Language, Games and Evolution

Evolutionary Game Theory

Gerhard Jäger

jaeger@ling.uni-potsdam.de

3rd NASSLLI, UCLA

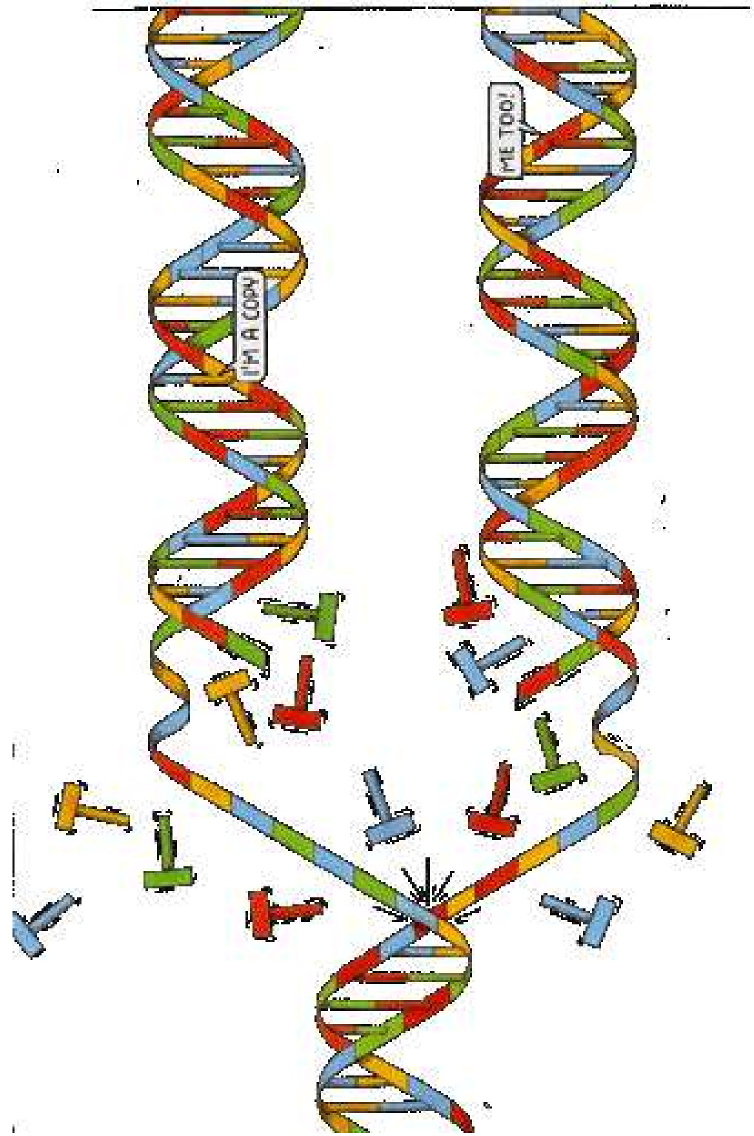
June 2004

Problems for classical GT

- multiple equilibria $\Rightarrow$ no predictions possible
- “perfectly rational player” is too strong an idealization

Evolutionary Game Theory

- populations of players
- individuals are (genetically) programmed for certain strategy
- individuals replicate and thereby pass on their strategy

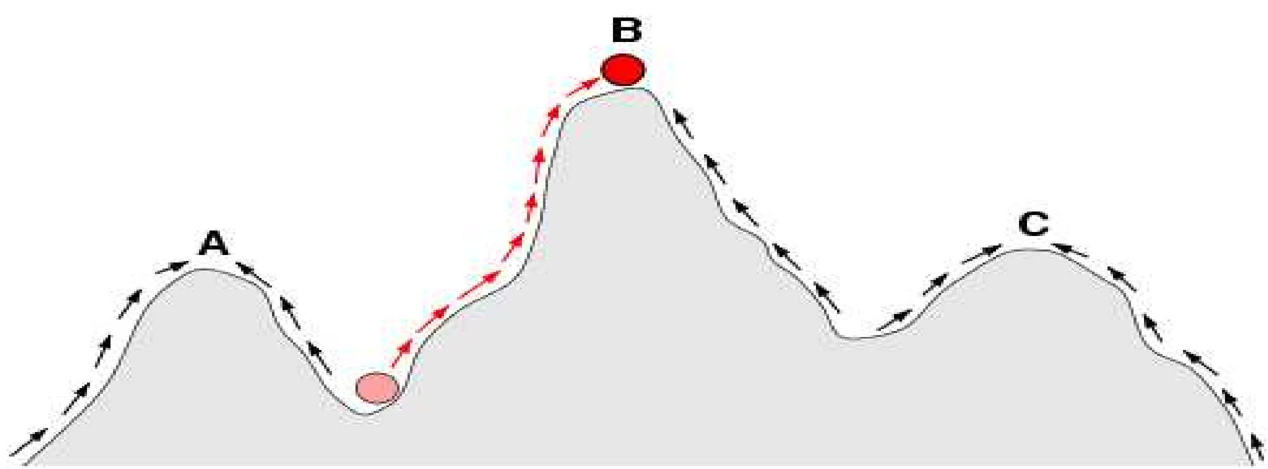


Utility and fitness

- number of offspring is monotonically related to average utility of a player
- high utility in a competition means the outcome improves reproductive chances (and vice versa)
- number of expected offspring (Darwin's “fitness”) corresponds to **expected utility** against a population of other players
- genes of individuals with high utility will spread

Evolutionary stability

- Darwinian evolution predicts ascent towards local fitness maximum
- once local maximum is reached: stability
- only random events (genetic drift, external forces) can destroy stability
- central question for evolutionary model: what are stable states?



Evolutionary stability (cont.)

- replication sometimes unfaithful (mutation)
- population is **evolutionarily stable** $\rightsquigarrow$ resistant against small amounts of mutation
- Maynard Smith (1982): static characterization of

  Evolutionarily Stable Strategies

  (ESS) in terms of utilities only
- related to Nash equilibria, but slightly different

Evolutionary stability (cont.)

Rock-Paper-Scissor

	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

- one Nash equilibrium: $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
- not evolutionarily stable though

Evolutionary stability (cont.)

Pigeon orientation game

- “players” are pigeons that go together on a journey
- A-pigeons can find their way back, B-pigeons cannot

	A	B
A	1	1
B	1	0

Evolutionary stability (cont.)

- A-is a non-strict Nash equilibrium, but nevertheless evolutionarily stable
- to be evolutionarily stable, a population must be able either
  - to fight off invaders directly (strict Nash equilibrium)
  - to successfully invade the invaders (non-strict Nash equilibrium)

Evolutionary Stable Strategy

- s is an Evolutionarily Stable Strategy iff
  - $u(s, s) \geq u(t, s)$ for all t, and
  - if $u(s, s) = u(t, s)$ for some $t \neq s$, then $u(s, t) > u(t, t)$.

Strict Nash Equilibria

$\subset$

Evolutionarily Stable Strategies

$\subset$

Nash Equilibria

The Replicator Dynamics

implicit assumption behind notion of ESS

- Populations are (practically) infinite.
- Each pair of individuals is equally likely to interact.
- The expected number of offspring of an individual (i.e., its fitness in the Darwinian sense) is monotonically related to its average utility.

can be made explicit in a dynamic model

Replicator Dynamics (cont.)

- easiest correlation between utility and fitness:

$$u(i, j) \quad = \quad \textit{expected number of offspring of an individual of type i in a j-population}$$

Replicator Dynamics (cont.)

suppose

- time is discrete
- in each round, each pair of players is equally likely to interact

Replicator Dynamics (cont.)

discrete time dynamics:

$$N_i(t+1) \quad = \quad N_i(t) + N_i(t)(\sum_{j=1}^{n} x_j u(i,j) - d)$$

$N(t)$ ... population size at time t
$N_i(t)$ ... number of players playing strategy s_i
$x_j(t)$ ... $\frac{N_j(t)}{N(t)}$
d ... death rate

Replicator Dynamics (cont.)

generalizing to continuous time:

$$N_i(t + \Delta t) = N_i + \Delta t(N_i \sum_{j=1}^{n} x_j u(i,j) - d)$$

thus

$$\frac{\Delta N_i}{\Delta t} = N_i(\sum_{j=1}^{n} x_j u(i,j) - d)$$

Replicator Dynamics (cont.)

if $\Delta t \rightarrow 0$

$$\frac{dN_i}{dt} = N_i(\sum_{j=1}^{n} x_j u(i,j) - d)$$

Replicator Dynamics (cont.)

size of entire population may also change:

$$
\begin{aligned}
N(t+\Delta t) &= \sum_{i=1}^{n}(N_i + \Delta t(N_i \sum_{j=1}^{n} x_j u(i,j) - d)) \\
&= N + \Delta t(N \sum_{i=1}^{n} x_i \sum_{j=1}^{n} x_j u(i,j))
\end{aligned}
$$

hence

$$
\frac{dN}{dt} = N(\sum_{i=1}^{t} x_i(\sum_{j=1}^{n} x_j u(i,j) - d))
$$

Replicator Dynamics (cont.)

let

$$\sum_{j=1}^{n} x_j u(i,j) = \tilde{u}_i$$

$$\sum_{i=1}^{n} x_i \tilde{u}_i = \tilde{u}$$

then we have

$$\frac{dN_i}{dt} = N_i(\tilde{u}_i - d)$$

$$\frac{dN}{dt} = N(\tilde{u} - d)$$

Replicator dynamics (cont.)

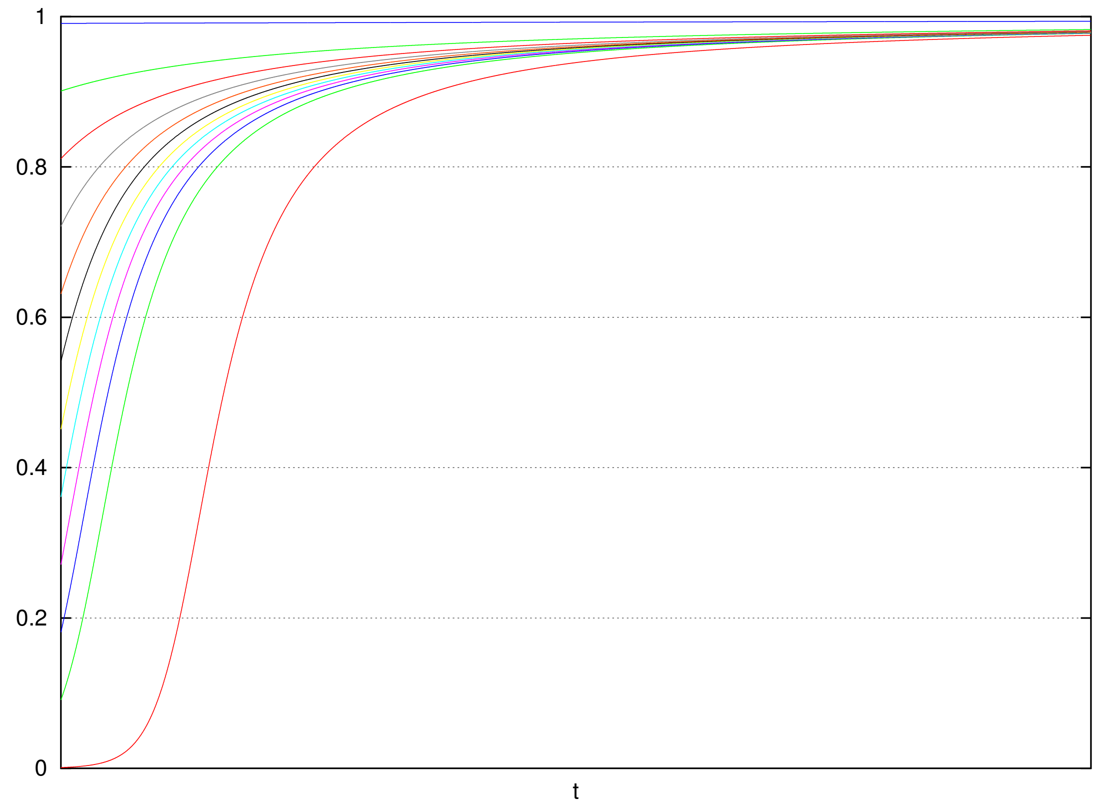
remember some calculus?

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

$$\begin{aligned}\frac{dx_i}{dt} &= \frac{(NN_i(\tilde{u}_i - d) - (N_iN(\tilde{u}_i - d)))}{N^2} \\ &= x_i(\tilde{u}_i - \tilde{u})\end{aligned}$$

Pigeon orientation

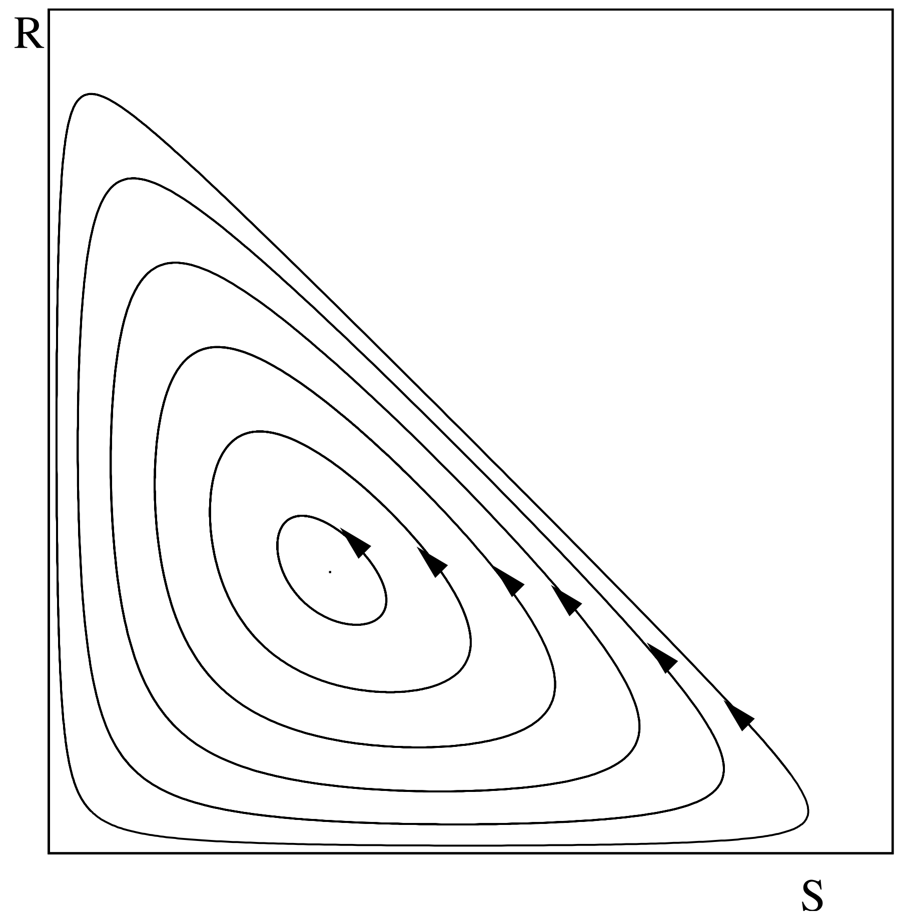
- ESSs correspond to **asymptotically stable states**
- a.k.a. **attractors**
- sample trajectories:



x-axis: time
y-axis: proportion of A-players

Rock-Paper-Scissor again

- three-strategy game: two independent variables
  - number of R-players
  - number of P-players
- number of S-players follows because everything sums up to 1
- supressing time dimension gives **orbits**



Asymmetric games

- symmetric games:
  - same strategy set for both players
  - $u_A(i,j) = u_B(j,i)$ for all strategies s_i, s_j
  - evolutionary interpretation: symmetric interaction *within one population*
- asymmetric games:
  - players have different strategy sets or utility matrices
  - evolutionary interpretation
    - different roles within one population (like incumbent vs. intruder, speaker vs. hearer, ...), or
    - interaction between disjoint populations
- evolutionary behavior differs significantly!

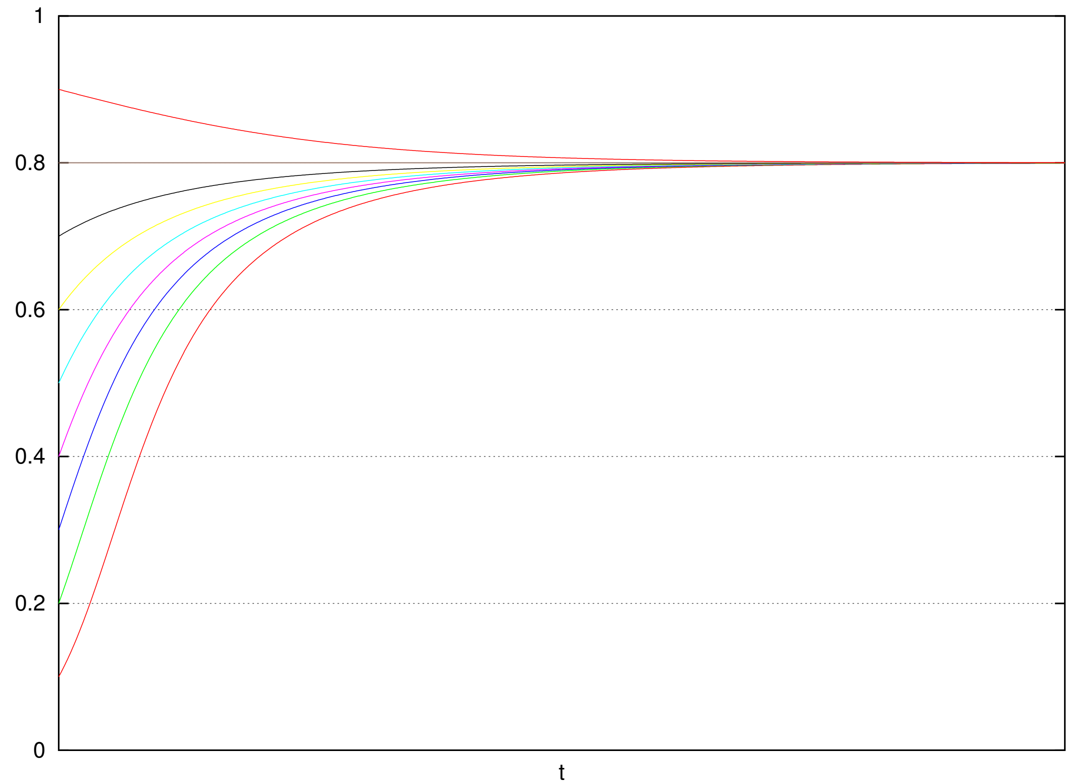
Asymmetric games (cont.)

Hawks and Doves

	H	D
H	1,1	7,2
D	2,7	3,3

- can be interpreted symmetrically or asymmetrically
- symmetric interpretation:
  - hawks prefer to interact with doves and vice versa
  - ESS: 80% hawks / 20% doves
  - both strategies have average utility of 2.2
  - trajectories:

Symmetric Hawk-and-doves



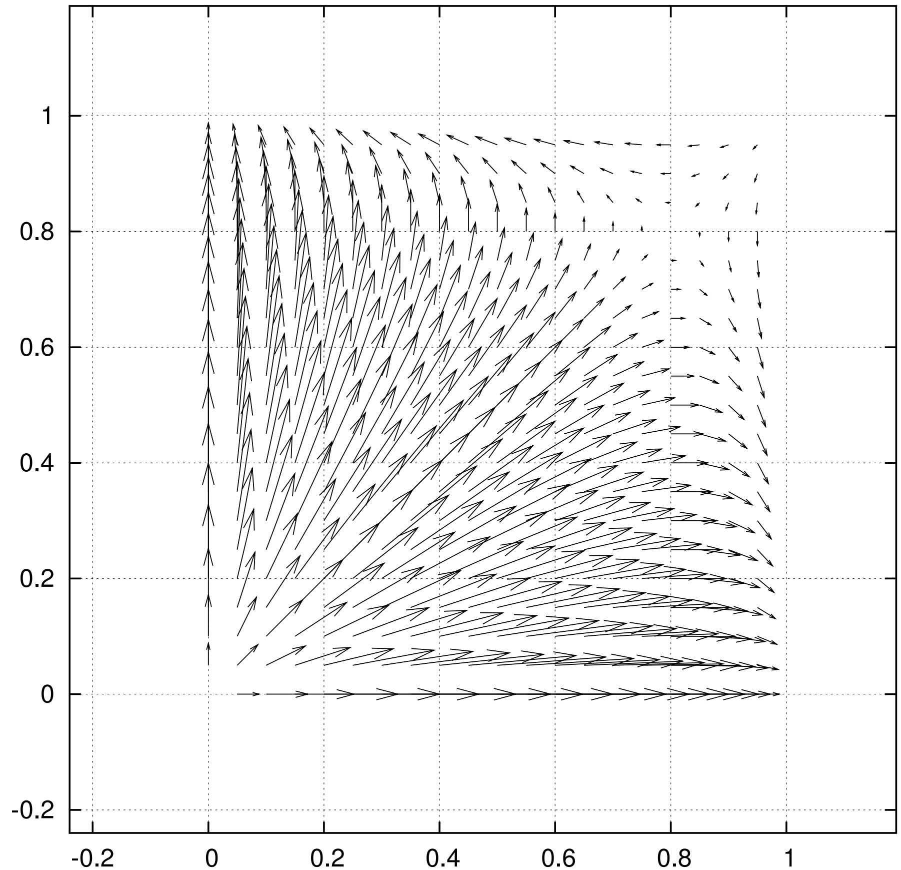
- if hawks exceed 80%, doves thrives, and vice versa
- 80:20 ratio is only attractor state

Asymmetric Hawks-and-doves

- suppose two-population setting:
  - both A and B come in hawkish and dovish variant
  - everybody only interacts with individuals from opposite “species”
  - excess of A-hawks helps B-doves and vice versa
  - population push each other into opposite directions

Hawks and doves

- 80:20 ratio in both populations is stationary
- not an attractor, but repellor



Asymmetric stability

- crucial difference to symmetric games:
  mutants do not play against themselves
- makes second clause of the symmetric ESS superfluous

In asymmetric games, a configuration is an ESS iff it is a strict Nash equilibrium.

Asymmetric replicator dynamic

$$\frac{dx_i}{dt} = x_i(\sum_{j=1}^{n} y_j u_A(i,j) - \sum_{k=1}^{n} x_k \sum_{j=1}^{n} y_j u_A(k,j))$$

$$\frac{dy_i}{dt} = y_i(\sum_{j=1}^{m} x_j u_B(i,j) - \sum_{k=1}^{n} y_k \sum_{j=1}^{m} x_j u_B(k,j))$$

x_i ... proportion of s_i^A within the A-population
y_i ... proportion of s_i^B within the B-population

Symmetrizing asymmetric games

- asymmetric games can be “symmetrized”
- correspondig symmetric game shares Nash equilibria and ESSs
- new strategy set:

$$S^{AB} = S^A \times S^B$$

- new utility function

$$u^{AB}(\langle i, j\rangle, \langle k, l\rangle) = u^A(i, l) + u^B(j, k)$$

Evolution in biology and linguistics

- correspondence between biology and linguistics

utterance	$\approx$	organism
language	$\approx$	species
dialect	$\approx$	deme
idiolect	$\approx$	lineage

Evolution in biology and linguistics

- concept of *evolution* can be applied to linguistic as well

genotype	$\approx$	grammatical knowledge (“langue”)
phenotype	$\approx$	utterances (“parole”)
replication	$\approx$	imitation

Mathematical models from evolutionary biology should be applicable to linguistics!

- Biological evolution is driven by variation and selection
- variation
  - Biology: mutations
  - Linguistics: errors, language contact, fashion...
- selection:
  - Biology: fitness = number of fertile offspring
  - Linguistics: communicative functionality, efficiency, social prestige, learnability, ...

EGT and pragmatics

Horn strategies: prototypical meanings tend to go with simple expressions and less prototypical meanings with complex expressions.

(1) a. John went to church/jail. (prototypical interpretation)
b. John went to the church/jail. (literal interpretation)

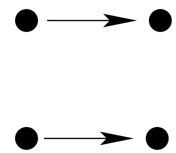
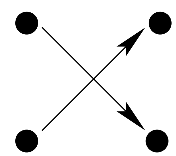
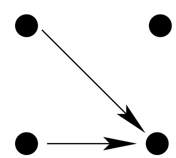
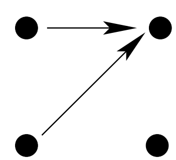
(2) a. I am going to marry you. (no indirect speech act)
b. I will marry you. (indirect speech act)

(3) a. I need a new driller/cooker.
b. I need a new drill/cook.

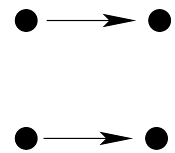
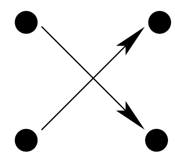
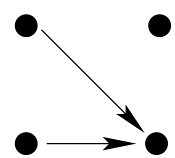
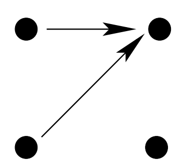
Horn strategies

- simple game:
  - players: speaker and hearer
  - two forms: f_0 (short) and f_1 (long)
  - two meanings: m_0 (frequent) and m_1 (rare)
  - speaker strategies: mappings from meanings to forms
  - hearer strategies: mappings from forms to meanings

Speaker strategies

- $S_1 : m_0 \mapsto f_0, m_1 \mapsto f_1$:  "Horn strategy"
- $S_2 : m_0 \mapsto f_1, m_1 \mapsto f_0$:  "anti-Horn strategy"
- $S_3 : m_0 \mapsto f_0, m_1 \mapsto f_0$:  "Smolensky strategy"
- $S_4 : m_0 \mapsto f_1, m_1 \mapsto f_1$:  "anti-Smolensky strategy"

Hearer strategies

- $H_1 : f_0 \mapsto m_0,\ f_1 \mapsto m_1$:  “Horn strategy’
- $H_2 : f_0 \mapsto m_1,\ f_1 \mapsto m_0$:  “anti-Horn strategy”
- $H_3 : f_0 \mapsto m_0,\ f_1 \mapsto m_0$:  “Smolensky strategy”
- $H_4 : f_0 \mapsto m_1,\ f_1 \mapsto m_1$:  “anti-Smolensky strategy”

Utility of Horn games

- whether communication works depends both on speaker strategy S and hearer strategy H
- two factors for functionality of communication
  - communicative success ("hearer economy")

$$\delta_m(S,H) = \begin{cases} 1 & \text{iff} \quad H(S(m)) = m \\ 0 & \text{else} \end{cases}$$

  - least effort ("speaker economy")

cost$(f)\ldots$ measure of complexity of expression

Utility of Horn games

$$u_s(S,H) = \sum_m p_m \times (\delta_m(S,H) - \textit{\textbf{cost}}(S(m)))$$
$$u_h(S,H) = \sum_m p_m \times \delta_m(S,H)$$

$p\ldots$ probability distribution over meaning types

Utility of Horn game

Let's make up some numbers:

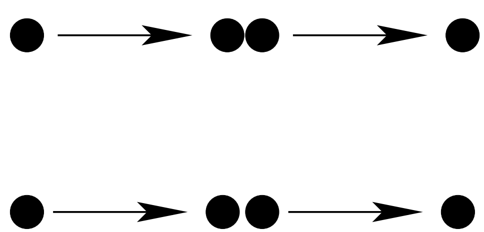
- $p(m_0) = .75$
- $p(m_1) = .25$
- *cost*$(f_0) = .1$
- *cost*$(f_1) = .2$

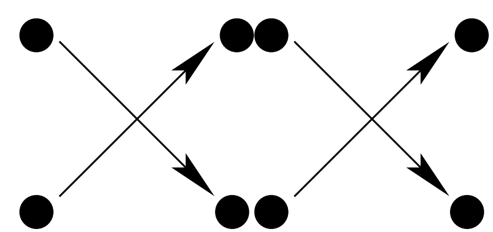
Utility of Horn game

	H_1		H_2		H_3		H_4	
S_1	.875	1.0	-.125	0.0	.625	.75	.125	.25
S_2	-.175	0.0	.825	1.0	.575	.75	.25	.075
S_3	.65	.75	.15	.25	.65	.75	.15	.25
S_4	.05	.25	.55	.75	.55	.75	.05	.25

Utility of Horn game

	H_1		H_2		H_3		H_4	
S_1	.875	1.0	-.125	0.0	.625	.75	.125	.25
S_2	-.175	0.0	.825	1.0	.575	.75	.25	.075
S_3	.65	.75	.15	.25	.65	.75	.15	.25
S_4	.05	.25	.55	.75	.55	.75	.05	.25





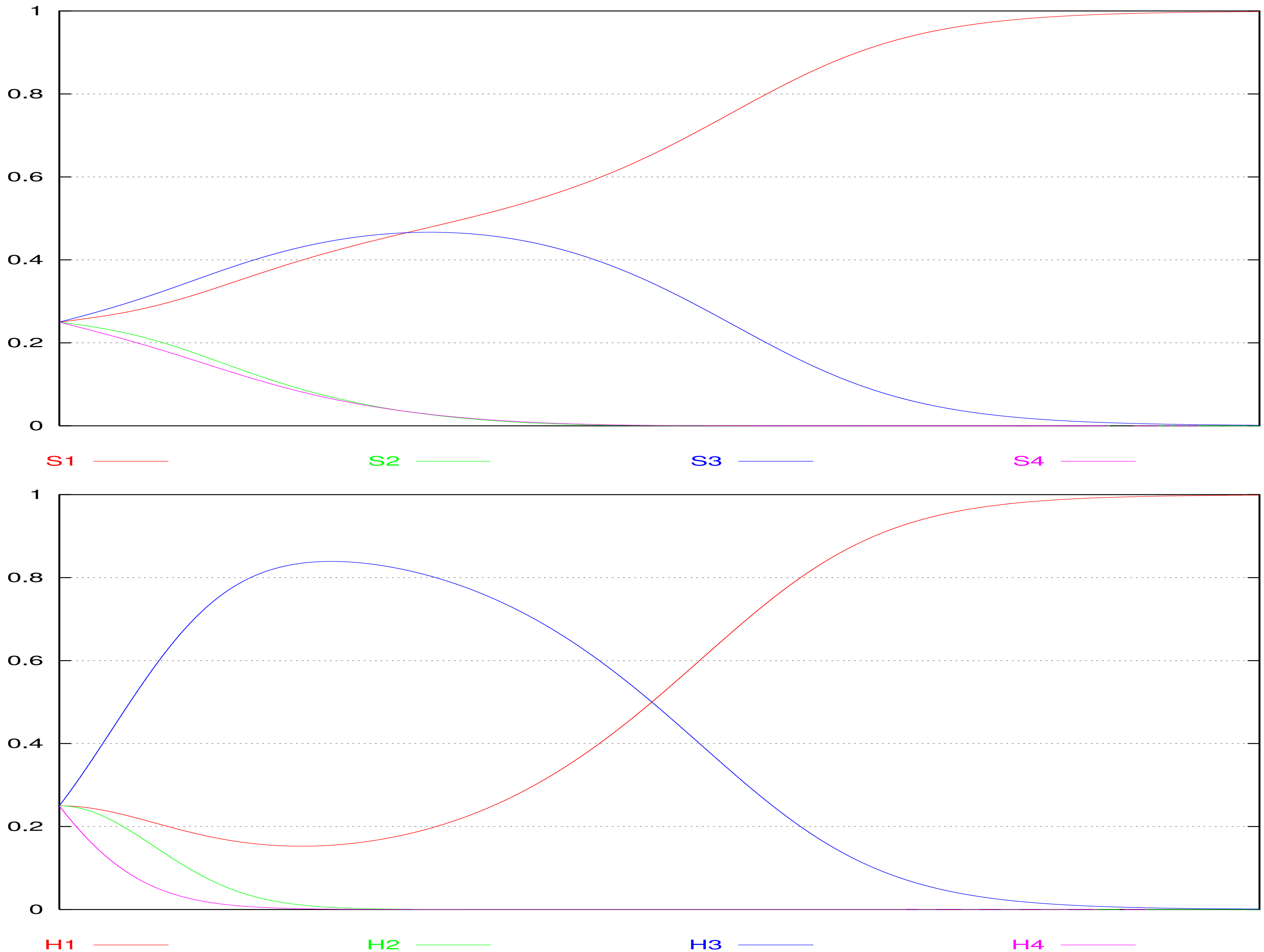
The problem of equilibrium selection

- both Horn and anti-Horn are evolutionarily stable
- EGT explains the aversion of natural against synonymy and ambiguity
- preference for Horn not directly explainable in standard EGT

The problem of equilibrium selection

- rationalistic considerations favor Horn over anti-Horn:
  - Horn strategy is **Pareto efficient** (nobody can do better in absolute terms)
  - Horn strategy **risk dominates** anti-Horn (if you know the population is in an equilibrium but you do not know in which one, going for Horn is less risky than anti-Horn)
- replicator dynamics favors Horn over anti-Horn:
  - complete random state evolves to Horn/Horn
  - basin of attraction of Horn is about 20 times as large as basin of attraction of anti-Horn (numerical approximation—does anybody know how to do this analytically?)

Trajectories starting from random state



The evolution of differential case marking

Ways of argument identification

- transitivity may lead to ambiguity

- three ways out
  1. word order
  2. agreement
  3. **case**

die Frau, die er kennt

the woman that he knows

die **Frau**, **die** ihn **kennt**

the woman that knows him

- Suppose one argument is a pronoun and one is a noun (or a phrase)

$$\{\text{I, BOOK, KNOW}\}$$

- both conversants have an interest in successful communication
- case marking (accusative or ergative) is usually more costly than zero-marking (nominative)
- speaker wants to avoid costs

speaker strategies	*hearer strategies*
always case mark the object (accusative)	ergative is agent and accusative object
always case mark the agent (ergative)	pronoun is agent
case mark the object if it is a pronoun	pronoun is object
	pronoun is agent unless it is accusative
⋮	⋮

Statistical patterns of language use

four possible clause types:

	O/p	*O/n*
A/p	he knows it	he knows the book
A/n	the man knows it	the man knows the book

statistical distribution (from a corpus of spoken English)

	O/p	*O/n*
A/p	pp = 198	pn = 716
A/n	np = 16	nn = 75

pn $\gg$ np

- functionality of speaker strategies and hearer strategies depends on various factors:
  - How often will the hearer get the message right?
  - How many case markers does the speaker need per clause — on average?

- speaker strategies that will be considered:

agent is pronoun	*agent is noun*	*object is pronoun*	*object is noun*
e(rgative)	e(rgative)	a(ccusative)	a(ccusative)
e	e	a	z(ero)
e	e	z	a
e	e	z	z
e	z	a	a
...	...	...	...
z	e	z	z
z	z	a	a
z	z	a	z
z	z	z	a
z	z	z	z

- hearer strategies:
  - strict rule: ergative means “agent”, and accusative means “object”
  - elsewhere rules:

1. SO: “The first phrase is always the agent.”
2. pA: “Pronouns are agents, and nouns are objects.”
3. pO: “Pronouns are objects, and nouns are agents.”
4. OS: “The first phrase is always the object.”

The game of case

- strategy space and utility function are known
- probability of meaning types can be estimated from corpus study
- hard to estimate how the complexity of a case morpheme compares to its benefit for disambiguation from the speaker perspective
- parameterized utility function

$$u(S,H) \;=\; \sum_m p_m \times (\delta_m(S,H) - k \times \textit{\textbf{cost}}(S(m)))$$

Utility of case marking

- let us assume $k = .1$

Speaker strategies	Hearer strategies			
	SO	pA	pO	OS
$eezz$	0.90	0.90	0.90	0.90
$zzaa$	0.90	0.90	0.90	0.90
$ezaz$	0.85	0.85	0.85	0.85
$zeza$	0.81	0.81	0.81	0.81
$zeaz$	0.61	0.97	0.26	0.61
$ezzz$	0.86	0.86	0.87	0.86
$zezz$	0.54	0.89	0.54	0.54
$zzaz$	0.59	0.94	0.59	0.59
$zzza$	0.81	0.81	0.82	0.81
$zzzz$	0.50	0.85	0.15	0.50

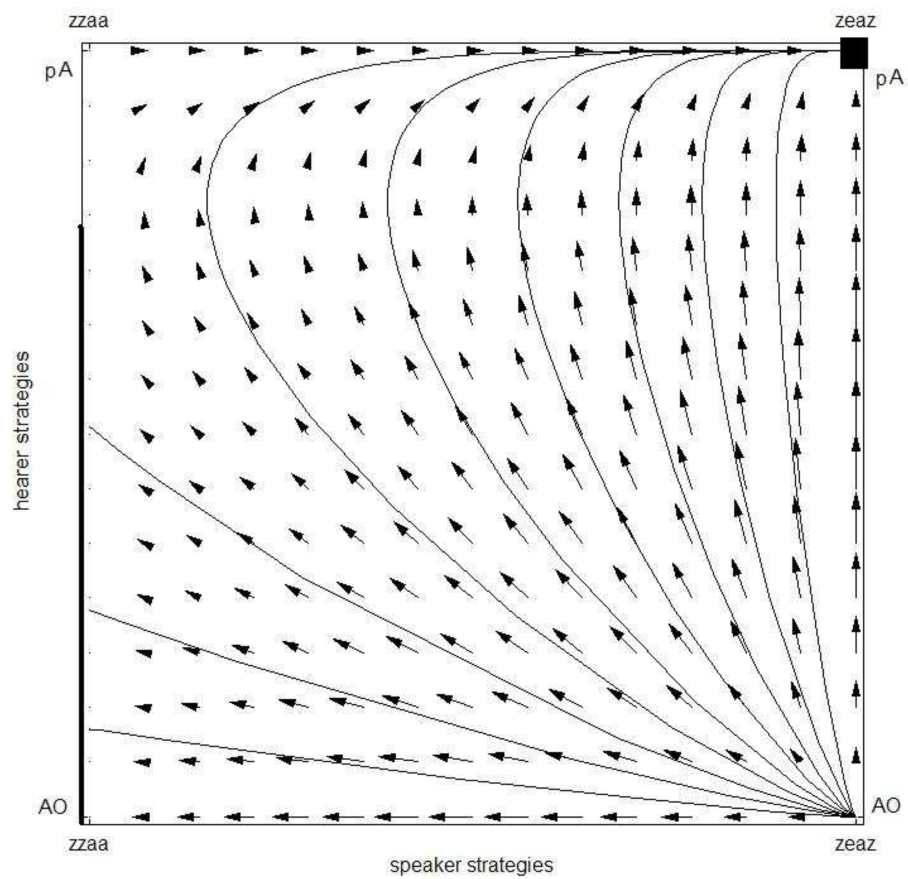
Utility of case marking

- only one evolutionary stable state: $zeaz/pA$ (*split ergative*)
- very common among Australian aborigines languages

Non-strict Nash equilibria

Why are non-strict Nash Equilibria unstable?

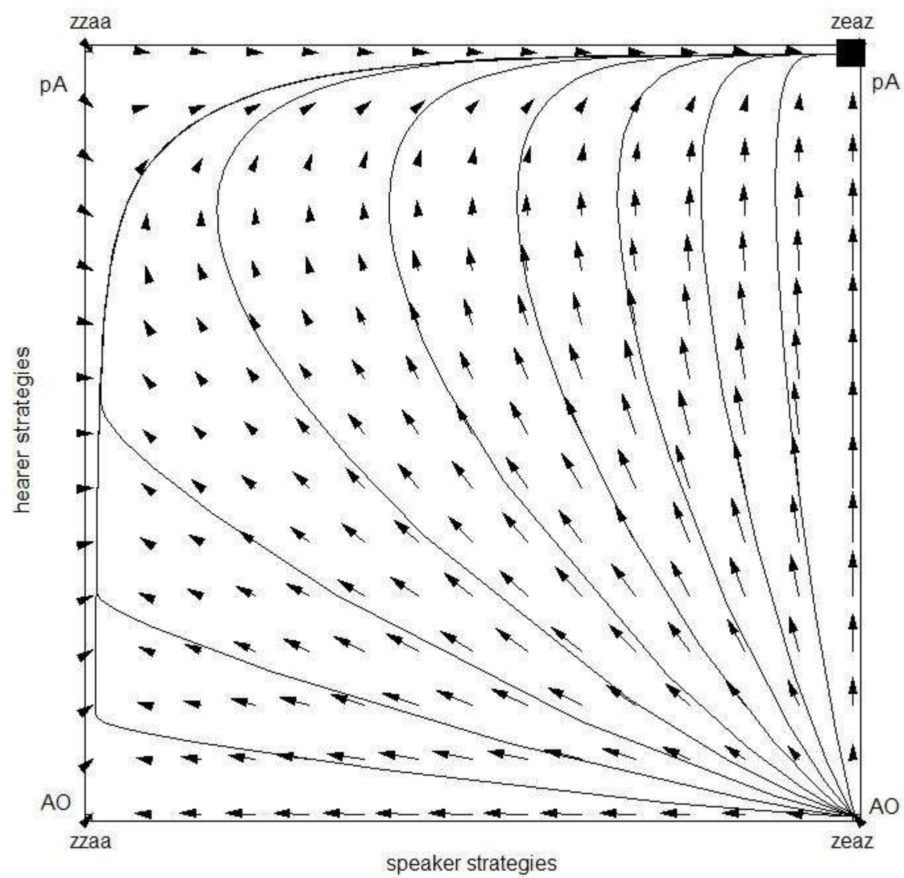
- Dynamics without mutation



Non-strict Nash equilibria

Why are non-strict Nash Equilibria unstable?

- Dynamics with mutation



Utility of case marking

If speakers get lazier...

- $k = 0.45$

Speaker strategies	Hearer strategies			
	SO	pA	pO	OS
$eezz$	0.550	0.550	0.550	0.550
$zzaa$	0.550	0.550	0.550	0.550
$ezaz$	0.458	0.458	0.458	0.458
$zeza$	0.507	0.507	0.507	0.507
$zeaz$	0.507	0.863	0.151	0.507
$ezzz$	0.545	0.538	0.553	0.545
$zezz$	0.505	0.861	0.148	0.505
$zzaz$	0.510	0.867	0.154	0.510
$zzza$	0.539	0.531	0.547	0.539
$zzzz$	0.500	0.849	0.152	0.500

Utility of case marking

... and lazier ...

- $k = 0.53$

Speaker strategies	Hearer strategies			
	SO	pA	pO	OS
$eezz$	0.470	0.470	0.470	0.470
$zzaa$	0.470	0.470	0.470	0.470
$ezaz$	0.368	0.368	0.368	0.368
$zeza$	0.436	0.436	0.436	0.436
$zeaz$	0.483	0.839	0.127	0.483
$ezzz$	0.473	0.465	0.480	0.473
$zezz$	0.497	0.854	0.141	0.497
$zzaz$	0.494	0.850	0.137	0.494
$zzza$	0.476	0.468	0.484	0.476
$zzzz$	0.500	0.848	0.152	0.500

Utility of case marking

... and lazier...

- k = 0.7

Speaker strategies	Hearer strategies			
	SO	pA	pO	OS
$eezz$	0.300	0.300	0.300	0.300
$zzaa$	0.300	0.300	0.300	0.300
$ezaz$	0.177	0.177	0.177	0.177
$zeza$	0.287	0.287	0.287	0.287
$zeaz$	0.431	0.788	0.075	0.431
$ezzz$	0.318	0.310	0.326	0.318
$zezz$	0.482	0.838	0.126	0.482
$zzaz$	0.457	0.814	0.101	0.457
$zzza$	0.343	0.335	0.350	0.343
$zzzz$	0.500	0.848	0.152	0.500

Utility of case marking

...

- k = 1

Speaker strategies	Hearer strategies			
	SO	pA	pO	OS
$eezz$	0.000	0.000	0.000	0.000
$zzaa$	0.000	0.000	0.000	0.000
$ezaz$	−0.160	−0.160	−0.160	−0.160
$zeza$	0.024	0.024	0.024	0.024
$zeaz$	0.340	0.697	−0.016	0.340
$ezzz$	0.045	0.037	0.053	0.045
$zezz$	0.455	0.811	0.099	0.455
$zzaz$	0.394	0.750	0.037	0.394
$zzza$	0.106	0.098	0.144	0.106
$zzzz$	0.500	0.848	0.152	0.500

Taking stock

$zeaz/pA$
split ergative
Australian languages

$zzaz/pA$	$ezzz/pO$
differential object marking	inverse DOM
English, Dutch, ...	—
$zezz/pA$	$zzza/pO$
differential subject marking	inverse DSM
several caucasian languages	Nganasan
$zzzz/pA$	$zzza/pO$
no case marking	
Chinese, Thai	

$zzzz/pA$

Taking stock

- only very few languages are not evolutionary stable in this sense
  $zzaa$: *Hungarian,* $ezza$: *Arrernte,* $eeaa$: *Wangkumara*
- curious asymmetry: if there are two competing stable states, one is common and the other one rare
- similar pattern as with Horn vs. anti-Horn

Alle equilibria are stable, but some equilibria are more stable than others.

Stochastic EGT

Random mutation and stability

- idealizations of standard Evolutionary Game Theory
  - populations are (practically) infinite
  - mutations rate is constant and low
- better model (Young 1993; Kandori, Mailath and Rob 1993)
  - finite population
  - mutation is noisy

Consequences of finite population model

- every mutation barrier will occasionally be taken
- no absolute stability
- if multiple Strict Nash Equilibria coexist, system will oscillate between them
- some equilibria are more stable than others
- system will spend most of the time in most robustly stable state
- stochastically stable states

A particular model

- discrete time/finite population version of replicator dynamics
- mutations occur rarely (most generations have no mutants at all)
- if mutation occurs, each individual in this generation has same probability to be a mutant
- mutation frequency and mutation rate equal for both populations
- each strategy is equally likely for a mutant (within its population)

A simulation

Stochastic stability

- punctuated equilibria
- long periods of dynamic stability alternate with short transition periods
- in the long run, more time in Horn state (67% vs. 26% in anti-Horn)
- simulation suggests that Horn is stable while anti-Horn is not
- can this be proved?

Analytic considerations

- Simple recipes for finding stochastically stable state in 2×2 games
- not easily extrapolated to larger games
- basic idea:
  - calculate the height of the invasion barrier of each ESS
  - the ESSs with maximal invasion barrier is stochastically stable

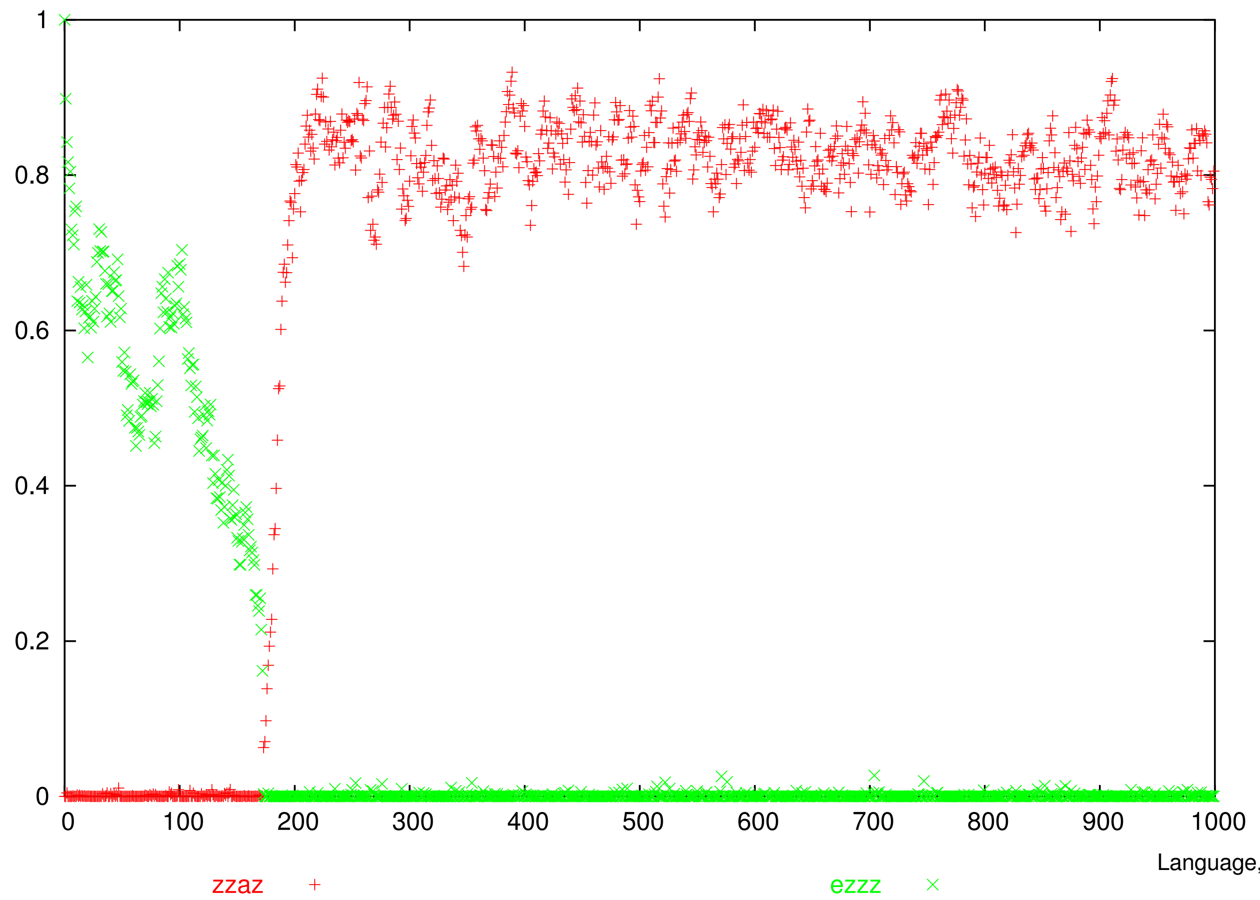
Analytic considerations

- invasion barrier = amount of mutations necessary to push the system into the basin of attraction of another ESS
- Horn $\Rightarrow$ anti-Horn: 50%
- anti-Horn $\Rightarrow$ Horn: 47.5%
- Hence:

Horn strategy is the only stochastically stable state

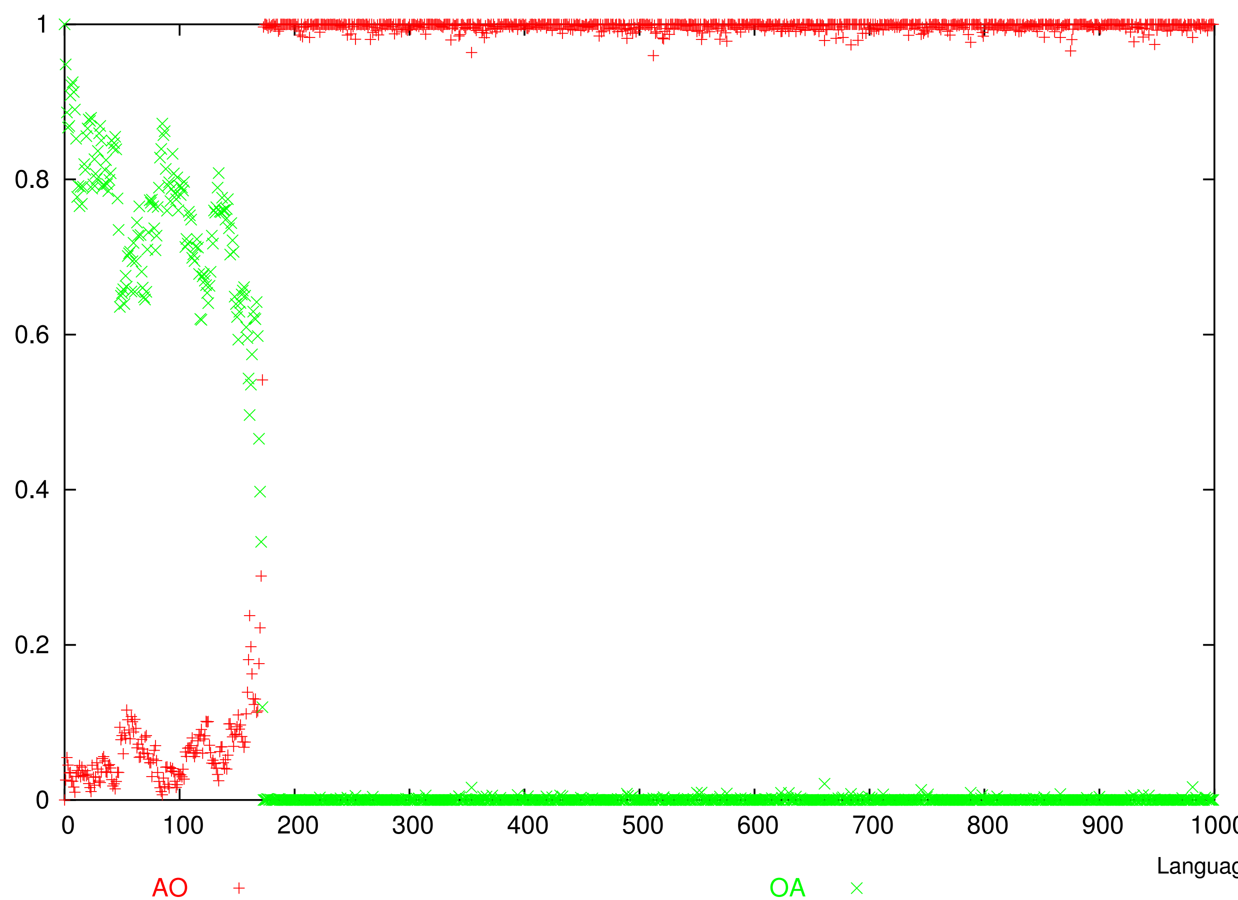
Stochastic evolution of case marking

- $k = 0.45$
- competition between $zzaz/pA$ and $ezzz/pO$
- evolution of speaker population:



Stochastic evolution of case marking

- $k = 0.45$
- competition between $zzaz/pA$ and $ezzz/pO$
- evolution of hearer population:



Analysis

- invasion barriers:
  - differential object marking: 45.2%
  - inverse differential subject markig: 2.06%

Differential object marking is stochastically stable; inverse differential subject marking is not.

- likewise, differential subject marking is stochastically stable while inverse differential object marking is not.

Stochastically stable states

$zeaz/pA$
split ergative
Australian languages

$zzaz/pA$
differential object marking
English, Dutch, ...

$zezz/pA$
differential subject marking
several caucasian languages

$zzzz/pA$
no case marking
Chinese, Thai

Conclusion

- out of $4 \times 16 = 64$ possible case marking patterns only four are stochastically stable
- vast majority of all languages that fit into this categorization are stochastically stable
- precise numbers are hard to come by though
- linguistic universals can be result of evolutionary pressure in the sense of cultural evolution

Iterated learning vs. iterated usage

- language is self-replicating system
- two modes of replication:
  1. (first) language acquisition
  2. language usage
- the modes differ in
  - selection pressure
  - source of variation
  - time scale

How do they interact?

Acqusition dynamics

- *replicator:* I-language in its entirety
- *interactors:* “teacher” (adult) and “student” (infant)
- *source of variation:* imperfect learning
- *time scale:* measured in decades

Usage dynamics

- *replicator:* components of I-language
  (lexical entries, constructions, ...)
- *interactors:* (mainly adult) language users
- *source of variation:* errors, language contact, ...
- *time scale:* detectable even within single text

The Iterated Learning Model

- formal model of acqusition dynamics
- many computational implementations (Hurford, Kirby, Briscoe, Niyogi, Berwick, ...)
- analytical mathematical formulation by Nowak (with various co-authors)

The Iterated Learning Model (cont.)

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

$$f_j(\mathbf{x}) \doteq \sum_k x_k U_{jk}$$

- main components:
  - fitness function f
  - learning matrix Q

Fitness

- *Biology:* fitness $\doteq$ expected number of fertile offspring
- *Linguistics:* communicative functionality, efficiency, social prestige, ...

Fitness (cont.)

- first approximation
  - finite number of languages $L_1, \cdots, L_n$
  - σ_{ij} ... average probability that a speaker using L_i is understood by a listener using L_j
  - c_i ... average complexity of utterances of L_i (length, entropy, whatever)
  - utility of communication between users of L_i and L_j:

$$U_{ij} = \frac{1}{2}(\sigma_{ij} + \sigma_{ji} - r(c_i + c_j))$$

Fitness (cont.)

- x_i ... relative frequency of users of L_i in proportion to total population

$$\sum_i x_i = 1$$

- $\mathbf{x}$ ... vector of relative frequencies $x_1, x_2, \cdots, x_n$
- fitness = average utility:

$$f_j(\mathbf{x}) \doteq \sum_k x_k U_{jk}$$

The learning matrix

- not every language is perfectly learnable
- Q_{ij} ... probability that an infant growing up in an L_i-environment acquires L_j

$$\sum_j Q_{ij} = 1$$

The learning matrix (cont.)

- simplest case:
  - identity matrix
  - infant always acquires language of environment

	L_1	L_2	L_3	$\cdots$
L_1	1	0	0	$\cdots$
L_2	0	1	0	$\cdots$
L_3	0	0	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment
- fitness (= abundance of offspring of users) of L_j

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment
- fitness (= abundance of offspring of users) of L_j
- abundance of infants that acquire L_i

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment
- fitness (= abundance of offspring of users) of L_j
- abundance of infants that acquire L_i
- death rate

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment
- fitness (= abundance of offspring of users) of L_j
- abundance of infants that acquire L_i
- death rate
- velocity of change of abundance of L_i-speakers

Acquistion dynamics

$$\frac{dx_i}{dt} = \sum_j x_j f_j(\mathbf{x}) Q_{ji} - x_i \sum_j x_j f_j(\mathbf{x})$$

- probability to learn L_i from an L_j-environment
- fitness (= abundance of offspring of users) of L_j
- abundance of infants that acquire L_i
- death rate
- velocity of change of abundance of L_i-speakers

Selection for learnability and fitness

Iterated language usage

- dynamics of E-language (= population of utterances)
- each utterance is produced and perceived by language users by means of underlying grammars (= I-languages)
- replication via imitation
- dynamics describes development of I-grammar frequencies within population of utterances

Iterated language usage (cont.)

Iterated language usage (cont.)

- simplest implementaton: **replicator dynamics**

$$\frac{dx_i}{dt} = x_i f_i(\mathbf{x}) - x_i \sum_j x_j f_j(\mathbf{x})$$

Iterated language usage (cont.)

- simplest implementaton: **replicator dynamics**

$$\frac{dx_i}{dt} = x_i f_i(\mathbf{x}) - x_i \sum_j x_j f_j(\mathbf{x})$$

- fitness of L_i (= expected number of imitations of an utterance from L_i)

Iterated language usage (cont.)

- simplest implementaton: **replicator dynamics**

$$\frac{dx_i}{dt} = x_i f_i(\mathbf{x}) - x_i \sum_j x_j f_j(\mathbf{x})$$

- fitness of L_i (= expected number of imitations of an utterance from L_i)
- abundance of utterances from L_i in next generation

Iterated language usage (cont.)

- simplest implementaton: **replicator dynamics**

$$\frac{dx_i}{dt} = x_i f_i(\mathbf{x}) - x_i \sum_j x_j f_j(\mathbf{x})$$

- fitness of L_i (= expected number of imitations of an utterance from L_i)
- abundance of utterances from L_i in next generation
- abundance of utterances from L_i in current generation

Iterated language usage (cont.)

- simplest implementaton: **replicator dynamics**

$$\frac{dx_i}{dt} = x_i f_i(\mathbf{x}) - x_i \sum_j x_j f_j(\mathbf{x})$$

- fitness of L_i (= expected number of imitations of an utterance from L_i)
- abundance of utterances from L_i in next generation
- abundance of utterances from L_i in current generation
- velocity of change of abundance of L_i-utterances

Iterated language usage (cont.)

- selection only for fitness — ignores learnability
- only homogeneous populations can be attractors

$\rightsquigarrow$ natural languages display high amount of optionality and non-determinism

Hybrid dynamics

- both modes of replication play a role in (cultural) language evolution
- adequate dynamics should capture both
- fitness of language is arguably negligible as factor for biological reproduction rate (at least on historical time scale)
- acqusition dynamics thus simplifies to

$$\frac{dx_i}{dt} = \sum_j x_j Q_{ji} - x_i$$

Hybrid dynamics (cont.)

- some fraction b $(0 \leq b \leq 1)$ of all utterances are uttered by language acquiring infants
- rest of utterances is uttered by adults and underlies the utterance dynamics
- leads to hybrid utterance dynamics:

$$\frac{dx_i}{dt} = (1-b)(x_i f_i - x_i \sum_j x_j f_j) + b(\sum_j x_j Q_{ji} - x_i)$$

selection for functionality and learnability

An example: Binding Theory

- Modern English: restrictions on coreference

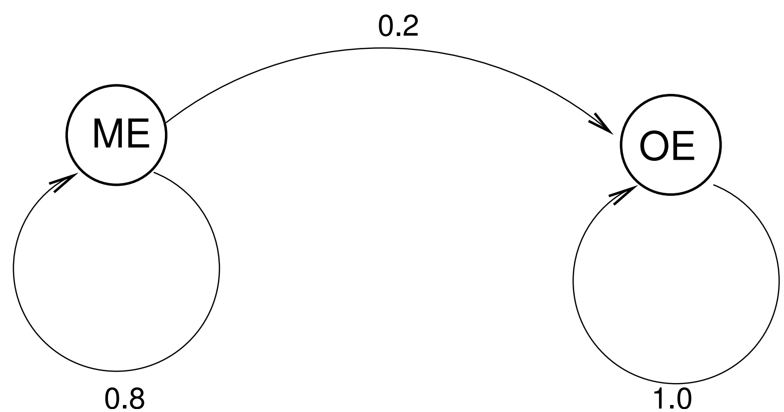
(4) a. Peter$_i$ sees him$_j$
    b. *Peter$_i$ sees him$_i$

- in Old English, (4b) is okay
- until a certain age, Modern English learning infants accept/produce structures like (4b)
- unlikely that OE infants underwent a stage corresponding to ME
- ME has less ambiguity and thus higher utility though

Binding Theory (cont.)

let us assume...

- acquisition probs.



- Q-matrix

	OE	ME
OE	1.0	0.0
ME	0.2	0.8

Binding Theory (cont.)

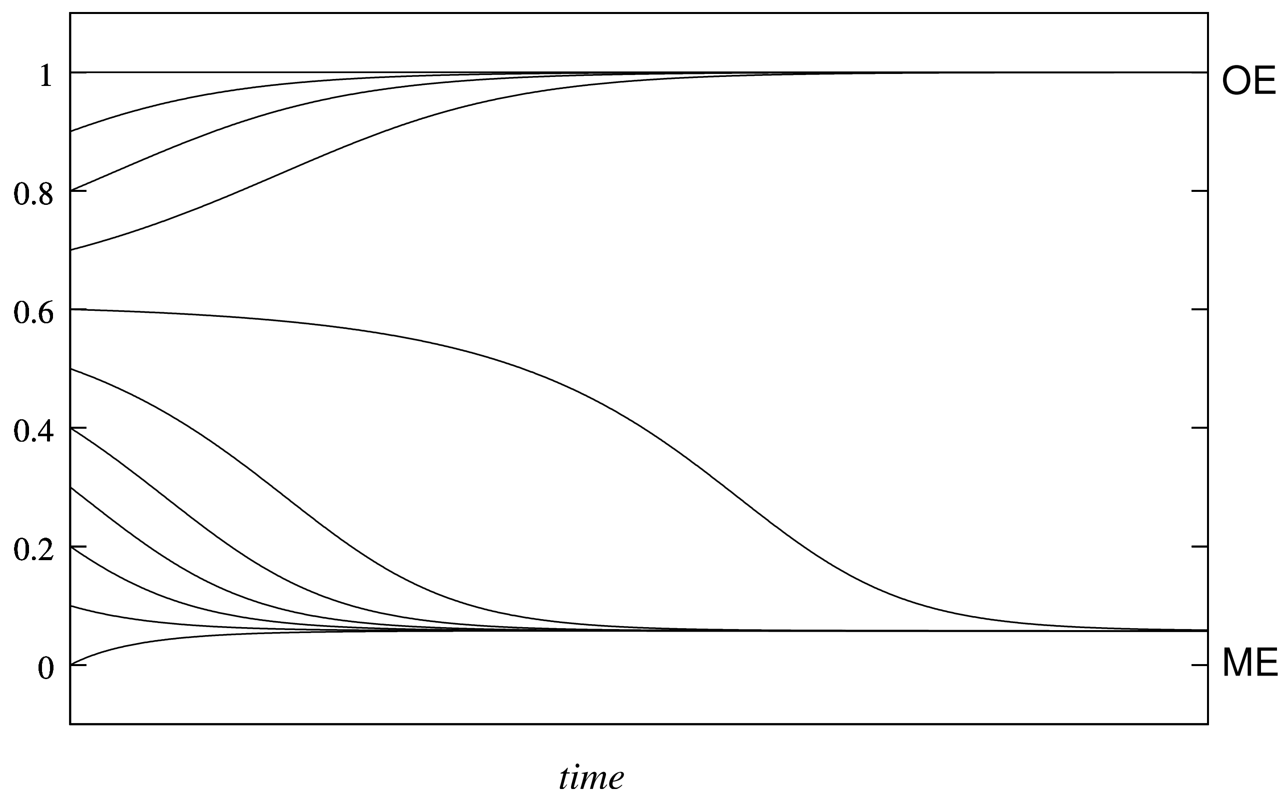
- U-matrix

	OE	ME
OE	0.9	0.8
ME	0.8	1

- $b = 0.05$

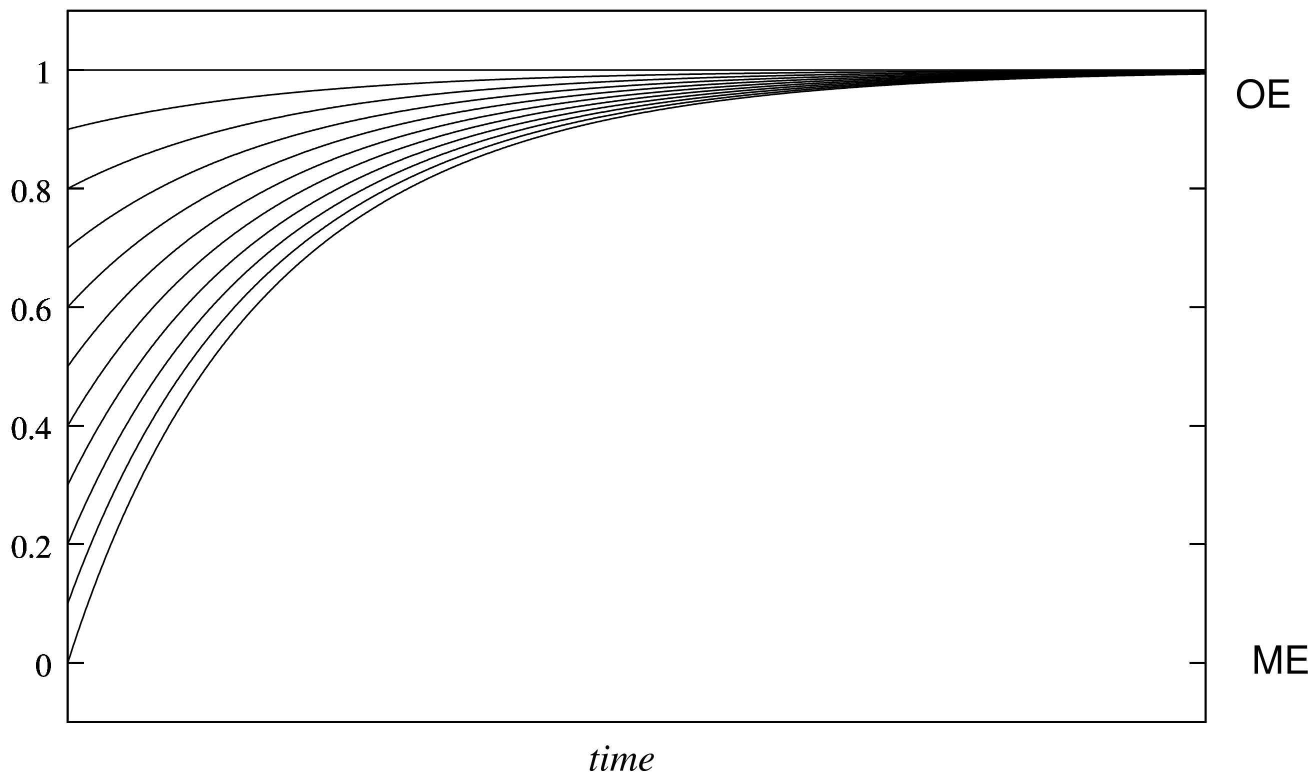
Binding Theory (cont.)

- two attractors (i.e. stable states)
  1. pure OE
  2. predominant ME (with a low probability of OE)



Binding Theory (cont.)

- acquisition dynamics also selects for high utility and high learnability
- learnability overrides utility though — only one attractor



Typology of case marking

- two kinds of accusative marking languages
  1. accusative is obligatory for all direct objects

     like Hungarian

(2) a. Szeretem a könyv**et**.
I-LIKE THE BOOK-ACC
"I like the book."

b. Egy ház**at** akarok.
A HOUSE-ACC I-WANT
"I want a house."

Typology of case marking (cont.)

2. accusative only on prominent object NPs

like Hebrew: only definites have accusative

(3) a. Ha-seret her?a **?et**-ha-milxama
THE-MOVIE SHOWED ACC-THE-WAR
b. Ha-seret her?a **(*?et-)**milxama
THE-MOVIE SHOWED (*ACC-)WAR
(from Aissen 2003)

Typology of case marking (cont.)

- utility matrix for competition between Hebrew and Hungarian type

	Hun	Heb
Hun	$.1100$	$.1060$
Heb	$.1060$	$.1734$

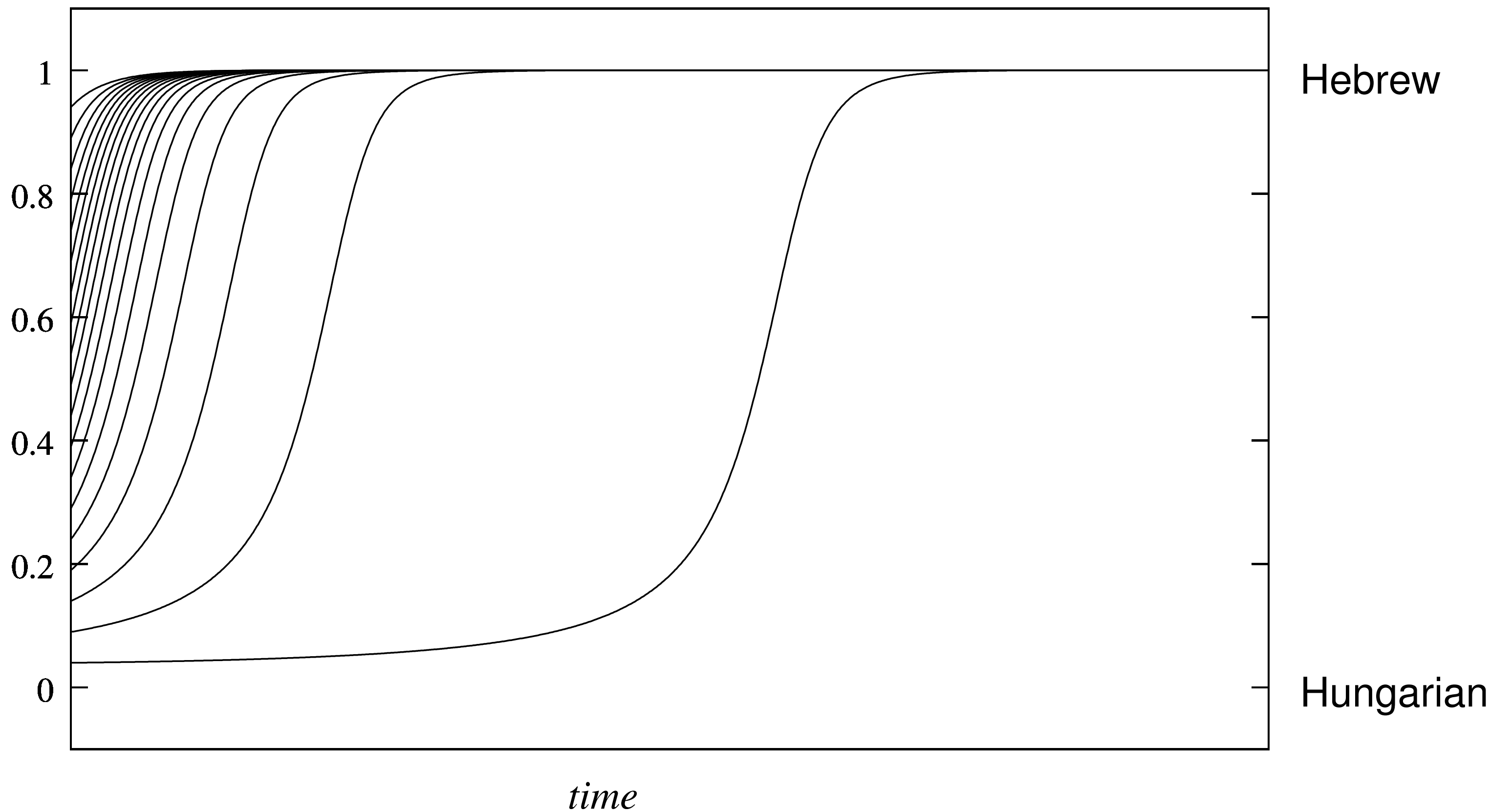
Typology of case marking (cont.)

- complicating factor: Hungarian style production grammar + Hebrew style comprehension grammar is also a possible language
- utility matrix for competition between Hebrew and Hungarian type

	Hun	Hun/Heb	Heb
Hun	.1100	.1100	.1060
Hun/Heb	.1100	.1100	.1417
Heb	.1060	.1417	.1734

Typology of case marking (cont.)

- usage dynamics predicts only Hebrew to be stable



Typology of case marking (cont.)

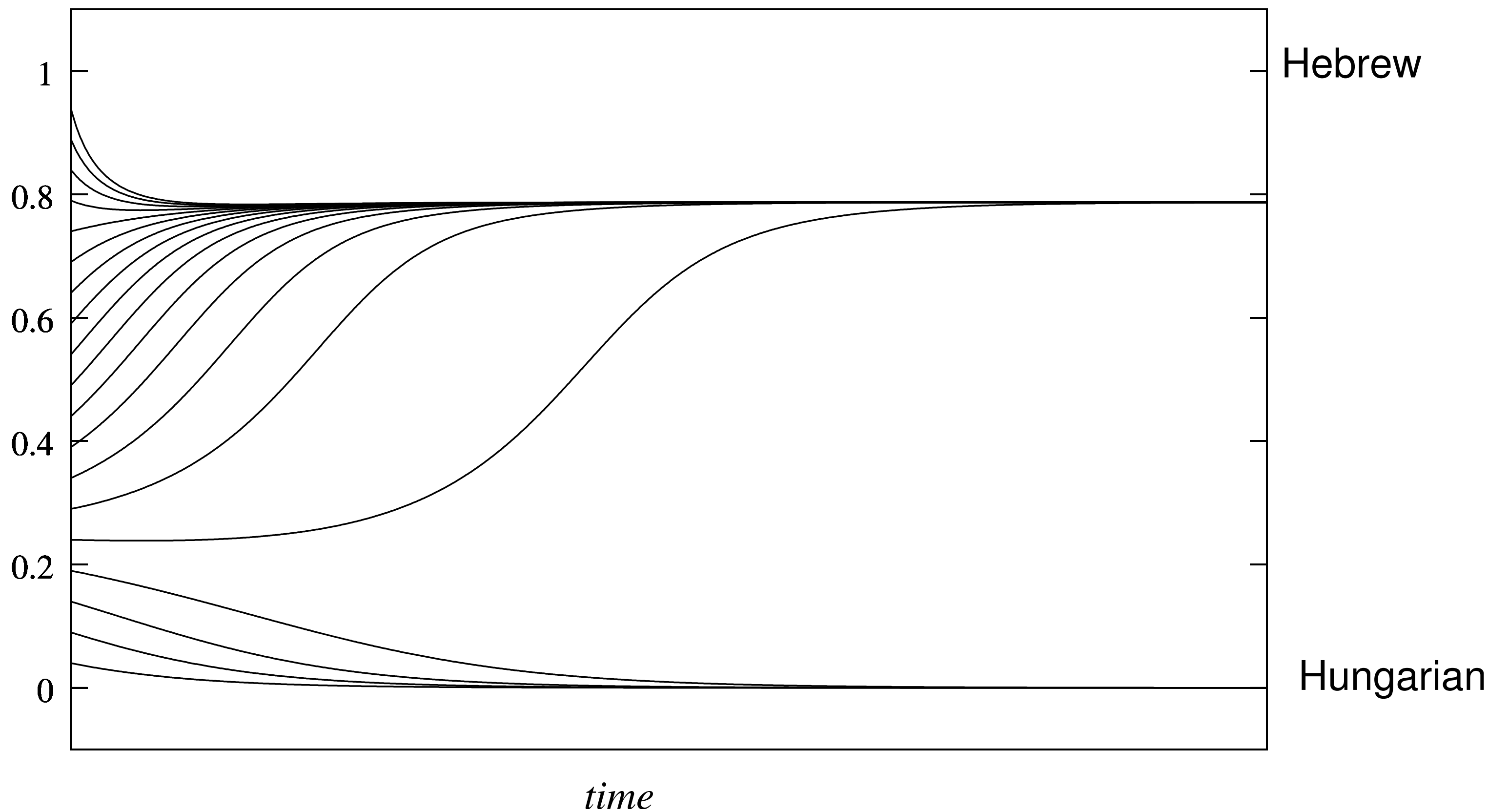
- Hungarian system ("All objects have accusative!") is arguably simpler than Hebrew system ("All **definite** objects have accusative!")
- acquistion matrix something like

	Hun	Hun/Heb	Heb
Hun	1.0	0.0	0.0
Hun/Heb	0.0	1.0	0.0
Heb	0.1	0.0	0.9

- $b = 0.1$

Typology of case marking (cont.)

- under hybrid dynamics (as under acqisition dynamics) both Hungarian and Hebrew style case systems are evolutionarily stable



Conclusion

- natural languages are shaped both by selection for learnability and selection for usability
- corresponds to replication via acqusition and replication via usage
- combined dynamics leads to refined typological predictions

Conclusion (cont.)

Question for future research

- *How can the parameters of these equations (fitness, learnability matrix) be determined in a non-circular way?*
- *Can we observe micro-evolution directly (psycholinguistics, corpus linguistics, ...) to validate formal models?*

Possible refinements

- Spatial EGT:
  - individuals are organized in a spatial structure
  - interaction mostly with neighors
  - offspring remain in neighborhood
  - ⇝ many interactions between kins
  - ⇝ fosters cooperation
- Network models
  - similar to spatial EGT
  - except: interaction determined by network relationships
  - fast spread of innovations in centralized networks
  - ...