepsilon)$, then $\dot{y}_0 > 0$.*

Proof. From the differential equation,

$$\begin{aligned}\dot{y}_0 &= (1-a)\left(-\varepsilon + (2\varepsilon - 1)x_{1,0} + \varepsilon^2 + \sum_{j=1}^{n} x_{j,0}^2\right)\varepsilon \\ &\ge (1-a)\left(-\varepsilon + (2\varepsilon-1)x_{1,0} + \varepsilon^2 + \frac{(1-\varepsilon)^2}{n}\right)\varepsilon,\end{aligned}$$

where we have once again used Lemma 8 to bound the summation. By substituting $x_{1,0} = 1/n - \kappa$, the inequality can be simplified to

$$\dot{y}_0 \ge (1-a)\left(\kappa - \varepsilon(1+2\kappa) + \left(1 + \frac{1}{n}\right)\varepsilon^2\right)\varepsilon.$$

The assumption that $\kappa > \varepsilon/(1-2\varepsilon)$ is equivalent to $\kappa > \varepsilon(1+2\kappa)$, so the right hand side is positive. □

The tricky part about interpreting this proposition is that a population state with $y_0 = 1$ might sill be unstable in the long term: It could move within the constraint $y_0 = 1$ to a state where $x_{1,0} > 1/n$, at which point Proposition 10 no longer applies and a perturbation can cause the population to be taken over by U_1, as this next proposition illustrates.

Proposition 11. *Let $\varepsilon > 0$ be small and suppose $y_0 = 1 - \varepsilon$ and $x_{1,1} = \varepsilon$. If $x_{1,0} > 1/2 - \varepsilon$, then y_0 is decreasing.*

Proof. For this proof, we use Lemma 8 to bound the summation in $\dot{y}_0$ from above,

$$\sum_{j=1}^{n} x_{j,0}^2 = x_{1,0}^2 + \sum_{j=2}^{n} x_{j,0}^2 \leq x_{1,0}^2 + (1 - \varepsilon - x_{1,0})^2.$$

This bound yields the inequality

$$\dot{y}_0 \leq 2\varepsilon(1-a)(x_{1,0} - (1-\varepsilon)) \left(x_{1,0} - \left(\frac{1}{2} - \varepsilon \right) \right).$$

If $1-\varepsilon > x_{1,0} > 1/2-\varepsilon$, then $\dot{y}_0$ is negative, and the perturbation will draw the population away from the region where $y_0 = 1$, indicating instability.

If $x_{1,0} = 1 - \varepsilon$, then we resort to the argument that a short time later, the population will change due to learning error to a different perturbation where $y_0 = 1-\varepsilon'$ and some other sub-population $x_{h,0} > 0$. Now $x_{1,0} < 1-\varepsilon'$, which implies that $\dot{y}_0 < 0$ and the population is moving away from the region where $y_0 = 1$. □

5.3. *Some remarks about these results.* Several remarks are in order. First, the Q matrix has mostly disappeared, so Propositions 7, 9, 10, and 11 hold regardless of the learning mechanism under U_0, except that it must not be perfect. In fact, it could be dynamic, depending on the population state for example, as long as it remains row stochastic.

Second, the fact that some of the propositions declare $y_0 = 1$ to be "stable" may be misleading. As noted before, the population could start in a state where $y_0 = 1$ and move within that constraint to a state in which y_0 begins to decrease. The simplest behavior for which $y_0 = 1$ would be truly stable is for the population to converge to a stable fixed point that satisfies Proposition 10, but it could also converge to a limit cycle or to a strange attractor, depending on what behaviors are available to a population restricted to U_0.

We can get more definite results from these propositions if we add assumptions that ensure that all population states with $y_0 = 1$ tend to fixed points. Any fixed points that are stable when only U_0 is allowed and that also fall under Proposition 10 are stable with respect to all perturbations, including those involving the introduction of U_1. Any such fixed points that fall under Proposition 11 are unstable. Some may be outside the hypotheses of both propositions, and we can say nothing more about them here.

A full bifurcation analysis of the fully symmetric case of the language dynamical equation with one universal grammar is worked out in (Mitchener, 2002). To apply those results here, we must add the assumption that for the learning matrix for U_0, all diagonal elements $Q_{i,i,0} = q$ and all off-diagonal elements $Q_{i,j,0} = (1-q)/(n-1)$. It follows from (Mitchener, 2002) that if only U_0 is present, then all populations tend to fixed points. The analysis shows that there is a constant $\hat{q}_1$,

$$\hat{q}_1 = \frac{2(n-1)(2+a(n-3)) + (a-1)n - 2(n-1)\sqrt{(1+a(n-2))(n-1)}}{(a-1)(n-2)^2}, \tag{26}$$

such that if $q < \hat{q}_1$, then the only stable fixed point in a population restricted to U_0 is one in which every grammar is represented equally. Thus, $x_{1,0} = 1/n$ and that fixed point is potentially unstable to perturbations involving U_1 because Proposition 10 does not apply. On the other hand, if $q > \hat{q}_1$, then there are n stable fixed points, and each G_j is used by a large part of the population in exactly one of them. These are called the *1-up fixed points* in (Mitchener, 2002) and *single grammar fixed points* in (Komarova *et al.*, 2001). At the one where G_1 has the majority, $x_{1,0} > 1/n$, so it does not fall under Proposition 10 and is potentially unstable to perturbations involving U_1. If q is sufficiently large, this fixed point moves so that $x_{1,0}$ approaches 1, so at some value of q, it will exceed 1/2. Then Proposition 11 will apply and the fixed point will definitely be unstable. At the other fixed points, $x_{1,0} < 1/n$, so they fall under Proposition 10 and are therefore stable. In short, if the learning process in U_0 is sufficiently reliable, that is $q > \hat{q}_1$, then U_0 can take over the population in a stable manner, but not through G_1. If learning is unreliable, then U_1 will eventually take over.

6. Ambiguous grammars

In this section we will generalize the case in Section 4 not by adding dimensions but by allowing the grammar specified by U_1 to be different from both of those specified by U_2, and by allowing the grammars to be ambiguous. The diagonal entries of A are allowed to be less than one. This case can exhibit a greater variety of behavior than was seen in Section 4, including stable coexistence of both universal grammars, and dominance by U_2. This form of the language dynamical equation has a total of eight free parameters. Rather than attempt a complete symbolic analysis, we will present one short proposition and some numerical results.

6.1. *Parameter values.* We assume that U_1 allows for one grammar G_1, and that U_2 allows for two grammars, G_2 and G_3. The Q matrix is allowed to be fully general,

$$Q_{i,j,1} = \begin{pmatrix} 1 & 0 & 0 \\ * & * & * \\ * & * & * \end{pmatrix}, \qquad Q_{i,j,2} = \begin{pmatrix} * & * & * \\ 0 & q_2 & 1-q_2 \\ 0 & 1-q_3 & q_3 \end{pmatrix}. \tag{27}$$

The entries filled with $*$ are always multiplied by some $x_{j,K}$ that is restricted to be zero, so they do not matter. Also, the matrix A is allowed to be fully general; we even allow the diagonal elements to be less than 1. The only constraint we place on A is that since it appears in the model only through $B = (A + A^T)/2$ we may as well assume A is symmetric. There are eight free parameters, six from the upper half of A and q_2 and q_3.

6.2. *Analysis and phase portraits.* The expressions for $\dot{x}_{1,1}$, $\dot{x}_{2,2}$ and $\dot{x}_{3,2}$ are unwieldy so they will not be written out. However, it turns out that $x_{1,1} = 1$, $x_{2,2} = x_{3,2} = 0$ is a fixed point for all parameter settings. The one symbolic result is the following:

Proposition 12. *The fixed point* $x_{1,1} = 1$, $x_{2,2} = x_{3,2} = 0$ *is unstable if* $-2A_{1,1} + A_{1,2}q_2 + A_{1,3}q_3 > 0$.

Proof. We reduce the system to two dimensions by replacing $x_{3,2}$ by $1 - x_{1,1} - x_{2,2}$. The trace of the Jacobian matrix of the reduced system at the fixed point in question is $-2A_{1,1} + A_{1,2}q_2 + A_{1,3}q_3$. If this is positive, then at least one of the eigenvalues of the Jacobian must have positive real part (see (Strogatz, 1994), p. 137). □

Roughly what this proposition means is that if G_1 is sufficiently ambiguous, and G_2 and G_3 are similar to it and can be learned reliably, then U_1 is unable to take over the population. This situation seems unrealistic, however, there is at least one reasonable interpretation. Suppose that G_1 is close to the union of G_2 and G_3, and contains many sentences that can be interpreted so as to have multiple meanings. Suppose further that many of these sentences are in G_2 or G_3 but with a single meaning. Thus U_2 has an advantage because it restricts its people to some less ambiguous language at the expense of imperfect learning, and this may be enough to destabilize a population where everyone has U_1. Proposition 12 is a mathematical expression of this situation. Note that when $A_{1,1}$ is restricted to be 1, the proposition never applies, and the stability of the U_1 fixed point must be determined by other means.

A number of phase portraits for a variety of parameter values are drawn in Figures 6–9 based on numerical computations. In particular, these phase portraits illustrate that with this general model, it is possible to have stable coexistence of U_1 and U_2, and it is possible for U_2 to dominate. Neither of these situations is possible in the limited case analyzed in Section 4.

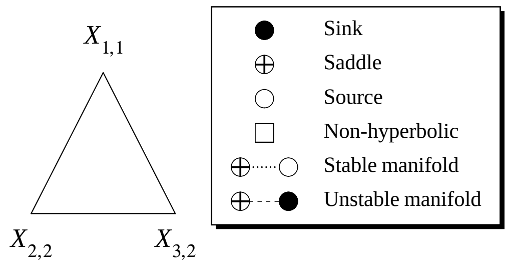
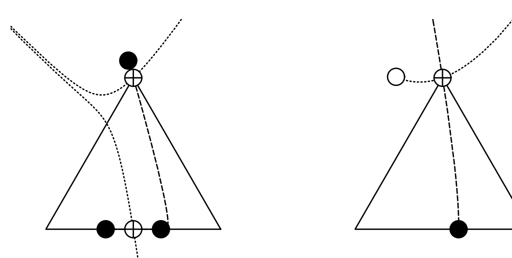


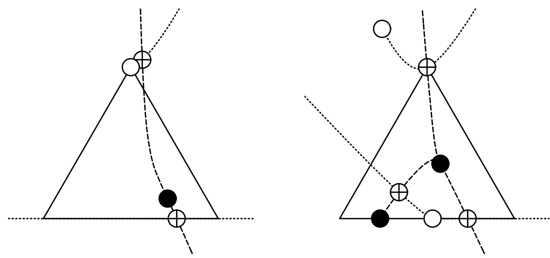
Figure 6. Key to phase portraits show in Figures 7 to 9. Some fixed points outside the simplex have been drawn for reference. The three corners of the triangle represent population states where everyone uses a single language, as indicated. The apex of the triangle represents $U_1 = 1$ and the base represents $U_2 = 1$.



$$\text{(a)}\ A = \begin{pmatrix} 0.66 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.7 \\ 0.7 & 0.7 & 0.9 \end{pmatrix}, \qquad \text{(b)}\ A = \begin{pmatrix} 0.45 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.7 \\ 0.7 & 0.7 & 0.9 \end{pmatrix},$$

$$q_2 = 0.95, q_3 = 0.95 \qquad q_2 = 0.8, q_3 = 0.85$$

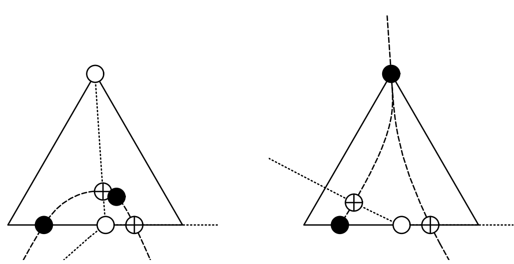
Figure 7. Two instances where U_2 dominates.



(a) $A = \begin{pmatrix} 0.4 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.2 \\ 0.7 & 0.2 & 0.9 \end{pmatrix}$, (b) $A = \begin{pmatrix} 0.5 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.2 \\ 0.7 & 0.2 & 0.9 \end{pmatrix}$,

$q_2 = 0.8, q_3 = 0.85$ $q_2 = 0.86, q_3 = 0.85$

Figure 8. Two instances of stable coexistence. In (b), U_2 can also take over, but only with G_2.



(a) $A = \begin{pmatrix} 0.3 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.2 \\ 0.7 & 0.2 & 0.9 \end{pmatrix}$, (b) $A = \begin{pmatrix} 0.7 & 0.6 & 0.7 \\ 0.6 & 0.9 & 0.2 \\ 0.7 & 0.2 & 0.9 \end{pmatrix}$,

$q_2 = 0.87, q_3 = 0.85$ $q_2 = 0.87, q_3 = 0.85$

Figure 9. Another instance of stable coexistence and an instance of exclusion.

7. Conclusion

The evolution of universal grammar is based on genetic modifications that affect the architecture of the brain and the classes of grammars that it can learn. At some point in the evolutionary history of humans, a UG emerged that allowed the acquisition of language with unlimited expressibility. In principle, UG can change as a consequence of random drift (neutral evolution), as a by-product of selection for other cognitive function, or under selection for language acquisition and communication. The third aspect is what we consider in this paper.

We explore some low-dimensional cases of natural selection among universal grammars. In particular, we study the competition between more specific and less specific UGs. Suppose two universal grammars, U_1 and U_2 are available, and U_2 admits two grammars, G_1 and G_2, while U_1 admits only G_1. If learning within U_2 is too inaccurate, then U_1 dominates U_2: For all initial conditions that include both U_1 and U_2, U_1 will eventually out-compete U_2. If learning within U_2 is sufficiently accurate, then for some initial conditions U_2 will win while for others U_1 will win; there is competitive exclusion. Note that accurate learning stabilizes less specific UGs. We can also find coexistence of two different UGs. We provide such an example where U_1 admits G_1 and U_2 admits G_2 and G_3.

A standard question in ecology is concerned with the competition between specialists that exploit a specific resource and generalists that utilize many different resources (May, 2001). Similarly, here we have analyzed competition between specialist UGs that admit few grammars and generalist UGs that admit many candidate grammars. This is an interesting similarity. There is also a major difference: In ecology the more individuals exploit a resource the less valuable this resource becomes, but in language the more people use the same grammar the more valuable this grammar becomes. Hence, the frequency dependency of the fitness functions work in opposite directions in the two cases.

The question that we want to understand ultimately is the balance between selection for more powerful language learning mechanisms that allow acquisition of larger classes of complex grammars, and selection for more specific UGs that limit the possible grammars. This paper provides mathematical machinery and a first step toward this end.

References

Ahlfors, L. V. (1979). *Complex Analysis.* Third edn. McGraw-Hill.

Aitchinson, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon.* Oxford: Basil Blackwell.

Andronov, A. A., E. A. Leontovich, I. I. Gordon, and A. G. Maier (1971). *Theory of Bifurcations of Dynamic Systems on a Plane.* Jerusalem: Keter Press.

Bickerton, D. (1990). *Language and Species.* Chicago: University of Chicago Press.

Cangelosi, A., and D. Parisi (eds) (2001). *Simulating the Evolution of Language.* Springer-Verlag.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1972). *Language and Mind.* New York: Harcourt Brace Jovanovich.

Ferrer i Cancho, R., and R. V. Solé (2001a). The small world of human language. *Proceedings of the Royal Society of London B*, **268**, 2261–2266.

Ferrer i Cancho, R., and R. V. Solé (2001b). Two regimes in the frequency of words and the origin of complex lexicons: Zipf's law revisited. *Journal of Quantitative Linguistics*, **8**, 165–173.

Ghazanfar, A. A., and M. D. Hauser (1999). The neuroethology of primate vocal communication: substrates for the evolution of speech. *Trends in Cognitive Sciences*, **3**(18), 377–384.

Gibson, E., and K. Wexler (1994). Triggers. *Linguistic Inquiry*, **25**, 407–454.

Gold, E. M. (1967). Language Identification in the Limit. *Information and Control*, **10**, 447–474.

Grassly, N., A. von Haesler, and D. C. Krakauer (2000). Error, population structure and the origin of diverse sign systems. *Journal of Theoretical Biology*, **206**, 369–378.

Hauser, M. D. (1996). *The Evolution of Communication.* Cambridge, MA: Harvard University Press.

Hauser, M. D., E. L. Newport, and R. N. Aslin (2001). Segmentation of the speech stream in a nonhuman primate: statistical learning in cotton-top tamarins. *Cognition*, **78**(3), B53–B64.

Hofbauer, J., and K. Sigmund (1998). *Evolutionary Games and Replicator Dynamics.* Cambridge University Press.

Hurford, J. R., M. Studdert-Kennedy, and C. Knight (eds) (1998). *Approaches to the Evolution of Language.* Cambridge University Press.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, **3**, 272–279.

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**(2), 102–110.

Komarova, N. L., P. Niyogi, and M. A. Nowak (2001). The evolutionary dynamics of grammar acquisition. *Journal of Theoretical Biology*, **209**(1), 43–59.

Krakauer, D. C. (2001). Kin imitation for a private sign system. *Journal of Theoretical Biology*, **213**, 145–157.

Lachmann, M., S. Szamado, and C. T. Bergstrom (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences, USA*, **98**, 13189–94.

Lai, C. S. L., S. E. Fisher, J. A. Hurst, F. Vargha-Khadem, and A. P. Monaco (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, **413**(6855), 519–523.

Lieberman, P. (1984). *The Biology and Evolution of Language.* Cambridge, MA: Harvard University Press.

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.

May, R. M. (2001). *Stability and Complexity in Model Ecosystems*. Princeton, NJ: Princeton University Press.

Mitchener, W. G. (2002). Bifurcation analysis of the fully symmetric language dynamical equation. *Journal of Mathematical Biology*. Submitted.

Niyogi, P. (1998). *The Informational Complexity of Learning*. Boston: Kluwer Academic Publishers.

Niyogi, P., and R. C. Berwick (1996). A Language Learning Model for Finite Parameter Spaces. *Cognition*, **61**, 161–193.

Nowak, M. A., and D. C. Krakauer (1999). The Evolution of Language. *Proceedings of the National Academy of Sciences, USA*, **96**, 8028–8033.

Nowak, M. A., J. Plotkin, and V. A. A. Jansen (2000). Evolution of Syntactic Communication. *Nature*, **404**, 495–498.

Nowak, M. A., N. L. Komarova, and P. Niyogi (2001). Evolution of universal grammar. *Science*, **291**, 114–118.

Nowak, M. A., N. L. Komarova, and P. Niyogi (2002). Computational and evolutionary aspects of language. *Nature*, **417**(6889), 611–617.

Pinker, S. (1990). *The Language Instinct*. New York: W. Morrow and Company.

Pinker, S., and A. Bloom (1990). Natural Language and Natural Selection. *Behavioral and Brain Sciences*, **13**, 707–784.

Ramus, F., M. D. Hauser, C. Miller, D. Morris, and J. Mehler (2000). Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys. *Science*, **288**(April), 349–351.

Strogatz, S. H. (1994). *Nonlinear Dynamics and Chaos*. Reading, MA: Perseus Books.

Studdert-Kennedy, M. (2000). Evolutionary implications of the particulate principle: Imitation and the dissociation of phonetic form from semantic function. *In:* Knight, C., J. R. Hurford, and M. Studdert-Kennedy (eds), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press.

Uriagereka, J. (1998). *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, MA: MIT Press.

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.

Wexler, K., and P. Culicover (1980). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.