Modeling language change for the worse

Some considerations from an evolutionary perspective

Gerhard Jäger

Tübingen University

Workshop *Language change for the worse*

Munich, May 28, 2017

In what sense can language A be "worse" than language B?

- A is less regular.
- A is more complex.
- A is harder to acquire.
- A is harder to use, e.g.,
  - A requires more articulatory effort from the speaker to get a certain message across.
  - A requires more cognitive effort from the speaker to plan a certain utterance.
  - A requires more cognitive effort (lexical access, parsing, ...) from the listener to understand an utterance.
- Certain concepts or distinctions cannot be expressed in A.
- ...

Can we make this more precise?

An analogy from biology

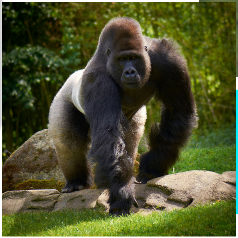
Evolution is **survival of the fittest.**

(Herbert Spencer; endorsed by Darwin)

Attributes of fitness

- size
- speed
- strength
- brain power
- ...






An analogy from biology

Definition (Fitness)

The *fitness* f_i of a trait i at time t:

$$f_i(t) \quad = \quad E\left({}^{w_i(t+1)}/_{w_i(t)}\right),$$

where $w_i(t)$ is the abundance of i-individuals at time t.

- Vulgo: Fitness of a trait is the *expected number of offspring* of individuals with that trait.

- Can we apply *fitness* to language?
- Can there be evolution that reduces fitness?

Language evolution

"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.)

Language evolution

standard assumptions about prerequisites for evolutionary processes (see for instance Richard Dawkins' work)

- population of **replicators** (for instance genes)
- (almost) faithful replication (for instance DNA copying)
- variation
- differential replication $\rightsquigarrow$ selection

Language evolution

What are the replicators?

- I-languages/grammars?
- E-languages/grammars?
- linguemes?
- rules?
- utterances (or features thereof)?

Perhaps Dawkins' conceptual framework is too narrow...

George R. Price

- 1922–1975
- studied chemistry; briefly involved in Manhattan project; lecturer at Harvard
- during the fifties: application of game theory to strategic planning of U.S. policy against communism
  - proposal to buy each Soviet citizen two pair of shoes in exchange for the liberation of Hungary
- tried to write a book about the proper strategy to fight the cold war, but *"the world kept changing faster than I could write about it"*, so he gave up the project
- 1961–1967: IBM consultant on graphic data processing

George R. Price

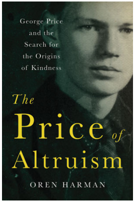
- 1967: emigration to London (with insurance money he received for medical mistreatment that left his shoulder paralyzed)
- 1967/1968: freelance biomathematician

George R. Price

- discovery of the **Price equation**
- leads to an immediate elegant proof of **Fisher's fundamental theorem**
- invention of **Evolutionary Game Theory**
  - Manuscript *Antlers, Intraspecific Combat, and Altruism* submitted to *Nature* in 1968; contained the idea of a mixed ESS in the Hawk-and-Dove game
  - accepted under the condition that it is shortened
  - reviewer: John Maynard Smith
  - Price never resubmitted the manuscript, and he asked Maynard Smith not to cite it
  - 1972: Maynard Smith and Price: *The Logic of Animal Conflict*
  - Price to Maynard Smith: *"I think this the happiest and best outcome of refereeing I've ever had: to become co-author with the referee of a much better paper than I could have written by myself."*

George R. Price

- 1968–1974: honorary appointment at the Galton Labs in London
- 1970: conversion to Christianity; after that, most of his attention was devoted to biblical scholarship and charity work
- around 1971: *The Nature of Selection* (published posthumously in 1995 in Journal of Theoretical Biology)
- early 1975: suicide



The Nature of Selection

"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."

The Nature of Selection

"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."

The Nature of Selection

Concepts of selection

- subset selection
- Darwinian selection

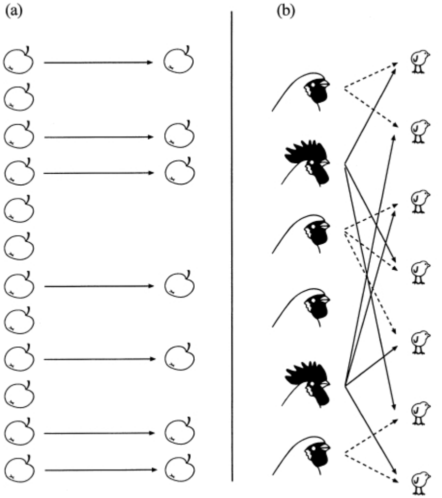


FIG. 1. Conventional concepts of selection. (a) Subset selection. (b) Darwinian selection.

The Nature of Selection

Concepts of selection

- common theme:
  - two time points
    - t: population before selection
    - t': population after selection

- partition of populations into N bins
- parameters
  - abundance w_i/w'_i of bin i before/after selection
  - quantitative character x_i/x'_i of each bin

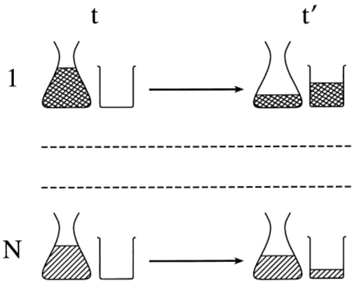


Fig. 2. A solution selection example.

The Nature of Selection

- each individual at t' **corresponds** to exactly one item at t
- nature of correspondence relation is up to the modeler — biological descendance is an obvious, but not the only possible choice
- partition of t-population induces partition of t'-population via correspondence relation

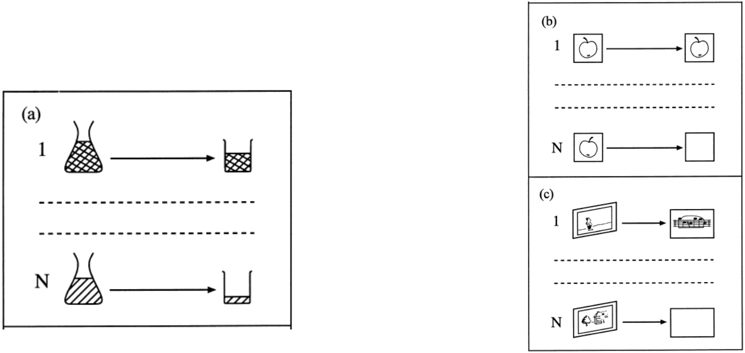
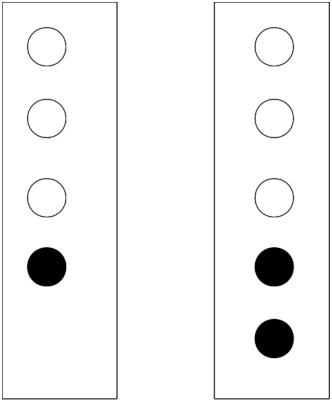


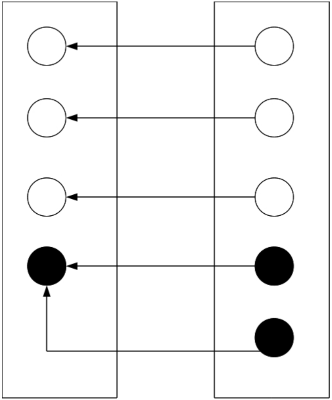
Fig. 3. Three selection examples arranged in the pattern of the general selection model. (a) The essential elements of the Fig. 2 example. (b) How the Fig. 1(a) example is fitted to the general model.

Schematic example



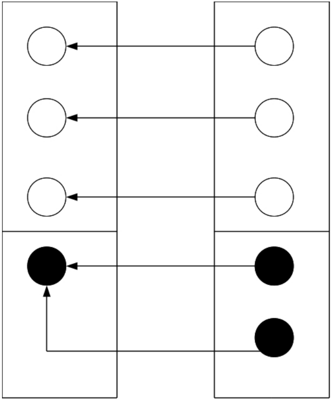
population at two points in time

Schematic example



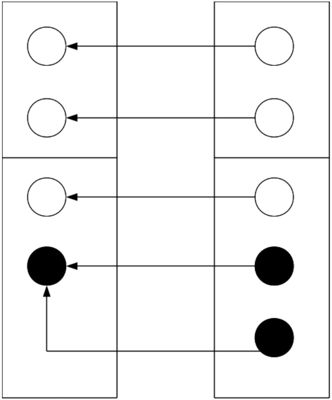
adding correspondence relation

Schematic example



adding partition structure

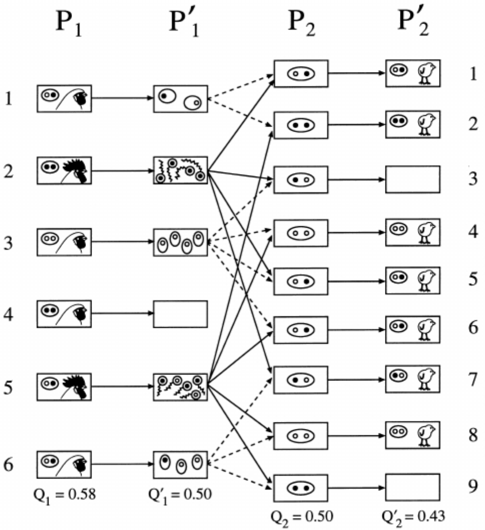
Schematic example



adding partition structure

The Nature of Selection

genetical selection:



The Price equation

Discrete time version

$$f\Delta x = Cov(f_i, x_i) + E(f_i \Delta x_i)$$

- $Cov(f_i, x_i)$: change of x due to natural selection
- $E(f_i \Delta x_i)$: change of x due to unfaithful replication

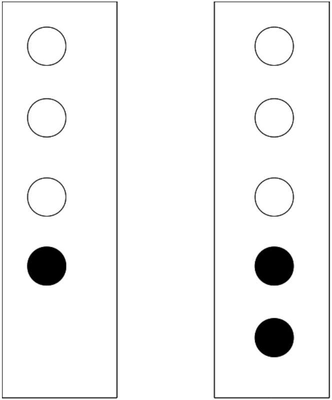
Continuous time version

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

The Price equation

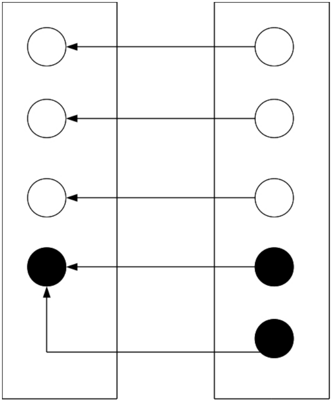
- important: the equation is a tautology
- follows directly from the definitions of the parameters involved
- very general; no specific assumptions about the nature of the replication relation, the partition of population into bins, the choice of the quantitative parameter under investigation
- many applications, for instance in investigation of group selection

Schematic example



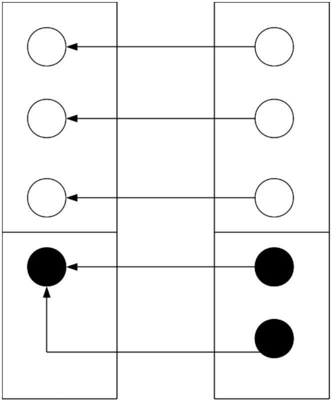
population at two points in time

Schematic example



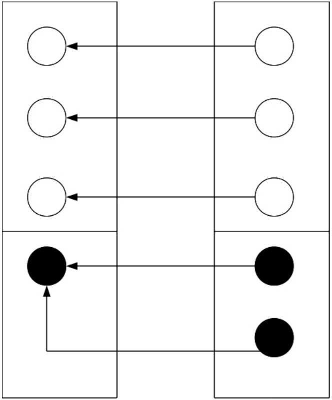
adding correspondence relation

Schematic example



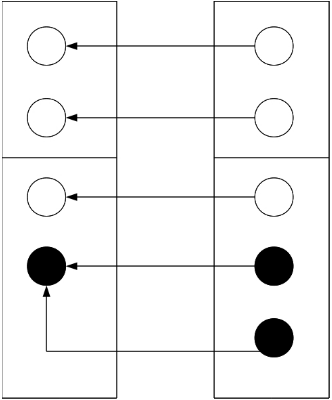
adding partition structure

Schematic example



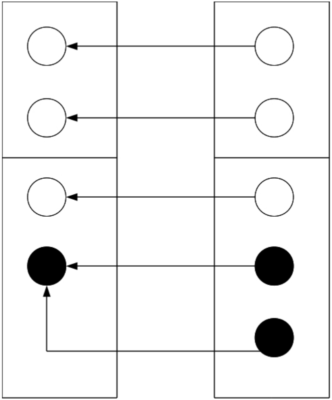
$$f\Delta x = Cov(f_i, x_i) + E(f_i \Delta x_i)$$
$$0.1875 = 0.1875 + 0$$

Schematic example



adding a different partition structure

Schematic example



$$
\begin{aligned}
f\Delta x &= Cov(f_i, x_i) + E(f_i \Delta x_i) \\
0.1875 &= 0.0625 + 0.125
\end{aligned}
$$

Applications of the Price equation

Fisher's Theorem

"The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time."

(R. A. Fisher, 1930)

- x can be any quantitative character, including fitness
- for $x = f$, we have

$$\dot{f} = Var_i(f_i) + E_i(\dot{f}_i)$$

  - $Var_i(f_i)$: increase in average fitness due to natural selection
  - $E_i(\dot{f}_i)$: decrease in average fitness due to
    - unfaithful replication (undirected or directed; cf. Lamarckian evolution)
    - deterioration of the environment

Applications of the Price equation

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

Group selection

- population of groups that each consists of individuals
- bins = groups
- first term:
  - covariance between a certain trait x and group fitness
  - corresponds to natural selection at the group level
- second term:
  - avarage change of x **within** group
  - corresponds to natural selection at the individual level
- for "altruistic" traits, first term would be positive but second term negative

Consequences of Price's approach

- no single "correct" way to model language evolution
- prerequisites for applying Price's approach:
  - two populations at different time points
  - natural assignment of items of the new population to items in the old population
- it is up to the model builder
  - what populations consist of (any measurable set would do)
  - the evolution of which character is studied (as long as it is quantitative in nature)
  - what the nature of the "replication" relation is — any function from the new population to the old one will do
  - how populations are partitioned into bins

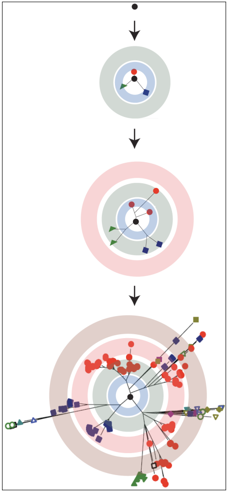
Language and fitness

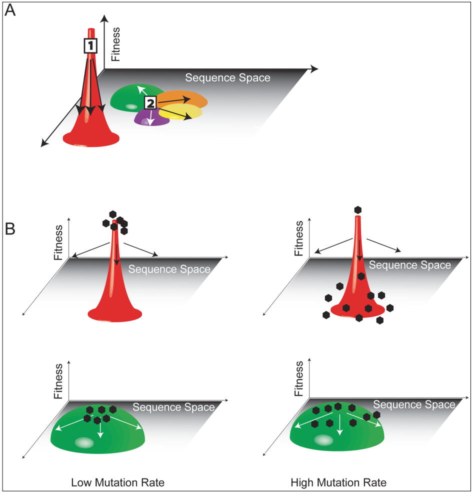
- Fitness of a lingueme $x \approx$ Propensity of x to be replicated.
- *Can language change for the worse?* translates to *Can a lingueme be replaced by less fit competitor?*
- General question: *Can evolution reduce fitness?*

Survival of the flattest

- $\dot{E}(f) < 0$ if $E(\dot{f}) < Var(f)$
- In words: fitness decreases if unfaithful replication/deterioration of the environment reduces fitness at a faster pace than the increase due to natural selection

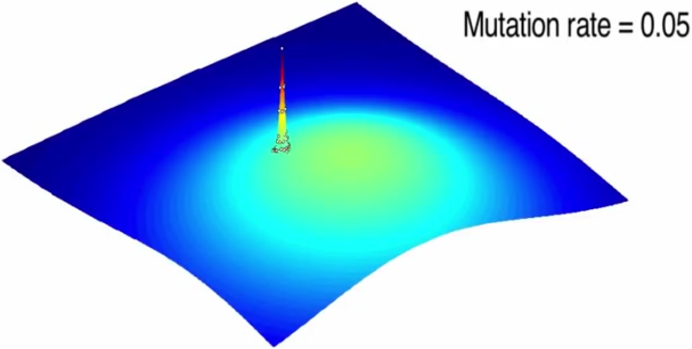
Survival of the flattest





(from Lauring and Andino 2010)

Survival of the flattest



At low mutation rate, the population stays on the highest peak

Population size, N = 2,304

Survival of the flattest

- applications to language:
  - little morphology $\rightarrow$ steep syntactic fitness landscape $\rightarrow$ low syntactic mutation rate $\approx$ fixed word order
  - Zipf's Law of Abbreviation
  - large number of L2-speakers increases mutation rate in acquisition
  - ...

Deterioration of the environment

Prisoner's Dilemma

	C	D
C	2	0
D	3	1

- Suppose a population consists of 50% cooperators and 50% defectors at time $t = 0$
- average payoffs ($=$ fitness):
  - $C : 1$
  - $D : 2$
  - population average: 1.5
  - $Var(f) = 0.25$

Deterioration of the environment

	$t = 0$	$t = 1$
proportion C	$1/2$	$1/3$
proportion D	$1/2$	$2/3$
fitness C	1	2
fitness D	$2/3$	$5/3$
expected fitness against itself	$3/2$	$4/3$
expected fitness agaings p_0	$3/2$	$5/3$

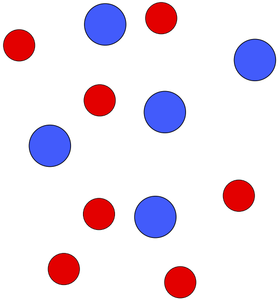
Price equation

$$\begin{aligned} f\Delta f &= Cov(f_i, f_i) &+& E(f_i \Delta f_i) \\ -0.25 &= 0.25 &+& -0.5 \end{aligned}$$

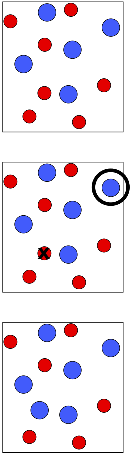
- the average fitness of the children, if placed into the parent generation, would exceed the parent's fitness
- however, the children interacting with the children do worse than the parents interacting with the parents

Weak selection in finite populations

- Suppose there are two variants (alleles, pronunciations, ...), A and B, with fitness f_A and f_B.
- $f_A > f_B$.
- If population is very large and can grow indefinitely, we expect both variants to coexist.
- Entirely different picture though if there is a finite upper limit on total population structure.



The Moran process



- Finite population of size N.
- In each time step, a random individual x is drawn for reproduction and another random individual y is drawn for replacement.
- y is replaced by a copy of x.

Neutral drift

- Let there be i type-A individuals (and $N - i$ type-B individuals).
- If A and B are picked with same probability for reproduction and replacement $\Rightarrow$ **neutral drift**
- Within finite time (monotonic in size of N), population will become monomorphic.
- probability of ending up in an only-A population: i/N

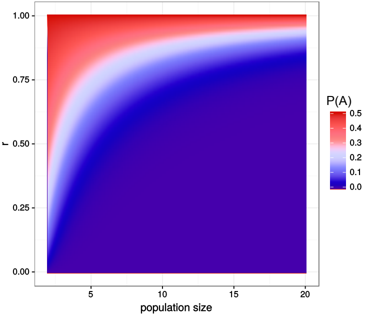
Weak selection

- Let p_A be the probability of an A-individual to be picked for replication (and same for p_B).
- Death probabilities are constant.
- Let $r \doteq p_A/p_B$
- The probability that a single A-mutant can flip an entire B-population to A is

$$P(B \to A) = \frac{r^{-1} - 1}{r^{-N} - 1}$$

- equilibrium probability

$$P(A) = \frac{P(B \to A)}{P(B \to A) + P(A \to B)}$$



In small populations, the probability that sub-optimal variants prevail is non-negligible.

(Nowak, 2006)

Summary

- Price equation is a versatile tool to model evolutionary processes beyond biology,
- including language change.

If we equate *worse* with *less fit*, there are three general scenarios how a Pricean system can change to the worse:

1. unfaithful replication, especially high rate of deleterious mutations

   fixation of word order, Zipf's law of abbreviation, language contact

2. deterioration of the environment

   loss of morphological marking via phonetic reduction, e.g. coherence of interrogative paradigm

3. stochastic effects in small populations

   higher morphological complexity in small populations (?), anti-DSM

Adam S. Lauring and Raul Andino. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens*, 6(7):e1001005, 2010.
Martin A. Nowak. *Evolutionary Dynamics. Exploring the Equations of Life*. Harvard University Press, Cambridge, Mass. and London, 2006.