Formal Models of the Evolution of Language

Jelle Zuidema

Institute for Logic, Language & Computation

Core Logic, Guest Lecture

"It is astonishing what language can do. With a few syllables it can express an incalculable number of thought, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thoughts corresponding to the parts of a sentence, so that the structure of the sentence serves as the image of the structure of the thoughts." – Gottlob Frege (1923)

(Examples from Comrie, 1981)

Isolating languages: No morphological variation for tense, case or plurality; Each word typically consists of a single morpheme. E.g. Vietnamese "*Khi tôi dên nhà ban tôi, chúng tôi*"

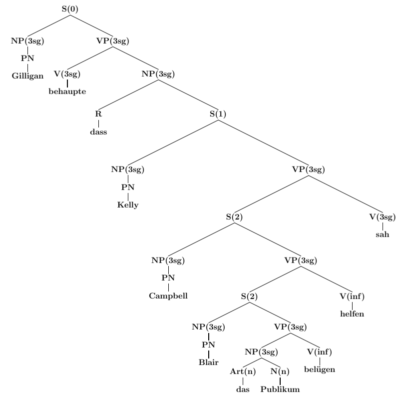
Khi tôi dên nhà ban tôi, chúng tôi bắt dâu làm bài.
when I come house friend I PLURAL I begin do lesson
"When I came to my friend's house, we began to do lessons."

Polysynthetic languages: Many lexical morphemes combined in a single word. E.g. Chukchi (Siberia) "*tɛmeyŋɛlevtɛpɛγtɛrkɛn*"

tɛ- meyŋɛ- levtɛ- pɛγt- ɛrkɛn
great head ache 1stSINGULAR IMPERFECT
"I have a fierce head-ache"

(1) a. Gilligan claims that Blair deceived the public.
b. Gilligan claims that Campbell helped Blair deceive the public.
c. Gilligan claims that Kelly saw Campbell help Blair deceive the public. (tail recursion)

(2) a. Gilligan behaupte dass Kelly Campbell Blair das Publikum belügen helfen sah. (center embedding)
b. Gilligan beweert dat Kelly Campbell Blair het publiek zag helpen bedriegen. (crossing dependencies)



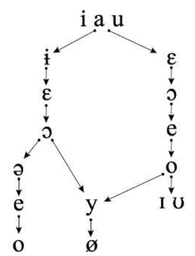


Figure 6.2: Vowel system hierarchy according to Crothers (1978).

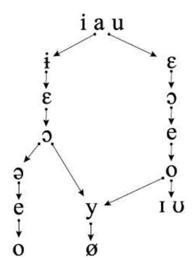
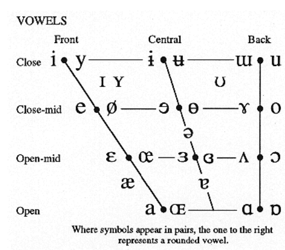


Figure 6.2: Vowel system hierarchy according to Crothers (1978).



What is the scope of existing linguistic phenomena, and what are the constraints on variation? (typology)

What is the scope of existing linguistic phenomena, and what are the constraints on variation? (typology)

Can we give a precise, predictive characterisation of language structure? ("periodic table")

What is the scope of existing linguistic phenomena, and what are the constraints on variation? (typology)

Can we give a precise, predictive characterisation of language structure? ("periodic table")

How do children acquire the complexities of language?

What is the scope of existing linguistic phenomena, and what are the constraints on variation? (typology)

Can we give a precise, predictive characterisation of language structure? ("periodic table")

How do children acquire the complexities of language?

Why are languages the way they are?

What is the scope of existing linguistic phenomena, and what are the constraints on variation? (typology)

Can we give a precise, predictive characterisation of language structure? ("periodic table")

How do children acquire the complexities of language?

Why are languages the way they are?

Plan for today

- The structure of evolutionary explanations;
- Evolutionary Game Theory;
- Communication as a Game;
- Biological Evolution;
- Cultural Evolution.

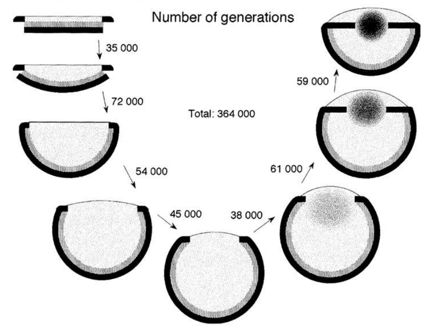


Fig. 1.6 A patch of light sensitive epithelium can be gradually turned into a perfectly focussed camera-type eye if there is a continuous selection for improved spatial vision. A theoretical model based on conservative assumptions about selection pressure and the amount of variation in natural populations suggest that the whole sequence can be accomplished amazingly fast, in less than 400 000 generations. The number of generations is also given between each of the consecutive intermediates that are drawn

3 key elements in an evolutionary scenario

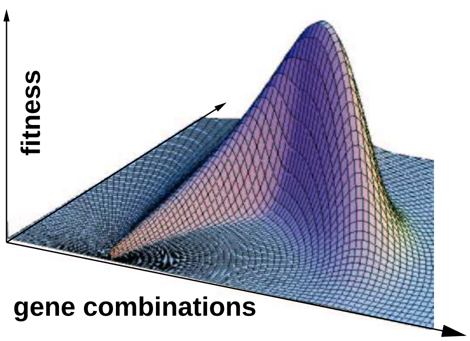
1. What is the scope of phenotypes that are "available" for evolution? . . . . strategy set
2.
3.

3 key elements in an evolutionary scenario

1. What is the scope of phenotypes that are "available" for evolution? . . . . strategy set
2. How well does each of these possible phenotypes fare? . . . . fitness function
3.

3 key elements in an evolutionary scenario

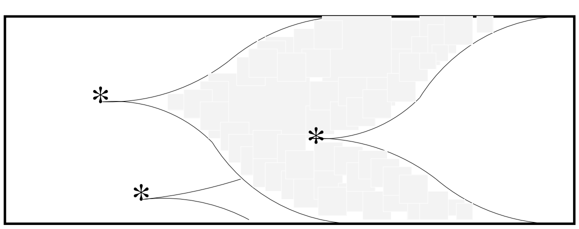
1. What is the scope of phenotypes that are “available” for evolution? . . . . strategy set
2. How well does each of these possible phenotypes fare? . . . . fitness function
3. Is there a sequence of possible phenotypes, each next one fitter than the previous, such that in can invade? . . . . fit intermediates



Limits to Optimality

“Natural selection tends only to make each organic being as perfect as, or slightly more perfect than, the other inhabitants of the same country with which it comes into competition. And we see that this is the standard of perfection attained under nature” (Darwin, 1872, p 163)

- biophysical and genetic constraints
- the speed of evolution
- mutational load
- fluctuating fitness
- frequency-dependent fitness
- correlation, levels of selection



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.

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 Communication

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

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

Communication

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.

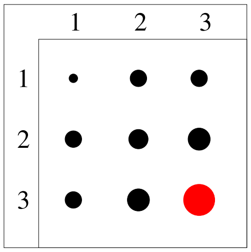
Sight

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

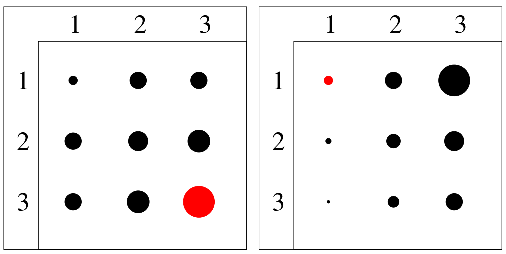
Communication

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

Evolutionary Game Theory



Evolutionary Game Theory



Evolutionary Game Theory (Maynard Smith & Price, 1973)

An Evolutionary Stable Strategy (ESS) is a strategy that cannot be invaded by any other strategy, because all other strategies have either a lower fitness when playing against the ESS, or if their fitness is equal, they have a lower fitness when playing against themselves.

That is, if $W(i,j)$ gives the fitness for a player playing strategy i against an opponent playing strategy j, then i is an ESS iff:

$$\forall j\,(W(i,i) > W(j,i) \vee W(i,i) = W(j,i) > W(j,j))$$

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)

A formalism for communication under noisy conditions

- Assume that there are M different meanings that an individual might want to express, and F different signals (forms) that it can use for this task.
- The communication system of an individual is represented with a production matrix S (S gives for every meaning m and every signal f, the probability that the individual chooses f to convey m);
- and an interpretation matrix R. (R gives for every signal f and meaning m, the probability that f will be interpreted as m).

- Signals can be more or less similar to each other and there is noise on the transmission of signals which depends on these similarities (confusion matrix U).
- Meanings can be more or less similar to each other, and the value of a certain interpretation depends on how close it is to the intention (value matrix V)

Example: alarm calls (vervets, squirls, ...)

Three different types of predators: from the air (eagles), from the ground (leopards) and from the trees (snakes).

The monkeys are capable of making a number (say 5) of different sounds that range on one axis (e.g. pitch, from high to low) and are more easily confused if they are closer together.

If one makes a mistake, typically not every mistake is equally bad.

$S =$

	sent signal				
intention ↓	1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
eagle	1.0	0.0	0.0	0.0	0.0
snake	0.0	0.0	1.0	0.0	0.0
leopard	0.0	0.0	0.0	0.0	1.0

$U =$

	received signal				
sent signal ↓	1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
1*kHz*	0.7	0.2	0.1	0.0	0.0
2*kHz*	0.2	0.6	0.2	0.0	0.0
3*kHz*	0.0	0.2	0.6	0.2	0.0
4*kHz*	0.0	0.0	0.2	0.6	0.2
5*kHz*	0.0	0.0	0.1	0.2	0.7

$R =$

received signal ↓	interpretation *eagle*	*snake*	*leopard*
1*kHz*	1.0	0.0	0.0
2*kHz*	1.0	0.0	0.0
3*kHz*	0.0	1.0	0.0
4*kHz*	0.0	0.0	1.0
5*kHz*	0.0	0.0	1.0

$V =$

interpretations ↓	intentions *eagle*	*snake*	*leopard*
eagle	9	5	1
snake	2	9	2
leopard	1	5	9

$S =$

intention ↓	sent signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
eagle	1.0	0.0	0.0	0.0	0.0
snake	0.0	0.0	1.0	0.0	0.0
leopard	0.0	0.0	0.0	0.0	1.0

$U =$

sent signal ↓	received signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
1*kHz*	0.7	0.2	0.1	0.0	0.0
2*kHz*	0.2	0.6	0.2	0.0	0.0
3*kHz*	0.0	0.2	0.6	0.2	0.0
4*kHz*	0.0	0.0	0.2	0.6	0.2
5*kHz*	0.0	0.0	0.1	0.2	0.7

$R =$

received signal ↓	interpretation *eagle*	*snake*	*leopard*
1*kHz*	1.0	0.0	0.0
2*kHz*	1.0	0.0	0.0
3*kHz*	0.0	1.0	0.0
4*kHz*	0.0	0.0	1.0
5*kHz*	0.0	0.0	1.0

$V =$

interpretations ↓	intentions *eagle*	*snake*	*leopard*
eagle	9	5	1
snake	2	9	2
leopard	1	5	9

Fitness (Utility, Payoff)

The fitness with a given S and R is:

$$W(S,R) = \sum_m \sum_f \sum_{f'} \sum_{m'} S_{mf} U_{ff'} R_{f'm} V_{mm'}$$

A language L is

$$L = \{S, R\}$$

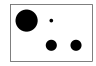
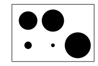
In a symmetric, cooperative game, the fitness for player 1 with L communicating with player 2 with L' is:

$$w(L, L') = \frac{1}{2}\left(W(S, R') + W(S', R)\right)$$

Visualising S and R

$$S = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.9 & 0.1 & 0.0 \\ m_2 & 0.0 & 0.5 & 0.5 \end{pmatrix} \qquad R = \begin{pmatrix} & m_1 & m_2 \\ f_1 & 0.7 & 0.3 \\ f_2 & 0.9 & 0.1 \\ f_3 & 0.0 & 1.0 \end{pmatrix}$$

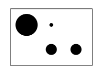
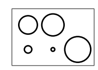
$$R^T = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.7 & 0.9 & 0.0 \\ m_2 & 0.3 & 0.1 & 1.0 \end{pmatrix}$$

Visualising S and R

$$S = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.9 & 0.1 & 0.0 \\ m_2 & 0.0 & 0.5 & 0.5 \end{pmatrix} \qquad R = \begin{pmatrix} & m_1 & m_2 \\ f_1 & 0.7 & 0.3 \\ f_2 & 0.9 & 0.1 \\ f_3 & 0.0 & 1.0 \end{pmatrix}$$

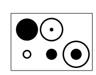
$$R^T = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.7 & 0.9 & 0.0 \\ m_2 & 0.3 & 0.1 & 1.0 \end{pmatrix}$$

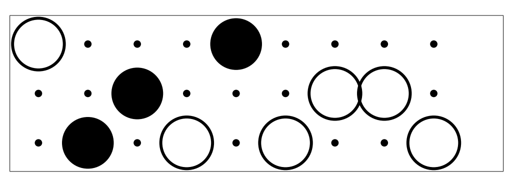



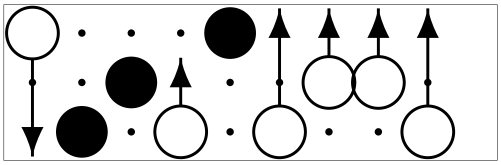
Visualising S and R

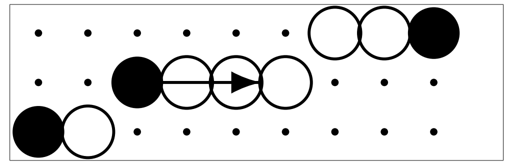
$$S = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.9 & 0.1 & 0.0 \\ m_2 & 0.0 & 0.5 & 0.5 \end{pmatrix} \qquad R = \begin{pmatrix} & m_1 & m_2 \\ f_1 & 0.7 & 0.3 \\ f_2 & 0.9 & 0.1 \\ f_3 & 0.0 & 1.0 \end{pmatrix}$$

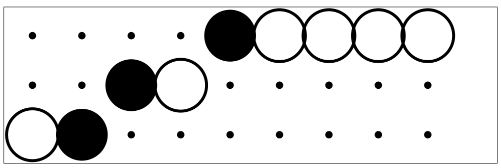
$$R^T = \begin{pmatrix} & f_1 & f_2 & f_3 \\ m_1 & 0.7 & 0.9 & 0.0 \\ m_2 & 0.3 & 0.1 & 1.0 \end{pmatrix}$$

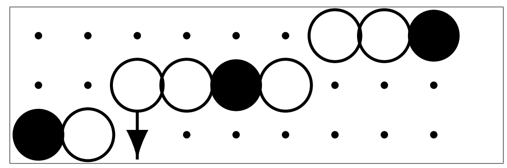


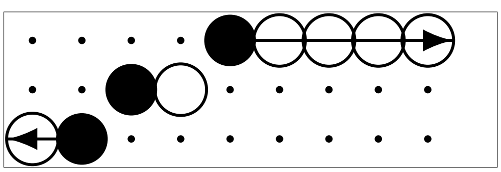


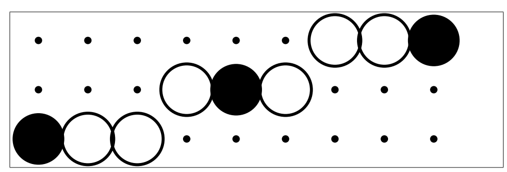


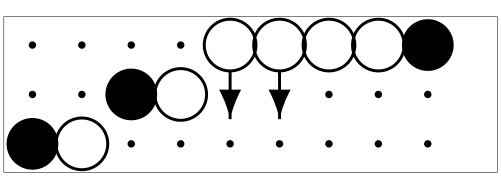




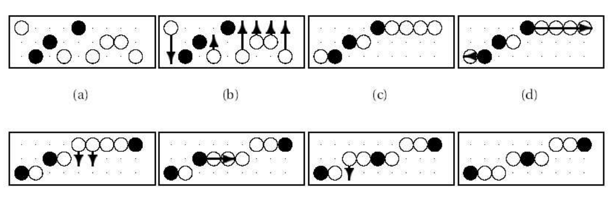








Only distinctive lexicons are ESSs



(a) (b) (c) (d)

(e) (f) (g) (h)

- Pragmatics (e.g. implicatures for quantifiers; Van Rooij, 2004)
  Mary or Lucy arrived late for the party.

- Pragmatics (e.g. implicatures for quantifiers; Van Rooij, 2004)
  Mary or Lucy arrived late for the party.
- Semantics – formal universals (e.g. bidirectionality of the sign - Hurford, 1989; compositionality - Nowak & Krakauer, 1999; Zuidema, 2003)

- Pragmatics (e.g. implicatures for quantifiers; Van Rooij, 2004)
  Mary or Lucy arrived late for the party.
- Semantics – formal universals (e.g. bidirectionality of the sign - Hurford, 1989; compositionality - Nowak & Krakauer, 1999; Zuidema, 2003)
- Semantics – substantive universals (e.g. color terms; Jaeger, 2006)

$$\begin{array}{c} \textit{white} \\ \textit{black} \end{array} < \textit{red} < \begin{array}{c} \textit{green} \\ \textit{yellow} \end{array} < \textit{blue} < \textit{brown} < \begin{array}{c} \textit{purple} \\ \textit{pink} \\ \textit{orange} \\ \textit{grey} \end{array}$$

- Pragmatics (e.g. implicatures for quantifiers; Van Rooij, 2004)
  Mary or Lucy arrived late for the party.
- Semantics – formal universals (e.g. bidirectionality of the sign - Hurford, 1989; compositionality - Nowak & Krakauer, 1999; Zuidema, 2003)
- Semantics – substantive universals (e.g. color terms; Jaeger, 2006)

$$\begin{array}{c} \textit{white} \\ \textit{black} \end{array} < \textit{red} < \begin{array}{c} \textit{green} \\ \textit{yellow} \end{array} < \textit{blue} < \textit{brown} < \begin{array}{c} \textit{purple} \\ \textit{pink} \\ \textit{orange} \\ \textit{grey} \end{array}$$

- Phonology – formal universals (e.g. combinatorial phonology; Zuidema & de Boer, 2005)
- Phonology – substantive universals (e.g. vowel systems – de Boer, 2000)
- Morphosyntax – formal universals (e.g. Universal Grammar/learnability; Nowak et al, 2001)
- Morphosyntax – substantive universals (e.g. case marking – Jaeger, 2003)

- `http://staff.science.uva.nl/∼jzuidema`
  → research (phd-thesis & bibliography)
- `jzuidema@science.uva.nl`
- `http://staff.science.uva.nl/∼vanrooy/`
- *Evolutionary Game Theory & Language Evolution* (MOLEGT6), 6 credits, semester II(1&2), Jelle Zuidema & Robert v. Rooij