

# Iterated Learning – Lecture 7

Jelle Zuidema

Institute for Logic, Language and  
Computation, Amsterdam

Evolution of Language 2014

# Components of an evolutionary explanation

- Explicitness requirement
  - Heritability
  - Strategy set
  - Fitness
  - Path of ever increasing fitness
    - starting point, end point, process

# Path: starting point

- We can try to reconstruct the evolutionary path through the *comparative method*
- All life originates from a single source – tree of life emerged from *speciation* events
- Every two species have a *last common ancestor* (LCA) a share traits by common descent (“*homology*”)
- Genetic or phenotypic distance can be used to reconstruct phylogenetic tree
- Distribution of traits over phylogenetic tree used to reason about traits LCA

# Path: end point

- We want an evolutionary explanation for modern language abilities;
- Many unresolved controversies about language universals and domain specificity of learning and processing biases
- Relative consensus about high-level “design features”:
  - symbolism,
  - combinatorial phonology
  - compositional semantics
  - hierarchical structure

# Path: process

- Fossil record:
  - no clear data about language origins
  - clear increase in brain size in last 2.5My
  - loss of air sacs between 2.5 and 1My BP
  - symbolic revolution from around 120Ky BP
- Evidence for selection
  - comparative method: analogy
    - problem: lack of model species (birds?)
  - population genetics: genetic hitchhiking
    - problem: lack of identified genes (FoxP2?)
  - optimality/ “argument from design”

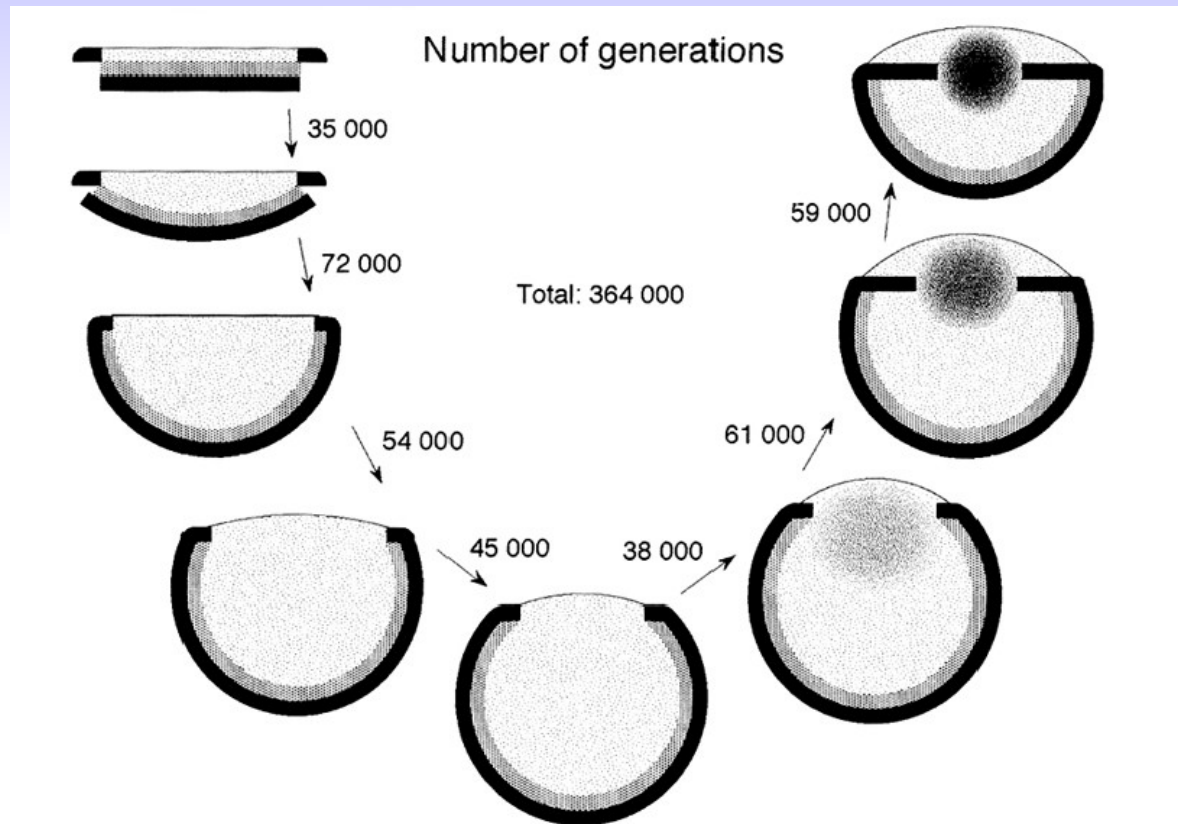
# Optimality

- Arguments about optimality can be sound in evolutionary reasoning, but must be handled with care
- In language evolution scenarios, we meet two special classes of challenges, having to do with language being:
  - A social trait
  - A culturally transmitted system

# Special challenge I: social traits

- Coevolution of senders and receivers
- Their interests do not necessarily align
- Mistake to assume receiving behavior will remain unchanged, while sending behavior evolves or v.v.
- Evolutionary Game Theory

# Evolution of the eye





## Language is not an eye!

Often, the fitness of an individual with a given phenotype does not only depend on the phenotype and environment (including other species), but also on the *frequency* of the phenotype in the population.

This is called: *Frequency-dependent Selection*

The prime example is the evolution of (code for) communication.

### Sight

individual ↓	population	
	bad eyes	good eyes
bad eyes	low	low
good eyes	high	high

### Communication

individual ↓	population	
	code A	code B
code A	high	low
code B	low	high

**Problem of cooperation** Why would senders be willing to send honest signals, and hearers be willing to receive and believe the signal?

Honest signaling theory (Zahavi, Maynard Smith, Grafen, Bergstrom)

**Problem of coordination** How is, after each innovation, a shared code established and maintained? And which code?

Coordination games (Lewis, Skyrms, Nowak, Hurford)

# The evolution of cooperation

- Kin selection: altruistic traits can be favored by natural selection under specific circumstances (Hamilton'64; West & Gardner'05)
  - carriers of the gene(s) for the trait must interact preferentially with other carriers (e.g., kin), while still competing with non-carriers
- Reciprocal altruism: conditional cooperation (tit-for-tat) can be favored under specific circumstances (Trivers, Axelrod)
  - individuals have repeated interactions and track past behavior

# Special challenge II: cultural traits

- Coevolution of languages and language users
- Cultural evolution of the appearance of design
- Moving targets argument (against arbitrary innate language universals)



# Feher et al. 2006

- After 7 generations, no demonstrable difference with natural song
- Zebra finch songs thus appears to be completely determined by their biology, but:
  - not an innate template, but determined by cognitive as well as anatomical and environmental factors;

# Is zebrafinch song learned or innate?

- The important features of zebrafinch song are obviously *learned*: without the exposure to songs in the sensitive period birds do not learn the species-typical song;
- The important features are obviously *innate*: all the knowledge of the species-typical song is apparently contained in their biology.

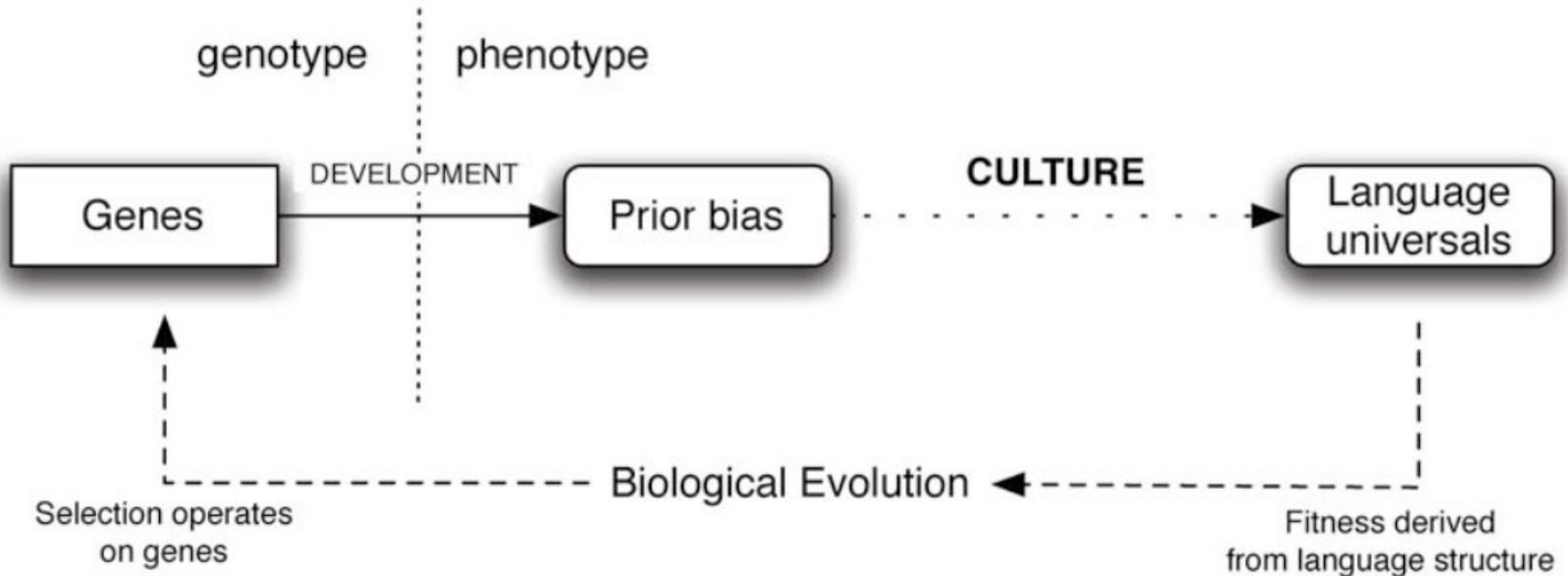
# Iterated Learning



(Kirby, 1999, 2000; Brighton, 2002; Zuidema 2003)



# Ontogeny, glossogeny, phylogeny



# Iterated Learning

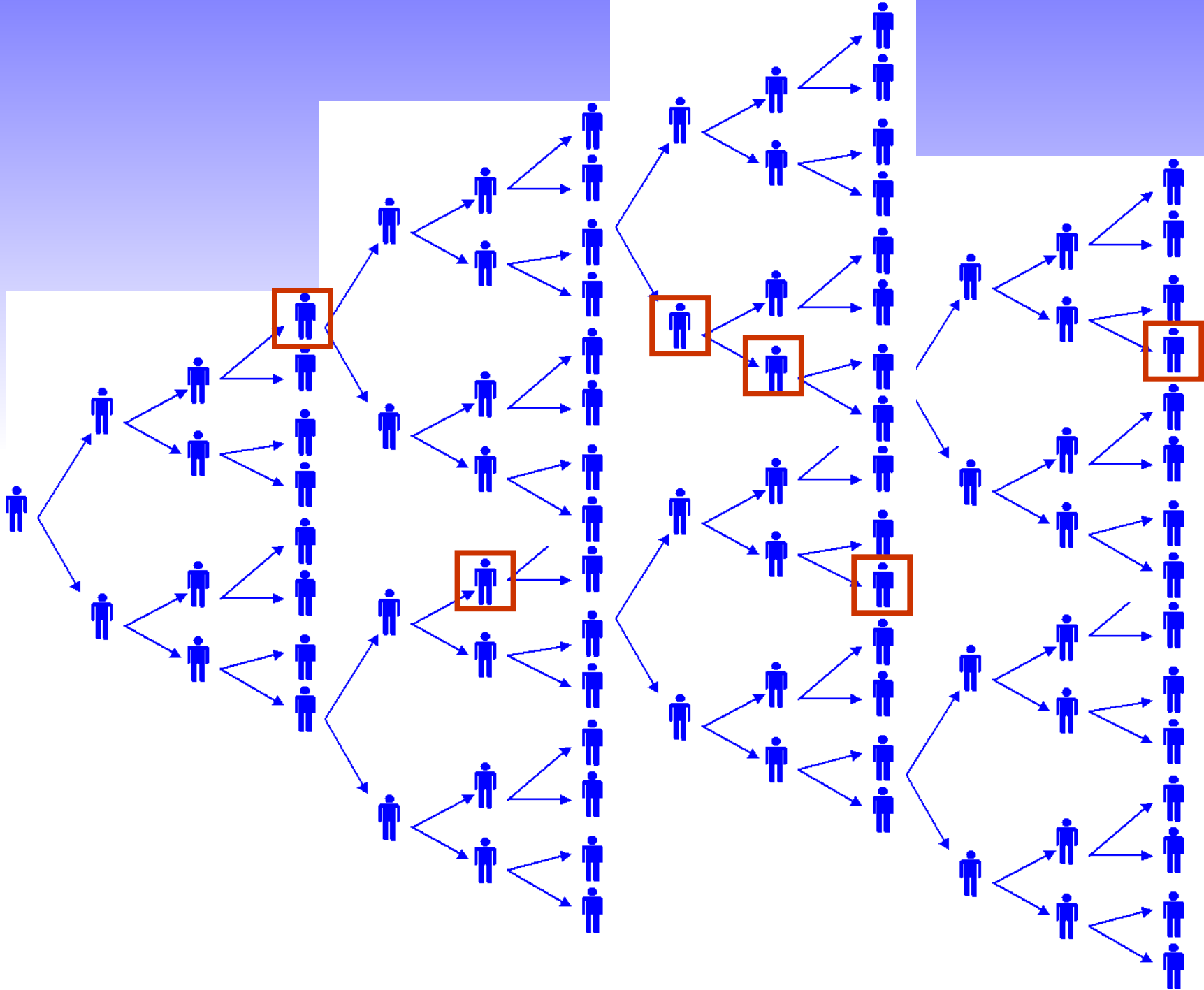
- Learners learn from learners
- Cultural transmission

Claim: in cultural transmitted systems one has to be extra careful to classify things as nature or nurture, adaptation or side-effect

# Cultural Adaptation Hypothesis

- Language is culturally transmitted and adapts to the *biases* of the learners;
- Given: population with many L1 and few L2 learners + unavoidable differences in learning biases:

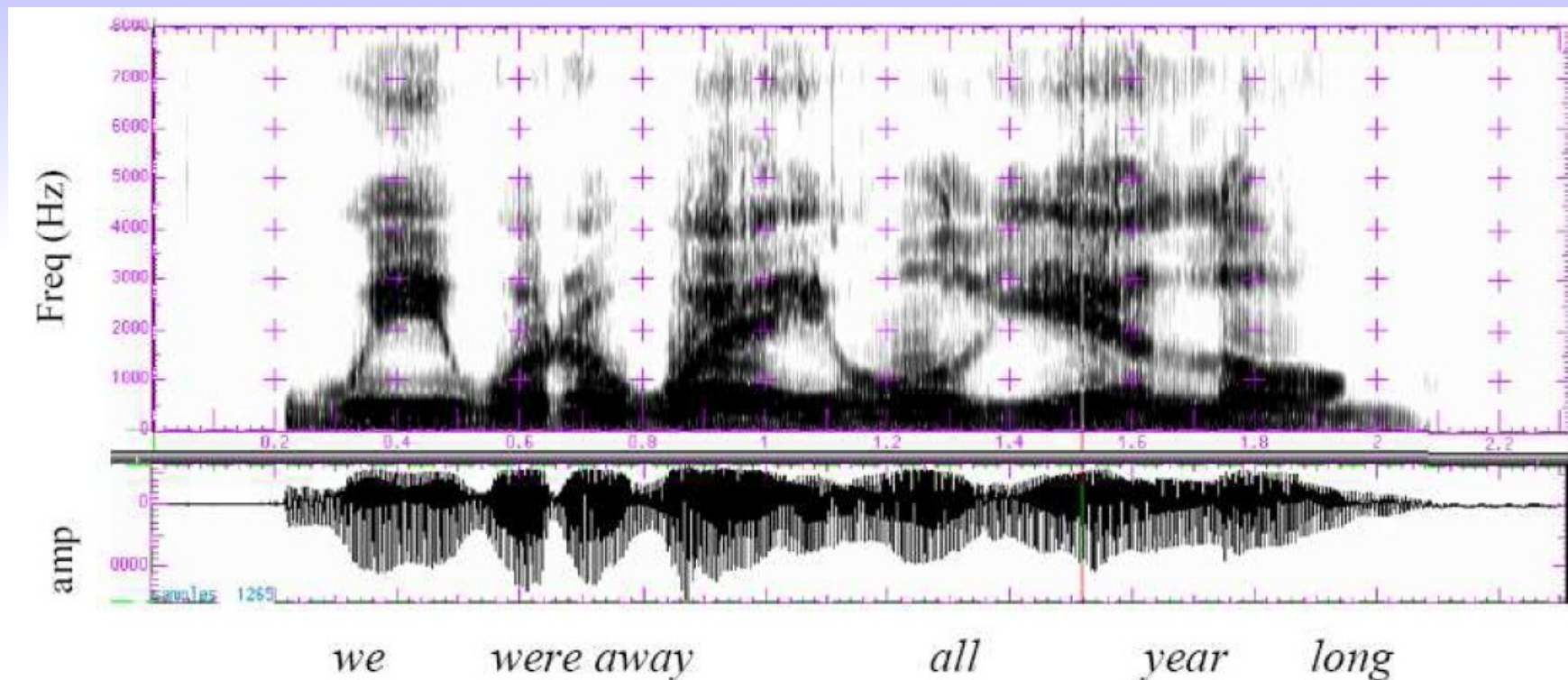
What will happen to the language in iteration?

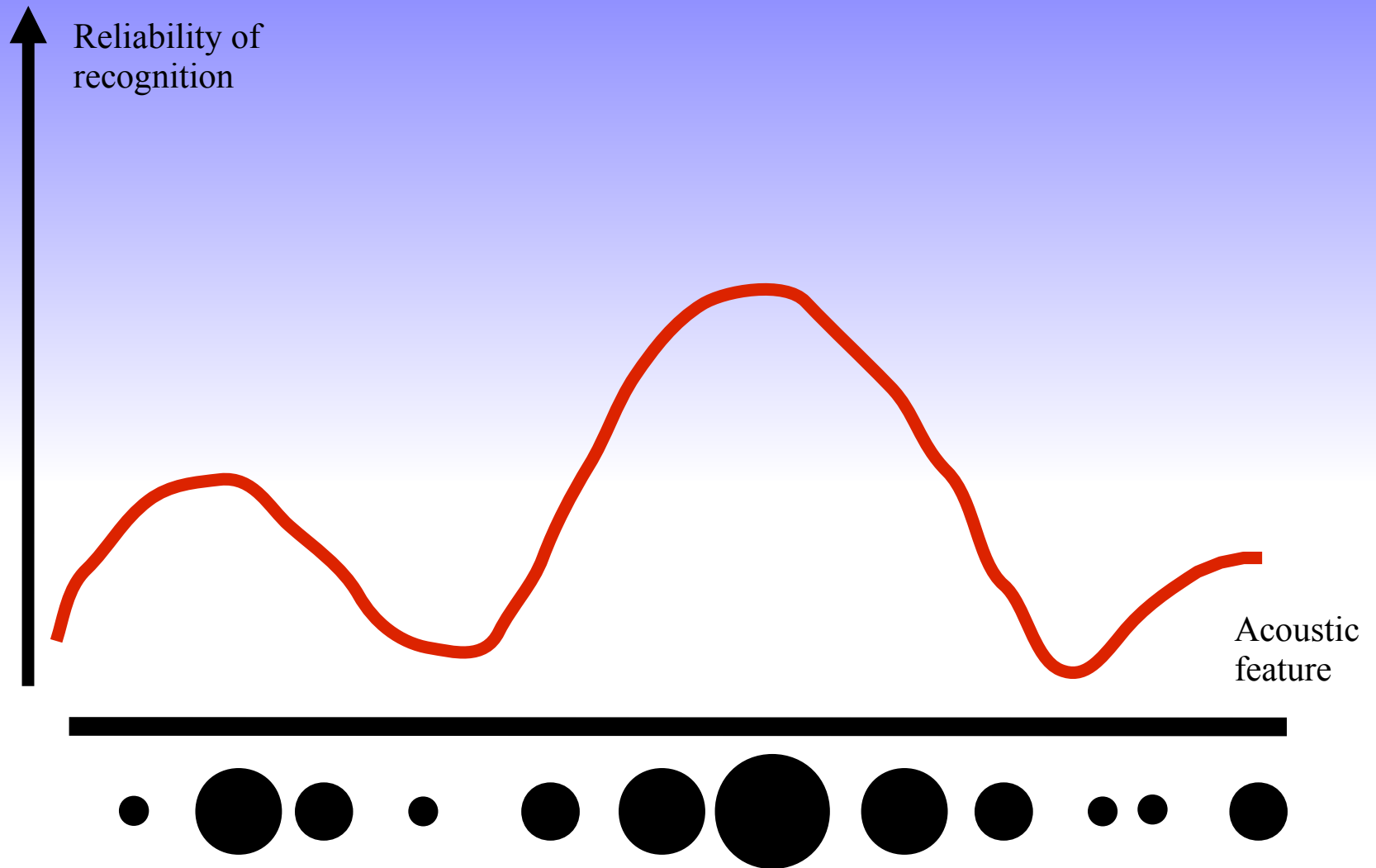


# Cultural Adaptation & the Critical Period

- In a population with many L1 and few L2 learners...
- With arbitrary, language-independent differences in their learning biases
- And cultural transmission of language...
- Language will adapt to the biases of L1 language learners, and L2 acquisition will be more difficult than L1.
- Critical Period effects come for free!

# Is human auditory perception specialized for speech?





(We choose random values from  $[0,1]$  on the diagonal of  $U$ ).

# 4-matrix model: payoffs

(Zuidema & Westermann, 2003)

**Speaker behavior** **Hearer behavior**  
(learned) (learned)

$$p \ a \ y \ \neq \ \sum_m \sum_f \sum_{f'} \sum_{m'} S_m U_{ff'} R_{f'm'} V_{m'm}$$

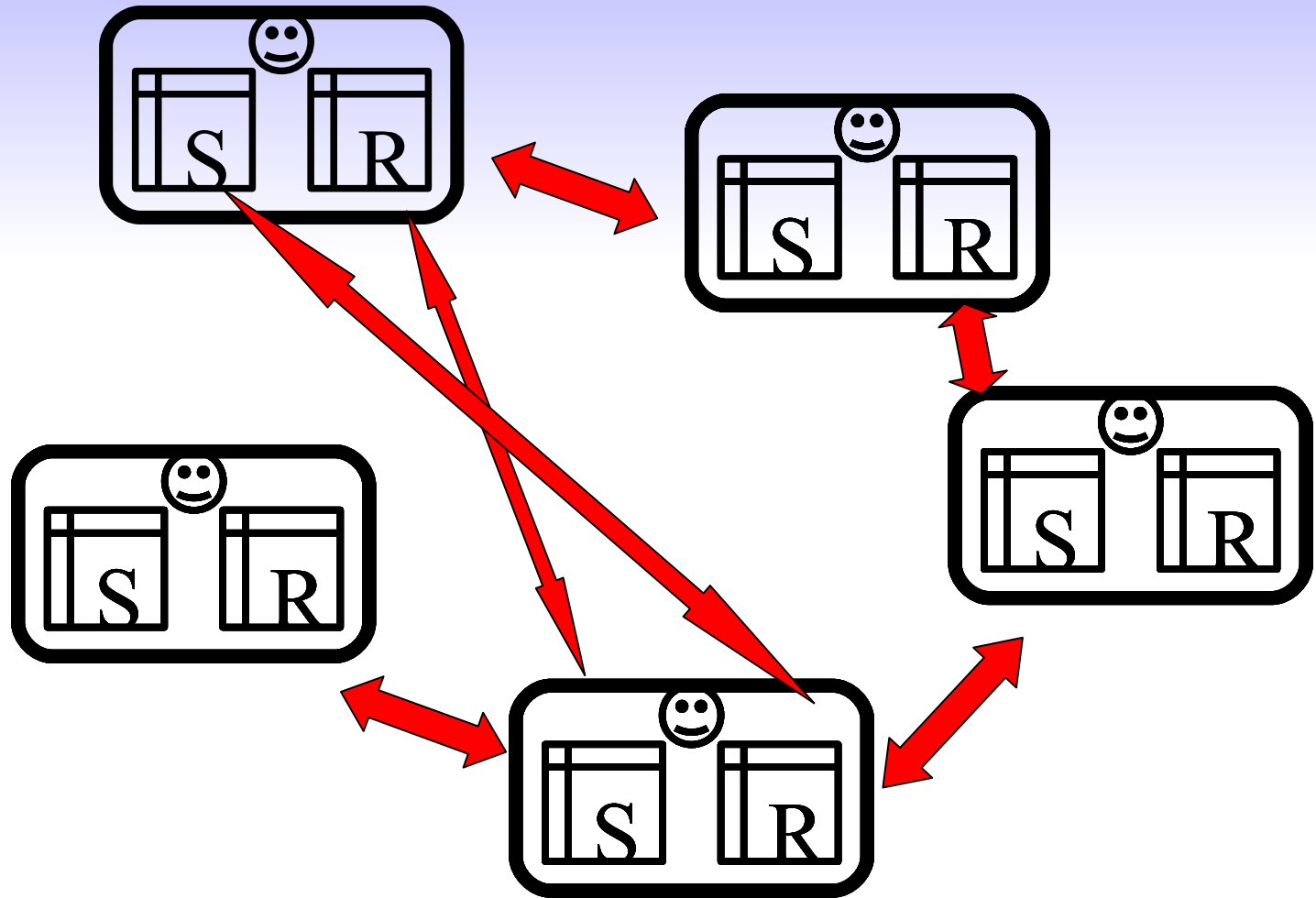
The diagram shows the payoff matrix equation:  $p \ a \ y \ \neq \ \sum_m \sum_f \sum_{f'} \sum_{m'} S_m U_{ff'} R_{f'm'} V_{m'm}$ . Red arrows point from the explanatory text below to specific parts of the equation: one arrow points from the text 'Confusability of alternate signals...' to the  $U_{ff'}$  term, another from 'Values of alternate interpretations...' to the  $V_{m'm}$  term, and a third from the 'Speaker behavior (learned)' label to the  $S_m$  term.

**Confusability of alternate signals:**  
articulation, acoustics, perception  
(assumed constant)

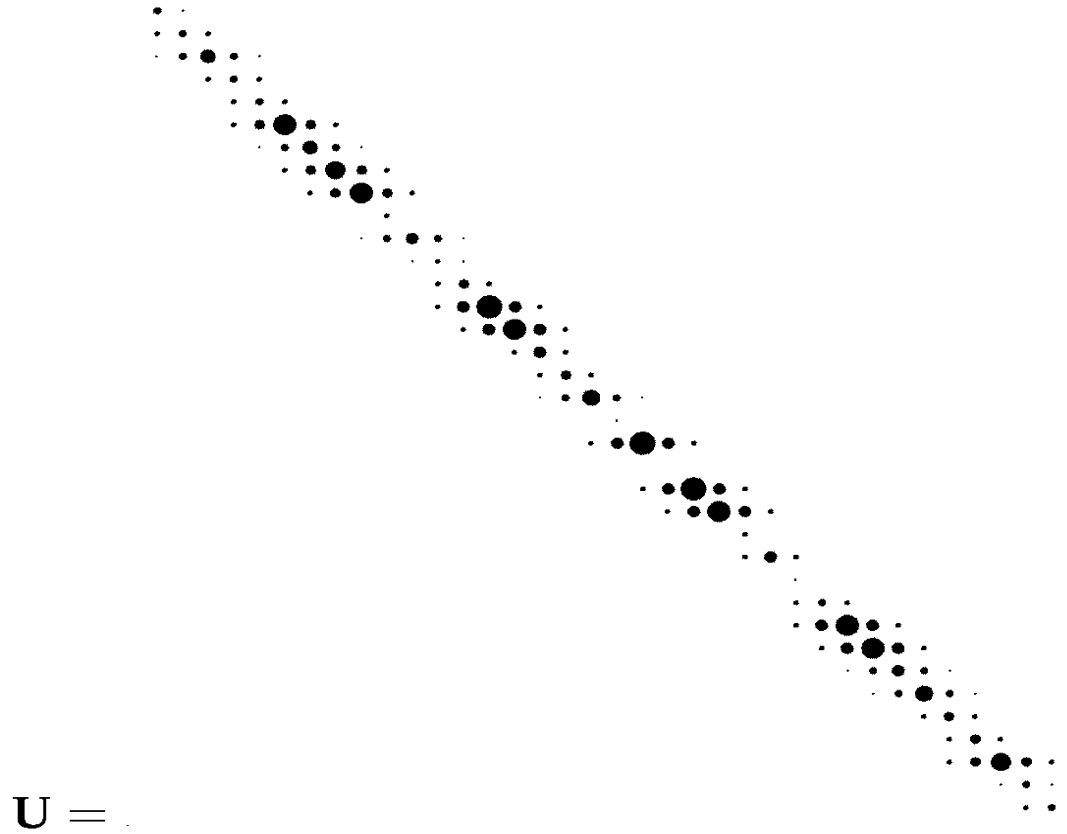
**Values of alternate**  
**interpretations**  
(assumed constant)



Agents in a population maximize payoff by adapting S and R  
(local hillclimbing)



## Random Perceptual Characteristics...



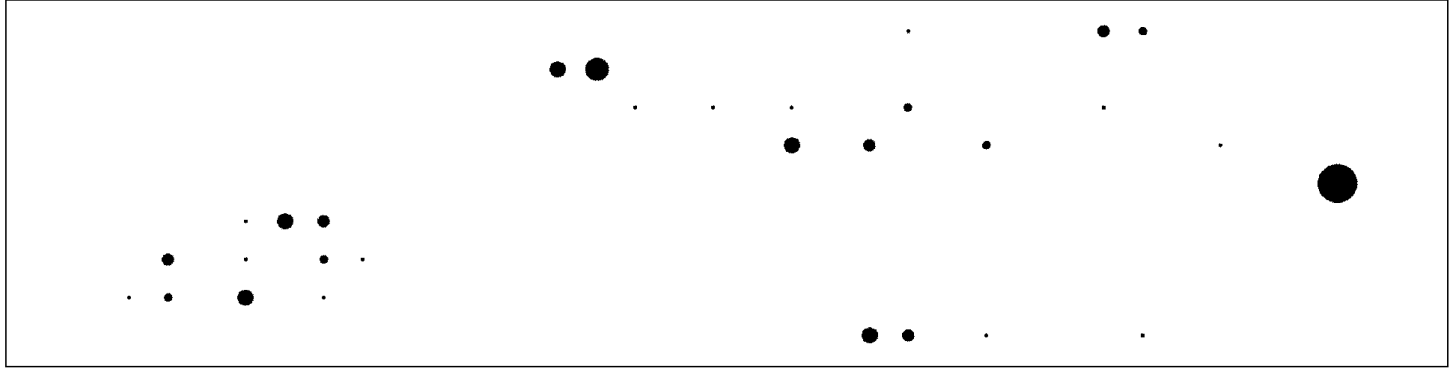
(after 100,000 learning steps)

U



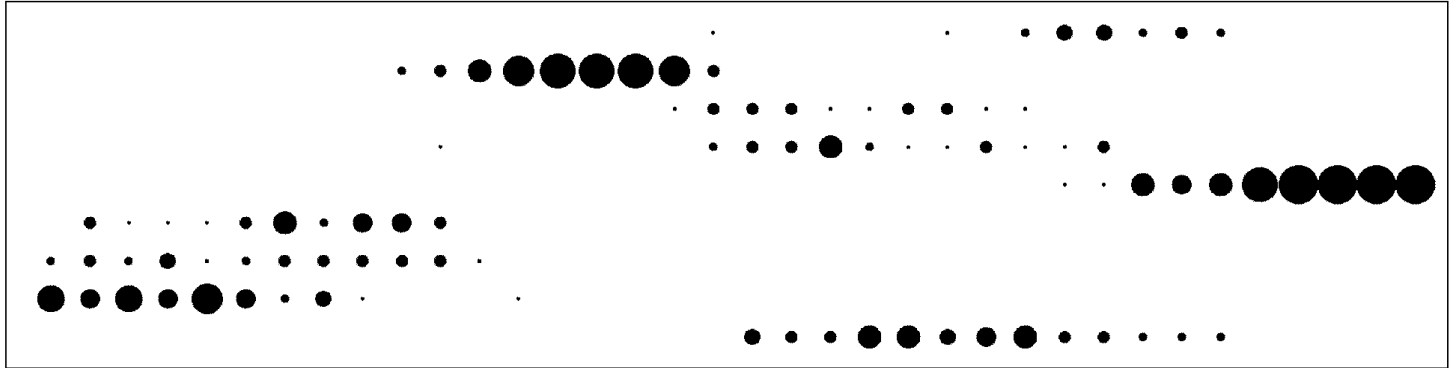
100000 S

S



R

$R^T$



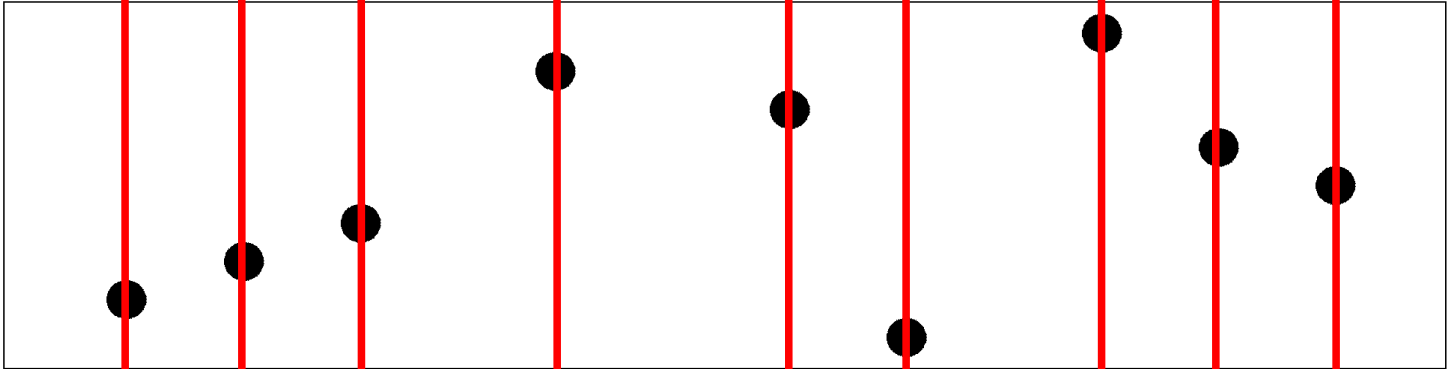
(after millions of learning steps)

U



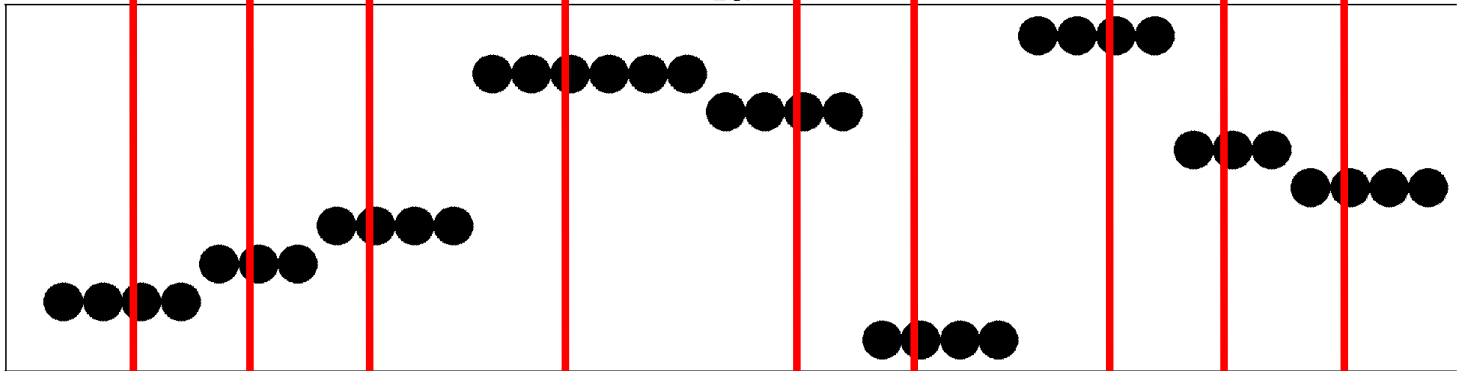
1000000 S:

S



R:

$R^T$



# Negative conclusions

- There is a tight fit between the perceptual characteristics and the “speech code”;
- But in this model, the perceptual characteristics did not evolve/adapt;
- Rather, through cultural evolution, the language has evolved to match perception;
- The model thus shows that **appearance of design cannot be taken as conclusive evidence** for (genetic) adaptation.

# Lessons learned

- Language and bird song are culturally transmitted systems
- Details of the final system are the result of interactions between
  - Innate biases,
  - perception,
  - production,
  - cognitive processing
  - in a population

# The real issues

- The real issue is not *nature or nurture*, but the *nature of nurture*.
- What are the biases that guide and constrain the learning process?
- ... how do they affect language & universalia?
- ... where do they come from?
- Are there language-specific adaptations that emerged in human evolution: specialization for language

# Poverty of the stimulus & language universals

- (1) The man is mean
- (2) Is the man mean?
- (3) The man who is feeding a donkey is mean
- (4) Is the man who is feeding a donkey \_ mean?
- (5) \*Is the man who \_ feeding a donkey is mean?





# 1980s nativism

“Any aspect of language that the speaker knows must either be learnable from positive evidence, that is to say, through exposure to sentences of the language, or be part of the innate equipment of the human mind” (Cook, 1983)

“Language shows signs of complex design for the communication of propositional structures, and the only explanation for the origin of organs with complex design is the process of natural selection.” (Pinker & Bloom, 1990)

# 1980s nativism

“Any aspect of language that the speaker knows must either be learnable from positive evidence, that is to say, through exposure to sentences of the language, or be part of the innate equipment of the human mind” (Cook, 1983)

“Language shows signs of complex design for the communication of propositional structures, and the only explanation for the origin of organs with complex design is the process of natural selection.” (Pinker & Bloom, 1990)

# 1980s nativism

“Any aspect of language that the speaker knows must **either** be learnable from positive evidence, that is to say, through exposure to sentences of the language, **or** be part of the innate equipment of the human mind” (Cook, 1983)

“Language shows signs of complex design for the communication of propositional structures, and the **only explanation for the origin of organs with complex design** is the process of natural selection.” (Pinker & Bloom, 1990)

# Universal Grammar

*Evolution*

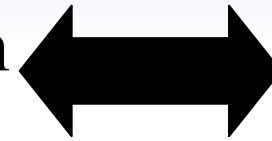
*Cognitive  
Science*

*Comparative  
Linguistics*

Language-  
specific  
adaptations



Language  
acquisition  
device

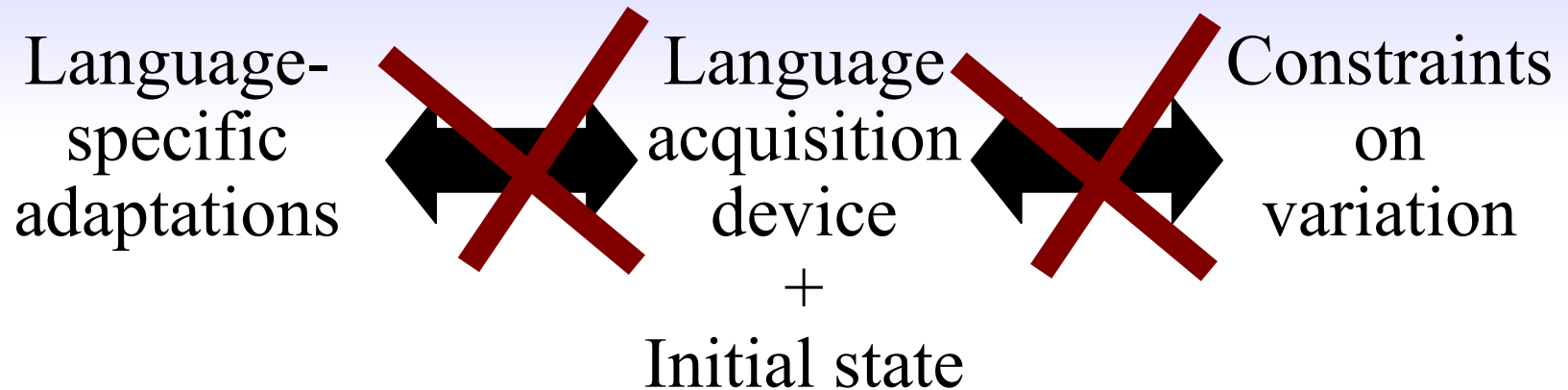


Constraints  
on  
variation

+

Initial state

# Universal Grammar



Iterated Learning:  
“appearance of design”  
can occur even without  
any biological  
evolution  
(Kirby, 1994/2000)

# Arguments for extensive language-specific adaptations

- Appearance of design
- Poverty of the stimulus
  - Formal learnability results
- Critical period
- Constraints on variation

# Conclusions: language-specific adaptations?

- Not the '*poverty of the stimulus*'
- Not the '*appearance of design*'
- Not the 'critical period'

How many of the (unique) properties of language and language acquisition are the result of language-independent cognitive skills + cultural transmission?



# Fallacies

- selection for the group,
- confusing Tinbergen's 4+1 why's,
  - Treating ultimate, proximate or glossogenetic explanations as mutually exclusive
  - Ignoring glossogeny (ignoring cultural nature)
  - Assuming the receiver's *mechanisms* are constant while the sender's are evolving or v.v. (ignoring social nature of language)
- applying models with implicit human-specific assumptions to other animals,
- assuming miracles
- argument from personal incredulity or authority

## Theories of the evolution of language

(Szamado & Szathmary, 2006)

**Hunting theories:** 'our intellect, interests, emotions, and basic social life all are evolutionary products of the success of the hunting adaptation.' (Washburn/Lancaster'68) use of language was to coordinate the hunting effort of the group.

**Motherese:** language evolved in the context of motherchild communication. Mothers had to put down their babies to collect food efficiently, and their only option to calm down babies was to use some form of vocal communication (Falk'04).

**Group bonding and/or ritual:** language evolved in the context of intergroup rituals, which first occurred as a kind of 'strike action' against non-provisioning males. Once such rituals were established, a 'safe' environment was created for further language evolution (Knight'98).

**Gossip:** menstrual ritual can be a costly signal of commitment; hence participating in such rituals can create female groups of shared interest in

which sharing information about the social life of others (i.e. gossiping) can be beneficial (Power'98).

**Status for information:** language evolved in the context of a so-called asymmetric cooperation, where information (that was beneficial to the group) was traded for status (Desalles'98).

**Sexual selection:** language is a costly ornament that enables females to assess the fitness of a male. According to this theory, language is more elaborate than a pure survival function would require (Miller'01).

**Language as a mental tool:** language evolved primarily for the function of thinking and was only later co-opted for the purpose of communication (Burling'93; also Chomsky's favourite just-so story).

**Grooming hypothesis:** language evolved as a substitution for physical grooming (Dunbar'98). The need for this substitution derived from the increasing size of the early hominid groups, which mean that physical grooming became more time consuming, whereas it was possible to 'groom' more than one individual simultaneously via vocal communication.

**Mating contract and/or pair bonding:** the increasing size of the early hominid groups and the need for male provisioning also necessitated 'social contract' between males and females (Deacon'97).

**Song hypothesis:** language evolved rapidly and only recently by a process of cultural evolution. The theory assumes two important sets of preadaptations; one is the ability to sing; the other is better representation abilities (i.e. thinking and mental syntax) (Vannechoutes/Skoyles'98).

**Tool making:** assumes a double homology: 'a homologous neural substrate for early ontogeny of the hierarchical organisations shared by two domains language and manual object combination and a homologous neural substrate and behavioural organisation shared by human and non-human primates in phylogeny. (Greenfield'91)

# Uniquely human traits

- Language
- Complex reasoning
- Mathematics
- Music
- Consciousness
- Music
- Cooperativity
- Life history, upright posture, opposable thumb, running

# What has happened?

- Earliest evidence for some (very limited) aspects of 'humanness' 2.5My BP – million years of stasis
- Last common ancestor all humans:
  - 140Ky-290Ky BP (mtDNA)
  - 70Ky BP (Ychromosome)
- Symbolic revolution +/-120Ky BP – enormous speed-up in developments
- Between 2.5My and 100,000 years BP something very significant has happened in hominin evolution!

# Requirements for plausible scenarios

- Explain how such a radical new phenotype can be based on relatively few genetic changes
  - Common causes
  - Hidden potential
- Explain how the unusual circumstances needed for the evolution of social traits can be sustained
  - Self-enforcing dynamic



# Requirements for plausible scenarios

- Explain how such a radical new phenotype can be based on relatively few genetic changes
  - Common causes – cognitive technology
  - Hidden potential – cultural evolution of language
- Explain how the unusual circumstances needed for the evolution of social traits can be sustained
  - Self-enforcing dynamic – green beard dynamics



# Linguistic sweep scenario

## Pre-existing

- hierarchical, conceptual structure
- non-combinatorial communication
- limited cooperativity & social cognition
- hidden potential for more complex cognition

*biological evolution*



- Biological adaptations to new niche
- larger social groups
  - increases in social intelligence, cooperativity & communication
  - increased reliance on learned, combinatorial signaling

*cultural evolution*



## Cultural adaptations

- learned communication system adapts to preexisting biases of hominin brain (can thus be much more complex than random code)
- communication system becomes representational system for internal thought too
- knowledge transfer from previous generations unlocks potential for complex cognition

*biological evolution*



- New cultural niche
- creates intense selection pressure for linguistic & cognitive skills

*cultural evolution*



# Conclusions

- Evidence on human evolution points to a radical change in the human phenotype >100.000 yrs BP
- Many human-specific traits are 'social' – require unusual circumstances to evolve
  - plausible scenarios need to show how
- Plausible scenarios for the evolution of humanness involve
  - common causes for various traits
  - ways to unleash hidden potential
- The language-first scenario does all this.