Language evolution and George Price's "General Theory of Selection"

Gerhard Jäger

University of Bielefeld

`Gerhard.Jaeger@uni-bielefeld.de`

May 2008

Abstract

In the early seventies, the bio-mathematician George Price developed a simple and concise mathematical description of evolutionary processes that abstracts away from the specific properties of biological evolution. In the article it is argued that Price's framework is well-suited to model various aspects of the cultural evolution of language. The first part of the paper describes Price's approach in some detail. In the second part, two case studies about its application to language evolution are presented.

1 Language evolution

Ever since the development of the evolutionary model in biology in the mid-nineteenth century, people have noted a certain affinity of the evolutionary logic and the development of natural languages. The following well-known citation from Darwin's *The descent of man* perfectly captures this intuition:

> "The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel. ...Max Müller has well remarked: 'A struggle for life is constantly going on amongst the words and grammatical forms in each language. The better, the shorter, the easier forms are constantly gaining the upper hand, and they owe their success to their inherent virtue.' To these important causes of the survival of certain words, mere novelty and fashion may be added; for there is in the mind of man a strong love for slight changes in all things. The survival or preservation of certain favoured words in the struggle for existence is natural selection." (Darwin 1871:465f.)

During the twentieth century, the theory of evolution in biology underwent a stunning development. The insights of genetics into the mechanism of heredity clarified the nature of the replication process, leading to the so-called "modern synthesis" of Darwinism with Mendelian genetics (or "Neo-Darwinism", as it is sometimes called). Mathematical frameworks like population genetics or evolutionary game theory led to very exact quantitative models that can be tested empirically with high precision.

In stark contrast to these developments in biology, the evolutionary perspective on natural languages largely remained a metaphor in linguistics during most of the last century. This is undoubtedly due to the fact that the predominant structuralist paradigm focuses on synchronic descriptions of languages rather than on their diachronic development. Also, generative grammar and related frameworks employ mathematical techniques from algebra and formal language theory, which are non-quantitative. As a consequence, languages appear to be discrete objects, while evolution requires a conceptualization of the domain of interest in terms of gradual differences and continuous change.

Nonetheless, the idea of language evolution has attracted a good deal of attention within the last ten years or so. There are at least three independent intellectual developments that led to this renaissance:

- Practical experience has shown that quantitative, statistical models of linguistic phenomena are by far more successful in computational linguistics then the more traditional approaches using discrete mathematics. As a result, quantitative models are taken seriously again in theoretical linguistics. These approaches lends themselves more readily for evolutionary modelling than the traditional algebraic framework (see for instance the work of Kirby 1999 or Wedel 2004).

- Evolutionary techniques are firmly established by now in neighbouring disciplines like artificial intelligence or artificial life. This serves as a source of inspiration for linguists with a background in computer science (like the work of Luc Steels and his co-workers, see for instance Steels 1996).

- Thanks to the work of popularizers like Richard Dawkins (Dawkins 1976) or Daniel Dennett (Dennett 1995), the idea of applying the Darwinian logic to cultural phenomena has gained some currency in the humanities in general. Various researchers from historical linguistics have taken up this approach (see for instance Croft 2000 or Ritt 2004).

Partially in parallel, there is also a revived interest in investigating the biological evolution of the human language faculty, as witnessed by publications like Pinker and Bloom (1990), Nowak et al. (2002), or Hauser et al. (2002).

So while there is a strong interest now in evolutionary approaches to linguistic issues, there is little consensus so far about how exactly language evolution should be conceptualized. The main topic of debate, as far as I can see, is the issue what are the *replicators* in language evolution. The term "replicator" (in the sense of a unit of evolution) was coined by Richard Dawkins in his 1976 book *The Selfish Gene.* According to Dawkins' view, the

basic unit of evolution in biology is the gene, the physical carrier of heritable information. Dawkins also argues that any evolutionary process must be based on a population of replicators, i.e. counterparts of genes. He actually invents a new term, "meme", as a unit of replication in cultural evolution.

If this logic is valid, the first step in developing a theory of language evolution is to identify the linguistic units of replication. This proves to be a surprisingly difficult task. There are essentially three modes of replication that play a role in the acquisition and usage of natural language:

1. the biological inheritance of the human language faculty,

2. first language acquisition, which amounts to a vertical replication of language competence from parents (or, more generally, teachers) to infants, and

3. imitation of certain aspects of language performance in language usage (like the repetition of words and constructions, imitation of phonetic idiosyncrasies, priming effects etc.)

It is fairly clear what replicators are for the biological evolution of the language faculty. Since this is just one aspect of biological evolution in general, the carriers of heritable information are of course the genes. For the other two aspects of language evolution, the question is not so easy to answer. What are replicators in iterated language acquisition — entire I-languages? Single Rules? Parameters? Lexical items? The same difficulties arise with respect to replication via language usage. Candidates for the replicator status are phonemes, morphemes, words, constructions etc., or single instances of them (i.e. features of utterances), or mental representations of such instances (so-called "exemplars") etc. A considerable amount of the recent literature on language evolution is actually devoted to foundational questions like this one.

The main point I want to make in this paper is that this issue is actually of little relevance in my view. For one thing, I tend to be sceptical about the usefulness of methodological discussions anyway. The proof of the pudding is in the eating — a certain approach is useful if (and only if) it leads to insightful analyses of linguistic facts. If this is missing, even the most sophisticated discussion of foundational issues will not make up for the lack of it. But quite apart from this general issue, I will try to argue that the programme for analysing cultural evolution that can be extracted from the work of George Price is perhaps better suited to conceptualize language evolution than Dawkins' memetics or related approaches that assume a very detailed analogy between the cultural and the biological sphere.

2 George Price's "General Theory of Selection"

George Price was certainly one of the more remarkable figures in twentieth century science, even though he has remained relatively obscure even in evolutionary biology, where he made several highly significant contributions. He was a trained chemist, but he dabbled

in many intellectual disciplines during his life, including computer science, economics, theology and political science. In 1967, by the age of forty-five, he turned his interest to evolutionary biology. Within the few years until his untimely death in 1975, he made at least three breakthrough discoveries there: he contributed decisively to the advent of evolutionary game theory (Smith and Price 1973), he developed the modern interpretation of R.A. Fisher's so-called "Fundamental Theorem of Natural Selection" (Price 1972b), and he developed the *Price equation*, a very simple and concise mathematical framework to describe evolution via natural selection (Price 1970, 1972a). A very recommendable short biography of this remarkable person is given by Schwartz (2000). Price's contributions to evolutionary biology are described in some detail by Frank (1995).

Around 1971, Price wrote a manuscript titled "The Nature of Selection". It was only published posthumously in 1995 (Price 1995). There he sketched a programme for a general theory of evolution (or "selection", as he calls it) which includes biological evolution in the neo-Darwinian sense but encompasses various other kinds of natural and cultural evolution as well. The abstract of the paper starts with:

> "A model that unifies all types of selection (chemical, sociological, genetical, and every other kind of selection) may open the way to develop a general 'Mathematical Theory of Selection' analogous to communication theory." (Price 1995:389)

The first paragraph of the paper deserves to be quoted in its entirety:

> "Selection has been studied mainly in genetics, but of course there is much more to selection than just genetical selection. In psychology, for example, trial-and-error learning is simply learning by selection. In chemistry, selection operates in a recrystallisation under equilibrium conditions, with impure and irregular crystals dissolving and pure, well-formed crystals growing. In palaeontology and archaeology, selection especially favours stones, pottery, and teeth, and greatly increases the frequency of mandibles among the bones of the hominid skeleton. *In linguistics, selection unceasingly shapes and reshapes phonetics, grammar, and vocabulary.* In history we see political selection in the rise of Macedonia, Rome, and Muscovy. Similarly, economic selection in private enterprise systems causes the rise and fall of firms and products. And science itself is shaped in part by selection, with experimental tests and other criteria selecting among rival hypotheses." (Price 1995:389, emphasis added)

Even though Price did not develop a theory of selection in the sense he probably envisioned it, the paper gives good arguments why the Price equation should be the cornerstone of such a theory. In the remainder of this section, I will recapitulate the main argumentation of Price's paper. The interested reader is of course referred to the original article, which is very readable and not overly technical.

4

First a note on terminology: Price's notion of "selection" is not completely identical to the notion of "evolution" that the present paper deals with. *Selection* can be a one-time process, starting with one state and terminating with a second state. *Evolution (via selection)* (as I use the term) is necessarily an iterated process, spanning several generations. Each generation step is one selection step in the Pricean sense. (Of course there are also notions of "evolution" that do not involve selection at all, which are not further considered here.)

This being said, let us turn to the main points of Price's paper. He distinguishes two senses of the term "selection" (which are both to be covered by a theory of selection). If you go to the marketplace and buy a few apples, you select among all the apples on sale those that you want to buy. This is *subset selection*, because the selected items form a subset of the set on which selection operates. The *Darwinian* notion of selection is different because selection operates on the parent generation while the selected items are the offspring (and these two sets are disjoint). Nevertheless, these two notions can be unified. In either case, we have two points in time, t (before selection takes place) and t' (after selection takes place). Furthermore, there is a set (or a "population") P of entities at t which selection operates on, and a set P' of selected items at t'. In case of subset selection, $P' \subseteq P$. In Darwinian selection, P is the parent generation and P' the offspring. In the example with selection of apples at the marketplace, P is the set of apples that are on sale, and P' is the set of apples that you buy.

Price points out that P and P' need not be finite sets. He also considers an example involving various chemical liquids that are filled from certain containers (time t) into other containers (time t'). Such non-atomic entities like liquids are usually mathematically modelled as infinite (in fact, continuous) sets. The central point for Price's notion of selection is that P and P' are *measurable* quantities. In case of finite sets, the most natural measure is just counting, but continuous measure functions like size, mass or volume, or even more abstract ones like probability, are also applicable. So P and P' are just two measurable quantities. Whatever measure function is applied, the number of items in P is denoted by w, and likewise w' is the number of items in P'. w and w' are non-negative real numbers.

Neither need P and P' be sets of objects of the same nature. Price considers Mussorgsky's creation of "Pictures of an Exhibition" as a case of selection. P is the set of paintings that the composer saw in the exhibition, and P' is the set of musical pieces that were inspired by paintings from P. Another example would be citations: P is a set of journal articles, and P' the set of references (in the sense of lines in the bibliography of some other journal article) to elements of P. Or P could be manuscripts and P' copies of the corresponding books and articles, etc. If P and P' are of a different nature, the measures that are used to obtain w and w' may of course be different. (For evolution via selection, this aspect is of little relevance because the selection process can only be iterated if P' is of the same nature of P.)

The next central ingredient of Price's theory is the idea that P is partitioned into a

disjoint portions or *bins*. So technically we have a family of quantities $p_1, \ldots, p_n$, such that

$$P = \bigcup_{i \le n} p_i \tag{1}$$

$$p_i \cap p_j = \emptyset \text{ if } i \ne j \tag{2}$$

In the apple example, the obvious partition would be the one where each bin contains exactly one apple. But other partitions are possible as well—like partitioning the apples according to size, or to color, or to price.

To take another example that does not involve subset selection: let P be the set of genes at a certain locus in the parent generation, and P' the corresponding set in the offspring generation. Then P could be partitioned into single molecular copies of the gene in question, i.e. each bin contains one DNA molecule. Alternatively, one might partition P according to alleles. In the latter case, we have few bins, each containing many molecules.

The set P' is partitioned as well, into the same number of bins as P. So we have

$$P' = \bigcup_{i \le n} p_i' \tag{3}$$

$$p_i' \cap p_j' = \emptyset \text{ if } i \ne j \tag{4}$$

Intuitively, there should be a natural relation between the content of of some bin p_i' and the content of the corresponding bin p_i. Let us again consider subset selection, as illustrated with the apple buying scenario. If I buy the apple in p_i, then $p_i' = p_i$. If, however, I do not buy p_j, then $p_j' = \emptyset$ is just empty. So p_i' is always the set of apples from p_i that I bought.

In the Mussorgsky example, p_i' is the set of musical pieces that were inspired by paintings from p_i. As for citations, p_i' is the set of references to articles in p_i, etc. In the examples with genes, p_i' is the set of DNA molecules that are direct copies of some molecule in p_i (at the relevant locus). Of course p_i' may be empty, or it may contain many more molecules than p_i.

The latter example is instructive because gene copying is mostly but not always *faithful*. Suppose we partitioned P according to alleles.[1] Then all molecules in p_1 will be instances of the same allele — call it r, while all molecules in p_2 are instances of a different allele, say s. DNA copying may involve mutations from s to r and vice versa. Since p_1' contains exactly the copies of genes in p_1 (including the non-faithful copies), p_1' may contain s-alleles next to r-alleles. The crucial point here is that the partitioning of P' is induced by the partitioning of P and the copying relation, not by some independent criterion (even if such a criterion was used to partition P). In Price's own words (p. 392):

> "We will say that a set P' is a *corresponding set* to a set P if there exists a one-to-one correspondence such that, for each member p_i of P there is a corresponding member p_i' of P' which (if not empty) is composed partly or

---

[1] According to Merriam-Webster, an allele is "any of the alternative forms of a gene that may occur at a given locus."

wholly of the same material of p_i, or has been derived directly from p_i, or contains one or more replicas of p_i or some part of p_i, or has some other special close relation to p_i" (emphasis in the original)

(Note that Price does not cleanly distinguish between the set of objects P and the set of bins that jointly constitute P, and—in case p_i is a singleton set—between the set and its only element. No confusion should arise from this though.)

The measure functions that assigned the numbers w and w' to P and P' respectively can also be applied to the various bins of P and P'. So w_i is the amount that is in bin p_i, and w_i' is the amount in p_i'.

The quantity

$$f = \frac{w'}{w}$$

is the growth rate or *fitness* of the entire system. If w and w' are just the number of objects in P and P' respectively, this is the average number of "descendants" that an element of P has in P'. Likewise, we can consider bin-wise fitness

$$f_i = \frac{w_i'}{w_i}$$

which gives the average number of descendants of an object from p_i.

A transition from a set P to a set P' as described so far can only be described as involving *selection* (in a non-technical sense) if the correspondence between elements of P and elements of P' is not random. Rather, whether or not there are many descendants of p_i in p_i' (i.e., whether f_i is high or low) should be correlated with some features of the objects in p_i. Features that lead to high fitness are *selected for*. A smart apple buyer, for instance, will only select high-quality apples (which can be judged from color, surface texture etc.), so only bins containing high quality apples in P will have a non-empty corresponding set in P'. Slightly more technically, high quality apples have a higher fitness than low quality ones. In this case, there is selection for the quality of apples. Analogously, influential papers have a high fitness according to the citation scenario because their corresponding sets contain many citations. An allele has a high fitness if many copies of it are transmitted from generation to generation, etc.

It sounds plausible to say that apples are selected for their quality etc. However, Price's framework is purely quantitative, and therefore selection—in the technical sense— can only operate on quantifiable characters (like size of an apple, or its price, its weight, the percentage of its surface which is red, ...). So let us say that there is is some function μ that measures some quantifiable trait of the objects in P (like total weight, total price etc.). Likewise μ' measures a corresponding trait of objects in P'. Note that μ measures this feature in a cumulative way—$\mu(P)$ is the *total* weight (size, price, ...) of all objects in P taken together. What we are actually interested in though is the *average* value of some objects from P (P') under μ (μ'). Let us call these average values x (before selection) and

x' (after selection). They are computed by the formulas:

$$x = \frac{\mu(P)}{w} \tag{5}$$

$$x' = \frac{\mu'(P')}{w'} \tag{6}$$

If P or P' are non-discrete sets, x and x' can be interpreted as something like the average density or concentration of the character μ. The average value of μ can also be calculated for the separate bins:

$$x_i = \frac{\mu(p_i)}{w_i}$$

$$x'_i = \frac{\mu'(p'_i)}{w'_i}$$

(If $w'_i = 0$ for some i, we simply stipulate that $x'_i = 0$ to make sure that this term is always defined.)

Suppose μ is a feature that is usually correlated with high fitness—like a large amount of red skin of an apple. Then objects from P with a high value under μ will have many counterparts in P'. Let us also assume that the degree of μ is passed on with little change from objects in P to their counterparts in P'. In this case, we expect that the average value of μ in P'—x'—will be higher than x. Conversely, if a high value of μ usually goes with low fitness, x' will be smaller than x. The *degree of change* of the average value of μ is notated as Δx. It is defined as

$$\Delta x = x' - x$$

Again this can also be calculated for each bin separately, so we have

$$\Delta x_i = x'_i - x_i$$

Using these definitions, the dynamics of selection can concisely be expressed by the *Price equation*:

$$f\Delta x = Cov(f_i, x_i) + E(f_i \Delta x_i). \tag{7}$$

We will have to look at each term of this equation separately. The term on the left hand side, $f\Delta x$, is the difference in the average value of μ between t' (after selection) and t (before selection), multiplied by fitness. So the entire equation is a *difference equation*. (This fact is a bit blurred by the fact that the difference of interest, Δx, is multiplied with f. However, we could as well divide both sides by f to obtain a canonical difference equation.) For the more interesting case that P and P' are sets of the same kind of objects and selection is iterated many times, such an equation can be seen as an update rule that describes the transition from one point in time to the next. More specifically, the equation
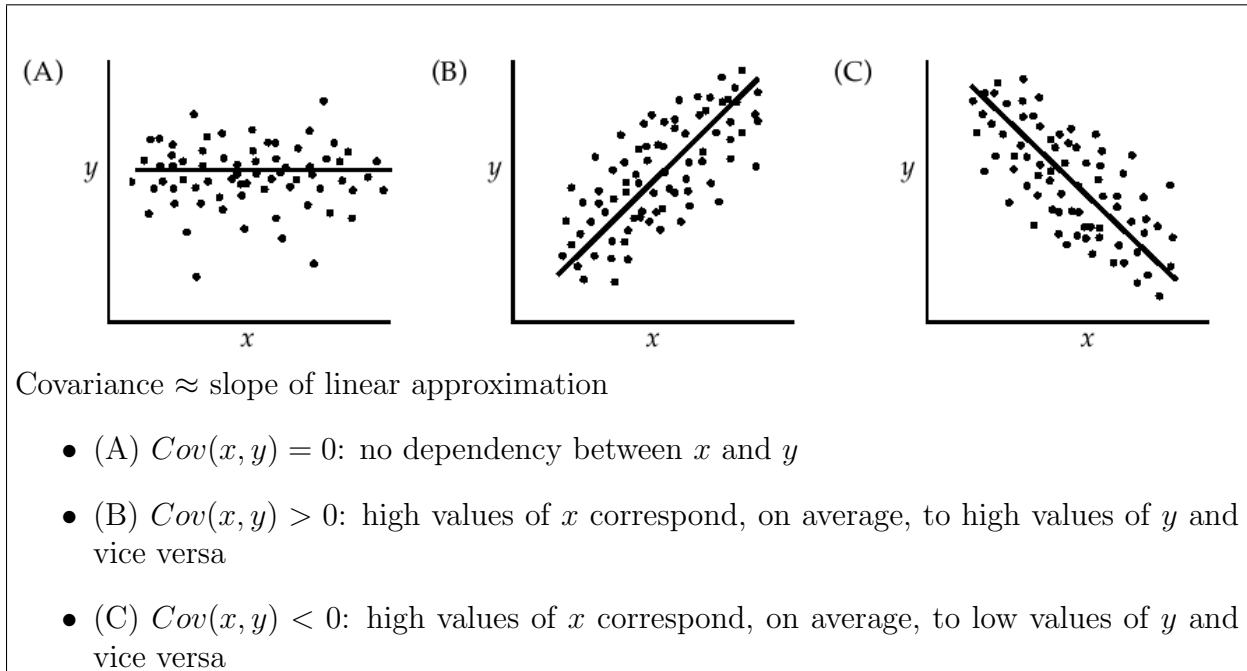
Figure 1: The concept of covariance

Covariance ≈ slope of linear approximation

- (A) $Cov(x, y) = 0$: no dependency between x and y

- (B) $Cov(x, y) > 0$: high values of x correspond, on average, to high values of y and vice versa

- (C) $Cov(x, y) < 0$: high values of x correspond, on average, to low values of y and vice versa

tells us how the value of x evolves over time. It is important to notice that it is up to the modeller to decide which quantitative character they want to study. x could be the average number of children that a speaker of Esperanto has, but it can also be the probability that a random sentence has OV word order. So essentially the Price equation can be used to study the evolution of any quantitative character, provided the dynamics of the system is well-understood. This, of course, is the difficult part, which is captured by the right hand side of the equation.

The first term, $Cov(f_i, x_i)$, stands for the *covariance* between fitness and the value of x_i, as compared between the different bins. Intuitively, the covariance of two random variables measures how strongly these two variables vary together. If the variables in question are independent of each other, the covariance is 0. If high values of the first variable tend to co-occur with high values of the second variable and vice versa, the covariance, is positive, and if high values of one variable mostly co-occur with low values of the other variable, covariance is negative. This is graphically illustrated in figure 1. Technically the covariance of the two random variables X and Y is defined as

$$Cov(X_i, Y_i) \;\; = \;\; (\sum_i p(i) X_i Y_i) - (\sum_i p(i) X_i)(\sum_i p(i) Y_i), \qquad (8)$$

where the index i gives the number of an event and $p(i)$ the probability of this event.

In the Price equation, the events in question are the bins $1, \ldots, n$. If some object is picked out from P at random, the probability that it comes from $p(i)$ is $\frac{w_i}{w}$. So spelled out,

the first term on the right hand side expands to

$$Cov(f_i, x_i) \;=\; (\sum_i \frac{w_i}{w} f_i x_i) - fx. \tag{9}$$

(I made tacit use from the fact that the average fitness is f, and the average value of x_i is x.)

If bins with a high average value of μ, i.e., with a high x_i, tend to have many offspring (meaning: large correspondence sets in P') and vice versa, there is a positive correlation between x_i and f_i. This means that $Cov(f_i, x_i)$ is positive. Conversely, if high values of x_i are correlated with low fitness, $Cov(f_i, x_i)$ is negative. This is intuitively unsurprising—if a certain character is correlated with high fitness, we expect it to be strongly represented in the next generation, i.e. Δx should be positive. Likewise, if a high value of μ is an indicator of low fitness, the average value of μ will decrease over time. So the first term on the right hand side simply covers the essential intuition of Darwinism: characters that lead to high fitness will spread in the population.

This logic only works if the degree of μ is passed on faithfully from parents to offspring. Price's framework does not require copying fidelity though. Unfaithful reproduction of μ is dealt with by the second term, $E(f_i \Delta x_i)$. $E(X)$ is the expectation value—i.e., the weighted average—of the random variable X. So the second term of the Price equation can be spelled out as

$$E(f_i \Delta x_i) \;=\; \sum_i \frac{w_i}{w} f_i \Delta x_i. \tag{10}$$

What is the value of Δx_i? We start with the apple example again. Suppose p_i contains a single apple that is selected. Then $p_i' = p_i$, and $x_i' = x_i$, hence $\Delta x_i = 0$. If, however, the apple in p_i is not selected, both $x_i' = 0$ and $f_i = 0$. Hence $f_i \Delta x_i = 0$ as well. So in this scenario, $E(f_i \Delta x_i) = 0$. Here each "surviving" object in P' is exactly identical to its counterpart in P, hence replication is entirely faithful. Let us consider an example where this is not the case.

Suppose there are two alleles of some gene, s and t. The fitness of s-genes is exactly 2, i.e. each s-gene has exactly two offspring. The fitness of the t-allele is only 0.5—only every second copy has one offspring on average. Also, suppose that there are 300 copies of the gene in question in P, 100 s-alleles and 200 t-alleles. So we have

$$\begin{aligned}
w &= 300 \\
w' &= 300 \\
w_s &= 100 \\
w_t &= 200 \\
w_s' &= 200 \\
w_t' &= 100 \\
f &= 1
\end{aligned}$$

10

$$
\begin{aligned}
f_s &= 2 \\
f_t &= \frac{1}{2}
\end{aligned}
$$

Note that w_s' is not necessarily the number of s-alleles in the offspring generation, but the number of offspring that have an s-parent! In fact, we assume that only t-alleles are entirely faithfully reproduced. An s-allele, however, has a 50% chance to mutate into an t-allele.

Let $\mu = \mu'$ be the function that counts the number of s-alleles in a set. Then we have

$$
\begin{aligned}
\mu(P) &= 100 \\
\mu(p_s) &= 100 \\
\mu(p_t) &= 0 \\
x &= \frac{1}{3} \\
x_s &= 1 \\
x_t &= 0 \\
\mu'(p_s') &= 100 \\
\mu'(p_t') &= 0 \\
x' &= \frac{1}{3} \\
x_s' &= \frac{1}{2} \\
x_t' &= 0 \\
\Delta x &= 0 \\
\Delta x_s &= -\frac{1}{2} \\
\Delta x_t &= 0
\end{aligned}
$$

This means that the first term of the Price equation is positive:

$$
Cov(f_i, x_i) = \frac{1}{3}
$$

x has a high value—1—within the s-subpopulation, and a low value—0—within the t-subpopulation. Also, s-alleles have a higher fitness than t-alleles. So high values of x go with high fitness and vice versa. Hence the covariance is positive.

The second term is negative though:

$$
E(f_i \Delta x_i) = -\frac{1}{3}
$$

This is so because the value of x may change from 1 to 0 under mutation, but not vice versa. So the average difference between x' and x under replication is negative. In the

example the number are chosen in such a way that P and P' actually have an identical composition and the two terms on the right hand side cancel each other out.

In this example, the covariance term covers the change of x that is due to selection, while the second term takes care of changes that are due to mutation. Under this interpretation, the Price equation is one way to express the so-called *replication-mutation dynamics* that is well-studied in theoretical biology (but mostly by different analytical means, especially by the so-called *quasi-species* model).

Depending on how the P-population is split into bins, the two terms of the equation may also receive a different interpretation though. Suppose there are two types of individuals in a species, altruistic ones and selfish ones. These are genetically determined traits. Altruistic individuals, by definition, behave in such a way that their interaction partners profit from the interaction, i.e. they receive a higher fitness than without the interaction. The altruists themselves do not profit from the interaction. (We may think of the two strategies of the well-known Prisoner's Dilemma in this context.)

Now suppose the population is structured into groups in such a way that everybody only interacts with partners within the same group. One possible application of the Price equation here is to identify the partition cells p_i with those groups. Then f_i is the fitness, i.e. the average reproduction rate, of an entire group. Let us furthermore assume that there are no mutations whatsoever and reproduction is asexual, so altruistic parents will have altruistic offspring only, and selfish parents selfish offspring.

If a group p_i contains many altruistic individuals, many of its member will benefit from interaction with the altruists and receive a high fitness. So f_i will be high. Conversely, a group consisting mainly of egoists will receive a low average fitness. So if x_i is the relative frequency of altruists within group p_i, high values of x_i go with high fitness and vice versa. So the first term of the equation, $Cov(f_i, x_i)$, is positive. This term measures *selection between groups*. The fact that it is positive means that groups with many altruists have an advantage over groups with many egoists.

The second term, however, measures *selection within groups*. On average a selfish individual will have more offspring than an altruistic one. So the percentage of altruists within one group will decrease from generation to generation. Hence Δx_i (and therefore also $f_i \Delta x_i$) will be negative for all i, which means that $E(f_i \Delta x_i)$ is also negative. So the qualitative interpretation here is that between-group selection favors altruism, while within-group selection favors selfishness. Depending on the relative strength of these two forces, the term on the left hand side will be positive, zero or negative. If between-group selection is strong enough, altruism may actually survive in the long run.

In biology, the Price equation is used mainly to study this kind of competition between group level and individual level selection. However, as the previous example illustrated, the framework is general enough to accommodate very diverse scenarios.

If this framework is used to model population dynamics, we are forced to assume discrete time steps, and one generation is completely replaced by the next generation when going from t to t'. A more realistic model would assume that some small portions of the population reproduce with small time intervals. If the population is large enough, this can legitimately be approximated by a process with continuous time. In Price (1972a) it is

shown that this continuous time approximation can be described by the following version of the Price equation:

$$\dot{E}(x) \quad = \quad Cov(f_i, x_i) + E(\dot{x}_i) \tag{11}$$

The expression $\dot{\phi}$ gives the first derivative of some variable ϕ against time. Intuitively this is the rate of change of that variable. If $\dot{\phi} = 0$, the value of ϕ does not change. $\dot{\phi} > 0$ means that the value of ϕ increases and vice versa.

Equation (11) is a *differential equation*—actually one of a particularly well-behaved kind, because it is an ordinary autonomous differential equation. By choosing different quantitative traits for μ/x, we can use (11) to set up a system of such differential equations that describes the dynamics of the domain which is modelled. The theory of this kind of equations is well-understood. In many cases it is even possible to solve them analytically, and even if this is not possible, there are established techniques to predict the qualitative long-run behaviour of the dynamical system in question. Also, there are good numerical algorithms to study the behaviour of such a system.

It is important to appreciate that the Price equation (both the discrete time version and the continuous time version) is a tautology. It follows directly from the assumptions about P and P', their correspondence relation, the partition into bins etc. Even though the derivation of the actual equation is not straightforward (and will not be explained here—the interested reader is referred to literature, for instance to Frank (1995)), it does not add any new information. The equations simply makes the implicit assumptions of the model explicit.

The main reason that I find Price's approach appealing for studying language evolution is not that it leads to systems of ordinary differential equations (even though this is certainly an asset). Rather, it imposes a certain intellectual discipline which I think is healthy. At the same time, it gives the modeller all the freedom that is needed to study a certain phenomenon, without enforcing a certain ontology that may be useful in one discipline but misplaced in another one. Let me spell out these two aspects in detail.

To apply Price's model—which means, in the end, to come up with an instance of his equation—the modeller has to be absolutely clear what is being modelled. There has to be absolute clarity about

- what the sets P and P' are,

- how P is partitioned into bins,

- what correspondence relation is assumed and what partition this relation imposes on P', and

- which quantitative character μ (or rather its average x) is being studied.

On the other hand, the approach is extremely flexible and general. P and P' can be any sets you like, as long as P' is, in whatever abstract sense, "later" than P and some well-defined correspondence relation can be established between the sets. P and P' even

need not consist of empirical objects. We could imagine a set of empirical objects M and its descendants N; and an abstract sample space Ω (which can for instance be identified with the interval $[0,1]$). Nothing prevents us from defining P as $M \times \Omega$ and P' as $N \times \Omega$. For instance, this would enable us to split an integral empirical object from M into disjoint fractions and to put them into different bins, which may be useful for a given application.

The only requirement on the correspondence relation is that it uniquely induces a partition of P' from a given partition of P. This is guaranteed if it is a function from P' to P.[2] The correspondence relation is the closest counterpart to the notion of "replication" in other conceptualizations of evolution. However, no specific requirements are made in Price's framework regarding the nature of correspondence. Neither need its domain and range (P' and P) be discrete sets with atomic elements ("replicators"), nor is there any requirement that there is any copying fidelity between correspondents. "Unfaithful" correspondence simply means that the second term (on the right hand side) of the equation is non-negligible. In principle this even covers dynamic systems where the elements are neither created nor destroyed but simply change states according to some transition probabilities (so-called Markov processes). In a Markov process, the fitness of each bin is a constant (namely 1), and the covariance term becomes 0. So the right hand side reduces to the second term.

One might object that such an extreme case cannot be called "evolution" anymore. It is certainly true that the covariance term of the equation pretty much captures the intuitive content of "evolution via replication and selection". However, reality does not care about the conceptual distinction between "evolutionary" and "non-evolutionary" processes, and it therefore strikes me as an advantage that Price's approach is tailored to capture the effects of selection and of non-faithful correspondence within one model.

3 Applications

In this section I will present two applications of the framework described above. My point here is not primarily to propose certain analyses but to give some examples of applications of the Price equation to demonstrate its versatility. The examples are essentially taken from the literature; its formalization in terms of the Price equation is new though.

3.1 Nowak's model of grammar evolution

In a series of publications, Martin Nowak and his co-workers developed a formal model of the evolutionary dynamics that is induced by iterated grammar acquisition (see for instance Komarova et al. 2001; Nowak et al. 2001). In the sequel I will use the version from Nowak (2006), chapter 13, as basis for discussion.

---

[2]To be perfectly precise, it has to be a measurable function under w/w' and μ/μ'. In the case of finite sets, every function is automatically measurable. For infinite measurable sets, measurability of the function means that the image of a measurable subset of P' is a measurable subset of P.

Grammars are culturally transmitted from the parent generation to infants. This is a form of replication, and thus a candidate for being the base of an evolutionary process. Nowak assumes that there are only finitely many grammars that are compatible with UG, $G_1, \ldots, G_n$.[3] The languages that are generated by these grammars need not be disjoint, but there may be a degree of mutual intelligibility of speakers of different grammars. Let a_{ij} be the probability that a sentence uttered by a speaker of grammar G_i is correctly understood by a speaker of grammar G_j. The chances that speakers of grammars G_i and G_j can communicate with each other (if both assume the roles of speaker and listener with equal probability) is then

$$F(G_i, G_j) \;\; = \;\; \frac{1}{2}(a_{ij} + a_{ji}) \tag{12}$$

Now suppose that the population is mixed, and that the number of speakers of grammar G_i is w_i, for all i. The total population size is $w = \sum_i w_i$. If everybody speaks with everybody else with equal probability, the chance of successful communication of a speaker of G_i are then

$$f_i \;\; = \;\; \sum_j \frac{w_j}{w} F(G_i, G_j) \tag{13}$$

(This is strictly speaking wrong because it assumes that people talk to themselves with the same probability as with everybody else. If the population is sufficiently large, this effect is negligible though, and the definition above is a licit simplification.)

The key assumption of Nowak's model is that communicative success immediately translates into fitness. Briefly put, an eloquent individual is supposed to have, on average, more offspring than a less eloquent one. That this is indeed the case seems plausible, given that eloquence raises social status, which in turn increases reproductive chances. Of course there are many other factors determining fitness, but it is legitimate to construct a simple model where the selective effects of communicative success are studied in isolation.

Nowak's assumption are somewhat stronger even, because the fitness differences induced by differential communicative success have to be quantitatively proportional to the differences in expected communicative success. This is perhaps too strong an assumption, but it is the null hypothesis as long as no more information about the quantitative relation between communicative and replicative success is available.

Ignoring the intricacies of sexual reproduction, the model assumes that each infant acquires its grammar from its parents. Grammars are discrete entities, and effects like bilingualism or language contact are not part of the model. Language acquisition may be imperfect though. There is an error matrix Q with the intended interpretation that a child acquiring its language from a parent using G_i will acquire G_j is Q_{ij}. If language acquisition is fairly accurate, Q_{ii} should be close to 1, but this assumption is not part of the model.

---

[3]This assumption is not uncontested, and franky, I consider it wrong because grammars are probalistic rather than algebraic entities. But it is nevertheless instructive to study the evolutionary consequences of the simplifying assumption of a finite grammar space.

Now all ingredients are in place to apply Price's analytical technique. The population before selection, P, is the set of speakers in the parent generation. P' is the children generation. We assume that every infant acquires its grammar from exactly one adult, so the the natural correspondence relation between a and b is: b acquired its grammar from a. An obvious way to split P into bins is given by the grammars of the speakers: p_i is the set of speakers in the parent generation using grammar G_i. This, together with the correspondence relation, induces a partition of P' as well. p'_i is the set of infants that acquired their grammar from a speaker of G_i. Note that the infants in p'_i need not all be speakers of G_i due to imperfect language acquisition.

Let $\mu^*(M)$ be the number of speakers of grammar G_{i*} within the set M. We define

$$x_i^* = \frac{\mu^*(p_i)}{w_i} \tag{14}$$

$$x_i'^* = \frac{\mu^*(p'_i)}{f_i w_i} \tag{15}$$

$$f = \sum_i \frac{w_i}{w} f_i \tag{16}$$

$$\Delta x_i^* = x_i'^* - x_i^* \tag{17}$$

Applied to the total population, Δx^* is the change of the relative frequency of G_{i*} within the population during the transition from the parent generation to the children generation.

The Price equation then says:

$$f \Delta x^* = Cov(f_i, x_i^*) - E(f_i \Delta x_i^*) \tag{18}$$

Let us look at the two terms on the right-hand side in turn.

The covariance term is defined as

$$Cov(f_i, x_i^*) = \sum_i \frac{w_i}{w} f_i x_i^* - f \sum_i \frac{w_i}{w} x_i^* \tag{19}$$

Now note that $x_{i*}^* = 1$ and $x_i^* = 0$ if $i \neq i^*$. So we can simplify to

$$Cov(f_i, x_i^*) = \frac{w_{i*}}{w} f_{i*} - f \frac{w_{i*}}{w} \tag{20}$$

$$= \frac{w_{i*}}{w}(f_{i*} - f) \tag{21}$$

The second term is defined as

$$E(f_i \Delta x_i^*) = \sum_i \frac{w_i}{w} f_i (x_i'^* - x_i^*) \tag{22}$$

Note that $x_i'^* = q_{ii*}$. So we can simplify to

$$E(f_i \Delta x_i^*) = \sum_i \frac{w_i}{w} f_i (q_{ii*} - x_i^*) \tag{23}$$

$$= \frac{w_{i^*}}{w} f_{i^*} (q_{i^*i^*} - 1) + \sum_{i \neq i^*} \frac{w_i}{w} f_i q_{ii^*} \tag{24}$$

$$= \sum_i \frac{w_i}{w} f_i q_{ii^*} - \frac{w_{i^*}}{w} f_{i^*} \tag{25}$$

Putting the two terms together, we get

$$f \Delta x^* = \frac{w_{i^*}}{w} (f_{i^*} - f) + \sum_i \frac{w_i}{w} f_i q_{ii^*} - \frac{w_{i^*}}{w} f_{i^*} \tag{26}$$

$$= \sum_i \frac{w_i}{w} f_i q_{ii^*} - \frac{w_{i^*}}{w} f \tag{27}$$

Without going into the details of the continuous time model, I just mention here that the version of the Price equation in this case looks almost identical, namely: Putting the two terms together, we get

$$\dot{x}^* = \sum_i \frac{w_i}{w} f_i q_{ii^*} - \frac{w_{i^*}}{w} f \tag{28}$$

This is exactly the *replicator-mutator equation* that Nowak uses in his model (even though he derives it in a different way). The bulk of Nowak's work about this model concerns conditions on learning precision that are necessary to guarantee stability of a coherent language in a population. For these investigations, the reader is referred to the original literature. My point in this subsection was to show how the Price framework can be used to model the iterated-learning notion of language evolution in a precise quantitative way, thereby taking the effects of imperfect learning into account.

3.2   Exemplar dynamics and blending inheritance

Exemplar based approaches to cognitive processing and representation, originally deriving from psychology, have gained high interest in several areas of linguistics in recent years (see for instance the articles in Gahl and Yu 2006, a special issue of *The Linguistic Review* on exemplar-based models in linguistics). The overarching idea of exemplar theory is that instances of linguistic events (both production and comprehension) are stored in a highly detailed fashion in memory. So even if two utterance tokens belong to the same type—like two utterances of the same word with the same meaning, they are memorized separately and in a detailed fashion. This may include specific information about phonetic parameters (like fundamental and formant frequencies of segments, their length etc.) as well as information about the specific syntactic context (i.e. cooccurrence with other lexical items) and the like. An exemplar is a detailed cognitive representation of an event, alongside with a categorization. So for instance, the exemplar of a vowel consists of a category (like the phoneme /a/) alongside with specific information about its phonetic representation like formant frequencies.

An important aspect of exemplar theories is that they assume a similarity metric over exemplars. Processing of new exemplars is based on analogy with similar exemplars. In

the simplest implementation of this idea, exemplars consist of points in an n-dimensional vector space, together with a category label. A new perceptual event is categorized by analogy with the closest exemplars in its neighborhood. Likewise, new exemplars of a given category are located in the neighborhood of previously stored exemplars of the same category.

Since new exemplars are generated in analogy to old ones, the dynamics of the exemplar "population" can be considered an evolutionary process, as has been observed by various authors like Batali (2002) and Wedel (2004). In this subsection I will analyse a very simple version of exemplar dynamics in detail. It is inspired by Pierrehumbert (2001).

Suppose exemplars are points in some n-dimensional vector space. We can think of them as phonetic events, where the dimensions are parameters like fundamental frequency, formant frequency, length, volume and the like (or rather their articulatory correlates). For simplicity, we only consider exemplars of one category, so the category labels can be ignored. In Pierrehumbert's model it is assumed that each exemplar has an activation level that decays over time. The impact of an old exemplar on new events is correlated with its activation, so that recent exemplars have the strongest impact. For simplicity's sake, I assume that the memory stores m-many exemplars, which all have the same activation level. If a new exemplar is added to the memory, one old item is picked out at random and removed instead.

In each cycle, a new exemplar is generated by picking out k-many exemplars from memory at random (with $k \geq 2$), and forming the average of this random sample. In a vector space, the average of k vectors is simply the arithmetic mean:

$$\mathrm{av}(s_1, \ldots, s_k) = \frac{1}{k} \sum_{i=1}^{k} s_i \tag{29}$$

Figure 2 gives some snapshots from a computer simulation of this model. Here the vector space has two dimension. The memory always contains $m = 100$ exemplars, and each sample consists of $k = 10$ items. At the beginning of the simulation, the memory is initialized with random exemplars (first picture). The cloud of exemplars constantly gravitates towards the centre, and after about 200 cycles, all items in memory are located in a small area. Pierrehumbert's simulations are somewhat more complex but show the same qualitative behaviour. She points out that this illustrates category formation via entrenchment. Even though the microscopic structure of the exemplar space is continuous (or at least very fine-grained), all exemplars of a category tend to gravitate towards the centre of the extension of this category, which leads to the emergence of categoricity at the macroscopic level.

This model is interesting for the issue of evolutionary modelling because it displays—as pointed out by Wedel (2004)—*blending inheritance*. Each new exemplar is causally connected to the sample of k old exemplars of which it is the arithmetic mean. So in a sense, each new exemplar has k parents. However, the features of the offspring are not faithfully inherited from any of the parents. Rather, they are a blend of the features of all parents. So this dynamics is qualitatively different from the neo-Darwinian conception
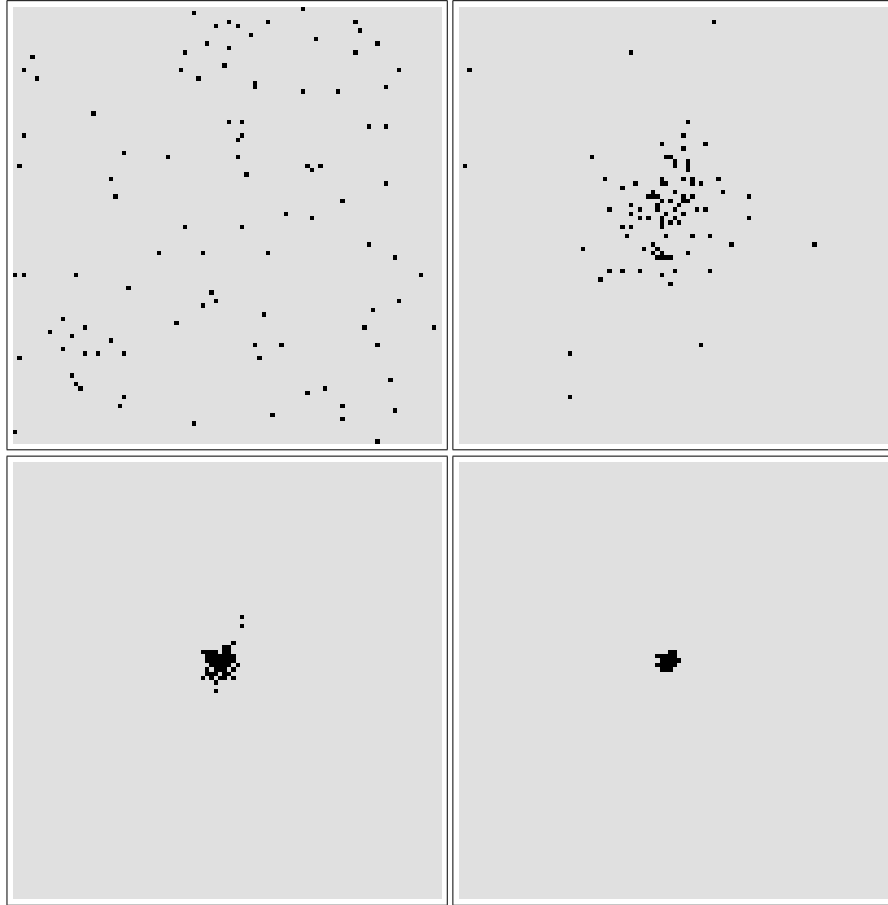
Figure 2: Simulation after 0, 90, 180, and 240 iterations

of biological evolution because there are no discrete units of inheritance that are faithfully passed on. Nonetheless, Price's framework is general enough to cope with this kind of evolution as well.

To fit this version of exemplar dynamics into Price's model, we have to decide what the sets P and P' are, what the correspondence relation is like etc. An obvious starting point is to say that t is a certain point in time when the memory contains a certain population of exemplars, and t' is the point in time when a new exemplar has been added (which replaces an old exemplar). It is not so clear so that the nature of the correspondence relation should be then. All the surviving old exemplars naturally correspond to themselves, but the new exemplar equally corresponds to all of its k "parents". Therefore we have to construct a slightly more abstract model. Let us say that P consists of k isomorphic copies of the memory at time t, and P' accordingly of k copies of the memory at t'. (If this is too abstract, you can also conceptualize this as splitting each exemplars into k equal parts.) Now each copy of the new exemplar corresponds to some copy of one of its parents—and each parent exemplar has a copy that corresponds to one copy of the offspring exemplar. All copies of old exemplars correspond to themselves. In this way each element of P' has

19

exactly one correspondent in P, as desired. For reasons that will become clear immediately, we furthermore assume that P' is not the set of exemplars one time step after t. Instead, we take P' to be the set of exemplars after g-many time-steps, where $g \geq 1$.

Since the dynamics that is investigated here is a random process, it is not possible to give deterministic expressions for f_i etc. All we know are probability distributions over possible trajectories. To obtain a deterministic dynamic system, we apply a common technique here: we let the size of all relevant sets grow to infinity and study the behaviour of the relevant parameters under this model inflation. Due to the law of large numbers, relative average values converge towards expected values. Since we know all probability distributions involved, we can actually calculate those values. So strictly speaking we do not study the behaviour of the actual, finite model, but the asymptotic properties of the model as the model size grows to infinity.

We will study the asymptotic behaviour of the exemplar model described if m, the number of exemplars, grows to infinity. We will also use two more parameters that grow to infinity, but at a slower rate:

- $b = \lceil m^{1/4} \rceil$ is the number of bins into which P is partitioned,[4] and

- $g = \lceil m^{1/2} \rceil$ is the number of time steps between t and t'.

In this way we ensure that $\lim_{m \to \infty} b, g = \infty$, but $\lim_{m \to \infty} \frac{b}{g} = \lim_{m \to \infty} \frac{g}{m} = 0$.

We assume that the n-dimensional vector space is partitioned into b cells. You can imagine this as an ever-finer grid that is laid over the vector space. p_i is the set of copies of exemplars at time t that are within the i-th partition cell. (We can assume, without restriction of generality, that each p_i is non-empty, because empty bins have a 0-weight and therefore have no impact on the exptectaction values that figure in the Price equation.) p'_i, accordingly, is the set of items that correspond to some item in p_i.

As P and P' are finite sets, the measure function w simply measures the cardinality of sets. So we have

$$ f_i \quad = \quad \frac{|p'_i|}{|p_i|}. $$

What is the cardinality of p'_i? The probability that the k copies of some exemplar from p_i do not survive within a single time step is $\frac{w_i}{w}$, so the expected number of items from p_i that do not survive is $\frac{kgw_i}{w}$. The probability that an exemplar from p_i is used to spawn a new exemplar (of which one copy belongs to p'_i) within a single round is $\frac{kw_i}{w}$. So the expected number of new items in p'_i that belong to newly created exemplars is $\frac{kgw_i}{w}$ as well. Hence the expected change in cardinality from p_i to p'_i is 0. Therefore the expected fitness f_i for each i is 1, and, using the notation $\alpha \to \beta$ for "α converges to β if m grows to infinity", we have

$$ Cov(f_i, x_i) \quad \to \quad 0, \tag{30} $$

---

[4] $\lceil x \rceil$ is the smallest integer $\geq x$.

no matter what parameter x we care to consider—the covariance of a constant with any random variable is always 0. So the entire dynamics is encapsulated in the second term of the Price equation this time.

The parameter x can be anything we like, so we start with studying the arithmetic mean c of all exemplars in memory. If v_j is the j-th exemplar, we have

$$c \;=\; \frac{1}{m}\sum_{j=1}^{m} v_j. \tag{31}$$

In physical terms, this can be interpreted as the centre of gravity of the memory, if each exemplar has the same mass. Let $c(d)$ be the coordinate of c at the d-th dimension.

Let us say that $p_i(d)$ is the arithmetic mean of the d-th dimension of the objects in p_i (which are all n-dimensional vectors). Likewise, $p_i'(d)$ is the average of the d-th coordinates of the copies in p_i'. (If $p_i' = \emptyset$, we stipulate that $p_i'(d) = 0$.) We define

$$
\begin{aligned}
x(d) &= E(p_i(d)) & (32)\\
x'(d) &= E(p_i'(d)) & (33)\\
\Delta x(d) &= x'(d) - x(d). & (34)
\end{aligned}
$$

It follows directly from the definition of c that $x(d)$ is the d-th coordinate of c at time t, and $x'(d)$ is the d-th coordinate of c at time t'.

The discrete-time Price equation now becomes

$$
\begin{aligned}
\Delta x(d) &= \sum_i \frac{w_i}{w}(p_i'(d) - p_i(d)) & (35)\\
&= \sum_i \frac{w_i}{w}p_i'(d) - x(d) & (36)
\end{aligned}
$$

Note that $|p_i'| = w_i$ and $p_i'(d) = |p_i'|^{-1}\sum_{v\in p_i'} v(d)$. Also, $|P| = km$. Hence we can simplify to

$$\Delta x(d) \;=\; \frac{\sum_{v\in P'} v(d)}{km} - x(d) \tag{37}$$

Let us consider the term $\frac{\sum_{v\in P'} v(d)}{km}$, which is the centre of gravity of all objects in P'. Each $v \in P'$ is either a survivor from P, or it is a new exemplar, i.e. the average of some sample of k objects from the previous round. Of the items in P, kg-many do not survive in P'. Each non-surviving exemplar from P is chosen at random, so the expected average of the non-survivors is identical with c, with $x(d)$ as its d-th coordinate. Likewise, each new exemplar is the average of a random sample of exemplars. Hence the expected d-th coordinate of each new exemplar is $x(d)$ as well. There are g-many new exemplars that are added between t and t'. As g converges to infinity as m grows to infinity, the average of the new exemplars converges to their expected value, $x(d)$. Hence the average of the

items in P', $\frac{\sum_{v \in P'} v(d)}{km}$, converges to $x(d)$ as well. So we have

$$\Delta x(d) \rightarrow x(d) - x(d) \tag{38}$$
$$= 0. \tag{39}$$

In words, the centre of gravity of the population of exemplars remains constant. At least this is the limit behaviour of this dynamics for large population—for smaller populations the centre of gravity undergoes some random drift due to sampling effects.

As a next step, we analyse the evolution of the variance of the population. (Recall that the variance of a random variable X is defined as the expected value of $(X - E(X))^2$.) To this end we can posit another instance of the Price equation. First some notation:

$$Var(P(d)) = \sum_i \frac{w_i}{w} (p_i(d) - x(d))^2 \tag{40}$$

$$Var(P'(d)) = \sum_i \frac{w_i}{w} (p'_i(d) - x(d))^2 \tag{41}$$

$$\Delta Var(d) = Var(P'(d)) - Var(P(d)) \tag{42}$$

Plugging the variance into the Price equation, we get:

$$\Delta Var(d) = E((p'_i(d) - x(d))^2 - (p_i(d) - x(d))^2) \tag{43}$$

What is $p'_i(d)$? It obviously holds that

$$p'_i(d) = \frac{1}{|p'_i|} \sum_{v \in p'_i} v(d) \tag{44}$$

Let D_i be the set of items from p_i that do not survive in p'_i. Also, let N_i be the set of items in p'_i that are copies of new exemplars that have been added between t and t'. Let $N_i(d) = \sum_{v \in N_i} v(d)$, and likewise for $D_i(d)$. We thus have:

$$|p'_i| \rightarrow w_i \tag{45}$$
$$|D_i| \rightarrow \frac{w_i}{w} gk \tag{46}$$
$$|N_i| \rightarrow \frac{w_i}{w} gk \tag{47}$$
$$p'_i = (p_i - D_i) \cup N_i \tag{48}$$
$$p_i \cap N_i = \emptyset \tag{49}$$

We therefore get

$$p'_i(d) = \frac{1}{|p'_i|} (\sum_{v \in p_i} v - \sum_{v \in D_i} v + \sum_{v \in N_i} v) \tag{50}$$

$$= p_i(d) - \frac{1}{|p'_i|} (\sum_{v \in D_i} v + \sum_{v \in N_i} v) \tag{51}$$

22

$$= p_i(d) - \frac{|D_i|}{|p_i'|}D_i(d) + \frac{|N_i|}{|p_i'|}N_i(d) \tag{52}$$

$$\rightarrow p_i(d) - \frac{gk}{w}D_i(d) + \frac{gk}{w}N_i(d) \tag{53}$$

The elements of D_i are drawn at random from p_i, so

$$D_i(d) \rightarrow p_i(d) \tag{54}$$

Each element of N_i is the average of k random samples from P. So the expected average of N_i is identical to the average of P, and we thus have

$$N_i(d) \rightarrow x(d) \tag{55}$$

Putting all this together, we get

$$p_i'(d) \rightarrow p_i(d) - \frac{gk}{w}p_i(d) + \frac{gk}{w}x(d) \tag{56}$$

$$= (1 - \frac{gk}{w})p_i(d) + \frac{gk}{w}x(d). \tag{57}$$

Therefore

$$p_i'(d) - x(d) \rightarrow (1 - \frac{gk}{w})(p_i(d) - x(d)) \tag{58}$$

$$(p_i'(d) - x(d))^2 \rightarrow (1 - \frac{gk}{w})^2(p_i(d) - x(d))^2 \tag{59}$$

$$(p_i'(d) - x(d))^2 - (p_i(d) - x(d))^2 \rightarrow ((1 - \frac{gk}{w})^2 - 1)(p_i(d) - x(d))^2 \tag{60}$$

$$\tag{61}$$

As abbreviation, we use

$$\alpha = ((1 - \frac{gk}{w})^2 - 1) \tag{62}$$

Obviously $\alpha < 0$. Equation (43) now reduces to

$$\Delta Var(d) \rightarrow \alpha Var(P(d)) \tag{63}$$

The continous time version of the Price equation takes an even simpler form. If we assume that one unit of time corresponds to m update steps, the slope of $Var(d)$ is the limit $\frac{\alpha}{g}$, which converges to 2 as m grows to infinity. Therefore we get

$$\dot{Var}(d) = -2Var(d) \tag{64}$$

As $\alpha < 0$, both versions of the Price equation predict that the variance of the exemplars decreases at an exponential rate and asymptotically approaches 0. Since the centre of

gravity remains constant over time, this means that in the long run, all exemplars are concentrated in an arbitrarily small environment around c. The continuous time version of the Price equation can even be solved analytically; all functions of the form

$$Var(d) \;=\; K\exp(-2t) \tag{65}$$

for some constant K are possible solutions, and each solution converges to 0 as t goes to $+\infty$.

4 Conclusion

The main purpose of this article is to bring the conceptual framework that underlies the Price equation to the attention of linguists that are interested in evolutionary modelling. Price's framework has several attractive features that are briefly recapitulated here:

- Price's framework is very general. It does not over-emphasize certain features of biological evolution that are specific to biology rather than to the notion of evolution via selection—like assuming discrete units of heritable information or the dual ontology of genotype and phenotype. Rather, it focuses on population dynamics as such.

- There are no specific requirements about what the nature of the populations involved or the correspondence relation betwen them is. It is thus clear that evolution and selection are a perspective under which empirical phenomena can be studied, rather than being objective properties of these phenomena. Identifying a certain set as an evolving population and a certain relation between stages of this set as replication (i.e. correspondence) is a matter of practicality and usefulness, not of truth or falsity.

- In particular, Price's framework does not require anything like copying fidelity of replicators to be applicable. If a certain process does in fact involve faithfully replicating entities, this simplifies the analysis because the second term of the Price equation can be dropped in this case. However, this is a matter of convenience, not of principle.

- While Price's framework admits considerable methodological freedom, it enforces an absolutely rigorous analysis, once the basic modelling decisions are made.

In the second part of the paper, I presented two case studies of applications of the Price framework to language evolution. The examples involved two quite different notions of linguistics replication: first language acquisition in the first case and exemplar imitation in the second case. Both processes involve non-faithful replication; imperfect learning in the first example and blending inheritance in the second one. The main point of the discussion of the examples was to illustrate possible applications of Price's framework in linguistics. In both cases, the Price style model confirmed previous findings: Nowak's different mathematical approach in the first case and Pierrehumbert's computational simulation results in the second case.

References

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In T. Briscoe, ed., *Linguistic evolution through language acquisition: Formal and computational models*, pp. 111–172. Cambridge University Press, Cambridge, UK.

Croft, W. (2000). *Explaining Language Change*. Longman, New York.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray, London.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, Oxford.

Dennett, D. C. (1995). *Darwin's Dangerous Idea*. Simon & Schuster, New York.

Frank, S. A. (1995). George Price's contributions to evolutionary genetics. *Journal of Theoretical Biology*, **175**(3):373–388.

Gahl, S. and A. C. L. Yu, eds. (2006). *Exemplar-Based Models in Linguistics*, volume 23 of *The Linguistic Review*. Walter de Gruyter, Berlin, New York. Special theme issue.

Hauser, M. D., N. Chomsky, and W. T. Fitch (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, **298**(5598):1569.

Kirby, S. (1999). *Function, Selection, and Innateness*. Oxford University Press, Oxford.

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

Nowak, M. A. (2006). *Evolutionary Dynamics. Exploring the Equations of Life*. Harvard University Press, Cambridge, Mass. and London.

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**:611–617.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. *Typological studies in language*, **45**:137–158.

Pinker, S. and P. Bloom (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, **13**(4):707–784.

Price, G. R. (1970). Selection and covariance. *Nature*, **227**(5257):520–521.

Price, G. R. (1972a). Extension of covariance selection mathematics. *Annals of Human Genetics*, **35**(4):485–490.

Price, G. R. (1972b). Fisher's 'fundamental theorem' made clear. *Annals of Human Genetics*, **36**(2):129–40.

Price, G. R. (1995). The nature of selection. *Journal of Theoretical Biology*, **175**(3):373–388.

Ritt, N. (2004). *Selfish Sounds and Linguistic Evolution: A Darwinian Approach to Language Change*. Cambridge University Press.

Schwartz, J. (2000). Death of an altruist. *Lingua Franca: The Review of Academic Life*, **10**:51–61.

Smith, J. M. and G. R. Price (1973). The logic of animal conflict. *Nature*, **246**(5427):15–18.

Steels, L. (1996). Emergent adaptive lexicions. In P. Maes, M. J. Mataric, J. A. Meyer,

J. Pollack, and S. W. Wilson, eds., *Fourth International Conference on Simulation of Adaptive Behavior; Cape Cod, Massachusetts*, pp. 562–567. MIT.

Wedel, A. (2004). *Self-Organization and Categorical Behavior in Phonology.* Ph.D. thesis, University of California at Santa Cruz.