Language Evolution in the Lab:
from computational models to human experiments in Evolutionary Linguistics

Simon Kirby

Language Evolution & Computation Research Unit

Linguistics & English Language, PPLS

University of Edinburgh

www.ling.ed.ac.uk/~simon

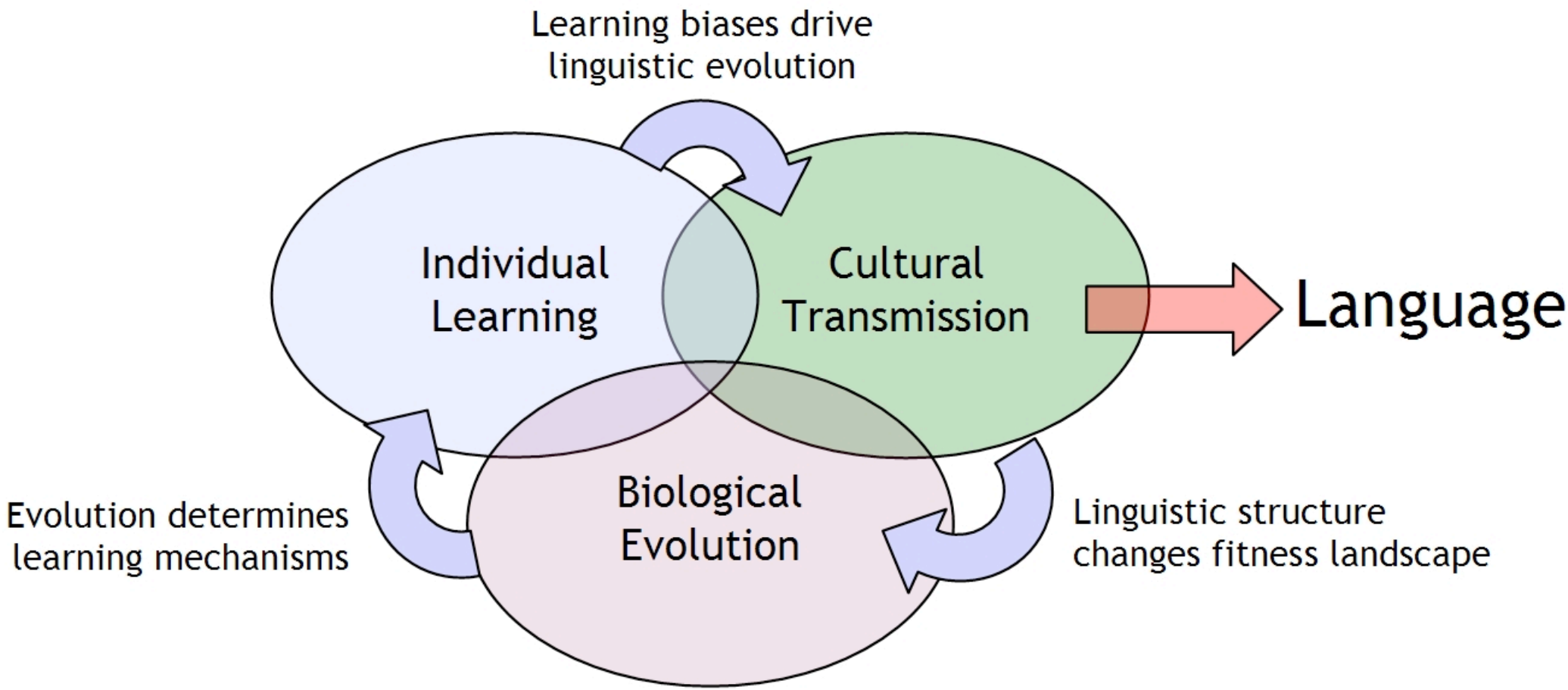
The evolutionary approach to language

The evolutionary approach to language

- Language is (almost) unique in nature

The evolutionary approach to language

- Language is (almost) unique in nature
- It involves three complex adaptive systems:
  - Biological evolution
  - Individual learning
  - Cultural transmission
- Language arises from the interaction of these three


Learning biases drive
linguistic evolution
Individual
Learning
Cultural
Transmission
Language
Biological
Evolution
Evolution determines
learning mechanisms
Linguistic structure
changes fitness landscape

Orthodox view
(e.g. Pinker & Bloom 1990)

- Domain-specific strongly-constraining innate knowledge

Orthodox view
(e.g. Pinker & Bloom 1990)

- Domain-specific strongly-constraining innate knowledge
- Two motivations:
  - **Poverty of the stimulus** - too hard to learn language without specific knowledge
  - **Complex design** - natural selection is the only possible mechanism

Orthodox view
(e.g. Pinker & Bloom 1990)

- Domain-specific strongly-constraining innate knowledge
- Two motivations:
  - **Poverty of the stimulus** - too hard to learn language without specific knowledge
  - **Complex design** - natural selection is the only possible mechanism
- These two are cast into doubt if we take cultural transmission seriously

From individuals to populations

From individuals to populations

- Nativist perspective associated with idealisation of single speaker/hearer in homogenous community
  - Model learning as acquiring target from a fixed set of data

From individuals to populations

- Nativist perspective associated with idealisation of single speaker/hearer in homogenous community
  - Model learning as acquiring target from a fixed set of data
- An alternative: *iterated learning*
  - Behaviour is shaped by observation/interaction with others whose behaviour is shaped in the same way

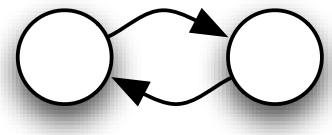
Iterated learning

Iterated learning

- One of the mechanisms of cultural transmission
- Implicated in language at every level:

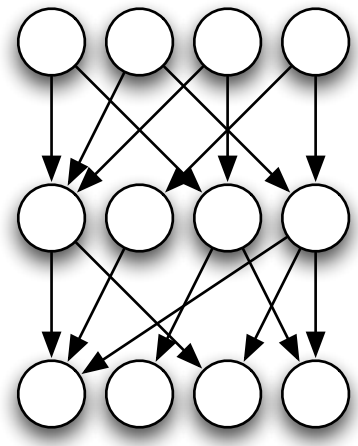
Iterated learning

- One of the mechanisms of cultural transmission
- Implicated in language at every level:
  - Two individuals who align through mutual priming



Iterated learning

- One of the mechanisms of cultural transmission
- Implicated in language at every level:
  - Two individuals who align through mutual priming
  - Whole populations in which language is transmitted over generations



Modelling this process

Modelling this process

- Computational, mathematical and robotic models of iterated learning over the years

Modelling this process

- Computational, mathematical and robotic models of iterated learning over the years
- Some key names:
  - *Batali, Steels, Brighton, Smith, Oudeyer, Zuidema, Baronchelli, Griffiths* (many more)...







Modelling this process

- Computational, mathematical and robotic models of iterated learning over the years
- Some key names:
  - *Batali, Steels, Brighton, Smith, Oudeyer, Zuidema, Baronchelli, Griffiths* (many more)...
- Consistent results: adaptive structure emerges
  - For example: compositional structure emerges from "holistic" protolanguages purely through repeated transmission







Sources of design

Sources of design

- How do we explain apparent design in nature?
  - Biological evolution by natural selection
  - Intelligent design by humans

Sources of design

- How do we explain apparent design in nature?
  - Biological evolution by natural selection
  - Intelligent design by humans
- Modelling results suggest a third source:
  - Cultural transmission leads to adaptive structure in language
- This is an "invisible hand" process (Keller 1994)

Design without a designer:
a cultural evolution experiment

- An experiment to demonstrate this kind of adaptation
  - Based on previous computational models of emergence of compositionality

Hannah Cornish

Kenny Smith

Kirby, Cornish & Smith (2008) *PNAS*

Design without a designer:
a cultural evolution experiment

- An experiment to demonstrate this kind of adaptation
  - Based on previous computational models of emergence of compositionality
- Try to rule out other sources of design:
  - Our participants do not evolve (!)
  - Set up to rule out intentional participant design (cf. experiments by Galantucci, Garrod, Healey, Theisen, Scott-Phillips...)

Hannah Cornish

Kenny Smith

Kirby, Cornish & Smith (2008) *PNAS*

Design without a designer:
a cultural evolution experiment

- Combine *diffusion chain* and *artificial language learning* studies

Design without a designer:
a cultural evolution experiment

- Combine *diffusion chain* and *artificial language learning* studies
- Cultural transmission of an "alien language"

Design without a designer:
a cultural evolution experiment

- Combine *diffusion chain* and *artificial language learning* studies
- Cultural transmission of an "alien language"

1. Start off with a random artificial language

Design without a designer:
a cultural evolution experiment

- Combine *diffusion chain* and *artificial language learning* studies
- Cultural transmission of an "alien language"
  1. Start off with a random artificial language
  2. Ask an experimental subject to try and learn this language and test them

Design without a designer: a cultural evolution experiment

- Combine *diffusion chain* and *artificial language learning* studies
- Cultural transmission of an "alien language"
  1. Start off with a random artificial language
  2. Ask an experimental subject to try and learn this language and test them
  3. Use their output on test as the language to teach the next subject in the experiment (and repeat)

Hypothesis

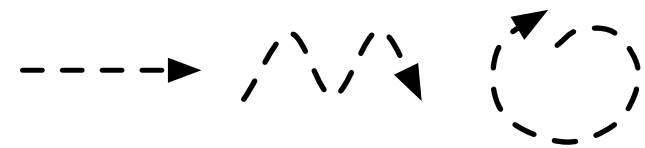
- There will be cumulative cultural adaptation of the language without intentional design by participants

Hypothesis

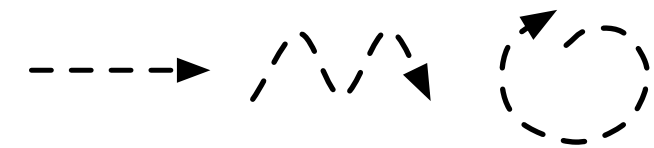
- There will be cumulative cultural adaptation of the language without intentional design by participants
- Two ways of verifying this:
  - The language should become easier to learn
  - The language should become structured

The Language

The Language

- A set of 27 possible "meanings"
  - Pictures with coloured objects in motion:
    - Three shapes 
    - Three colours 
    - Three motions 

The Language

- A set of 27 possible "meanings"
  - Pictures with coloured objects in motion:
    - Three shapes □ ○ △
    - Three colours 
    - Three motions 
- A large set of possible "signals"
  - Random sequences of between two and four syllables chosen from a set of nine
  - No spaces

Procedure

Procedure

- Language divided randomly into two sets:
  - SEEN set: 14 string-picture pairs
  - UNSEEN set: remaining 13 string-picture pairs

Procedure

- Language divided randomly into two sets:
  - SEEN set: 14 string-picture pairs
  - UNSEEN set: remaining 13 string-picture pairs
- Subjects trained on SEEN set

Procedure

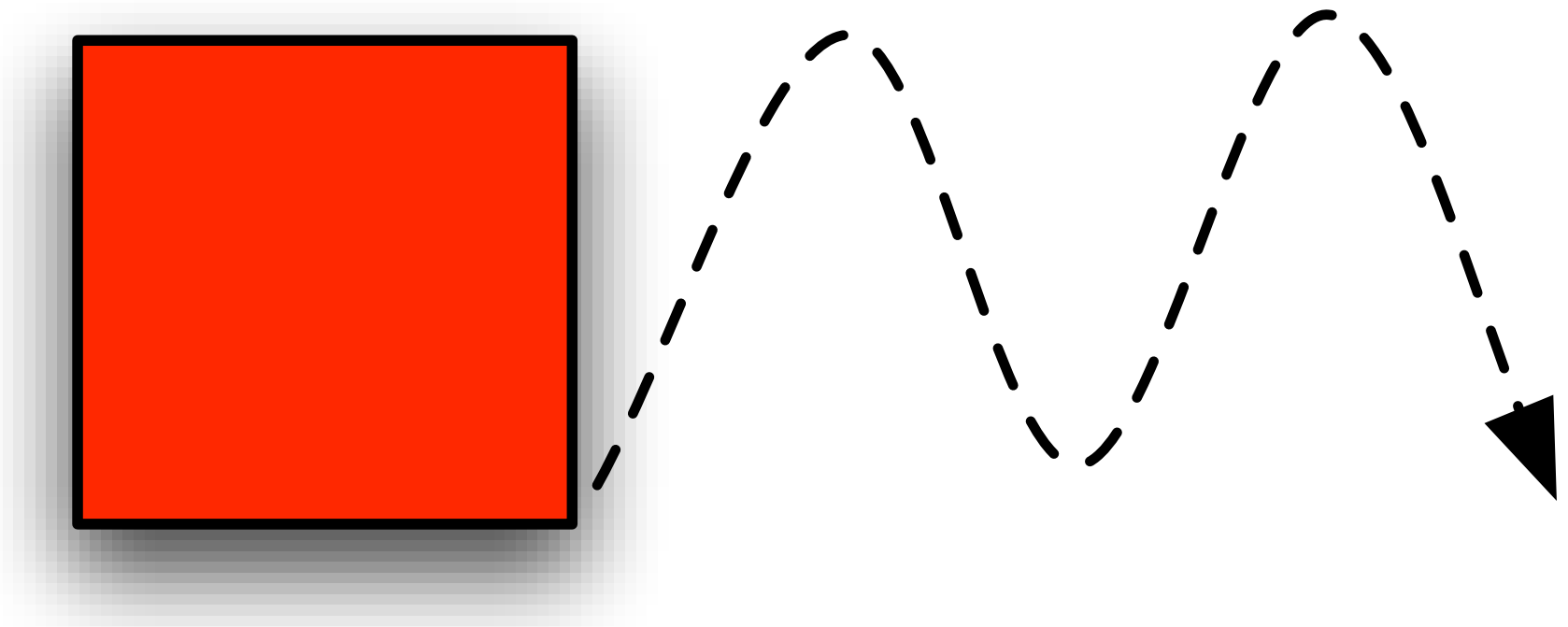
- Language divided randomly into two sets:
  - SEEN set: 14 string-picture pairs
  - UNSEEN set: remaining 13 string-picture pairs
- Subjects trained on SEEN set
- String displayed for 1 second, then string and picture for a further 5 seconds

Procedure

- Language divided randomly into two sets:
  - SEEN set: 14 string-picture pairs
  - UNSEEN set: remaining 13 string-picture pairs
- Subjects trained on SEEN set
- String displayed for 1 second, then string and picture for a further 5 seconds
- Tested on complete set,
  - randomly redivided into new SEEN and UNSEEN sets for next generation

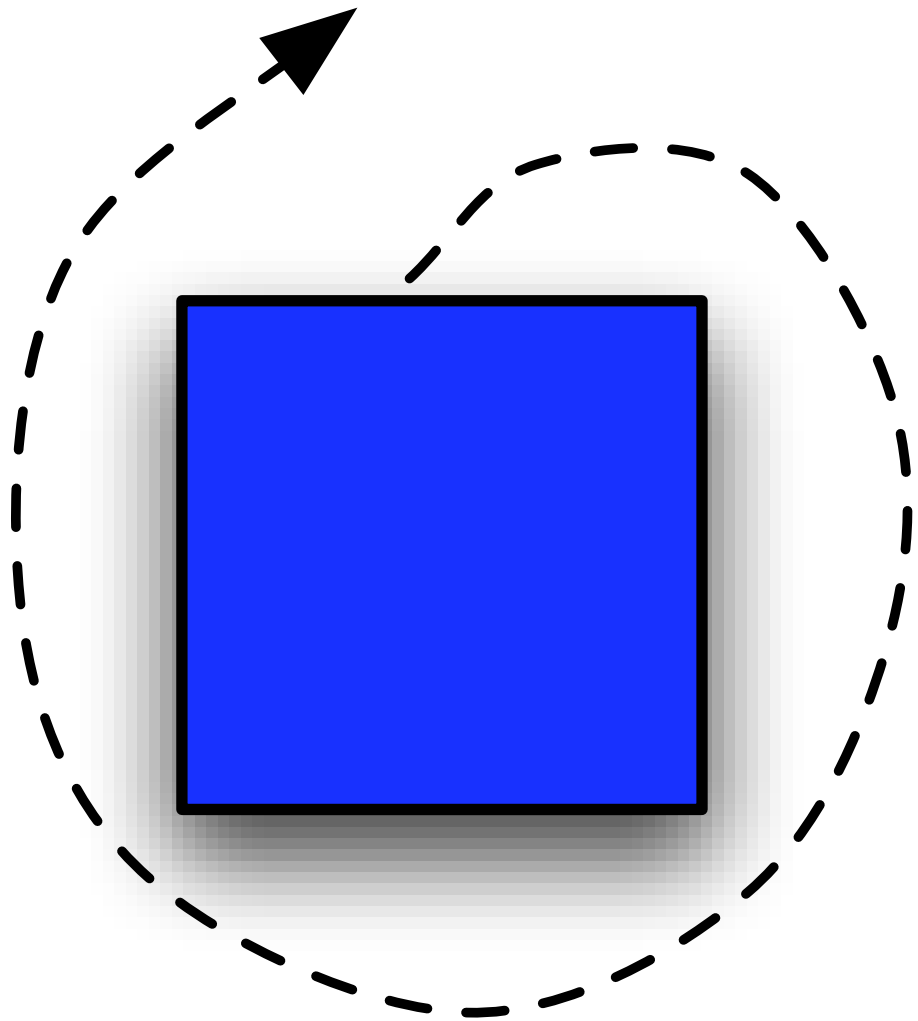
kihemiwi

kihemiwi

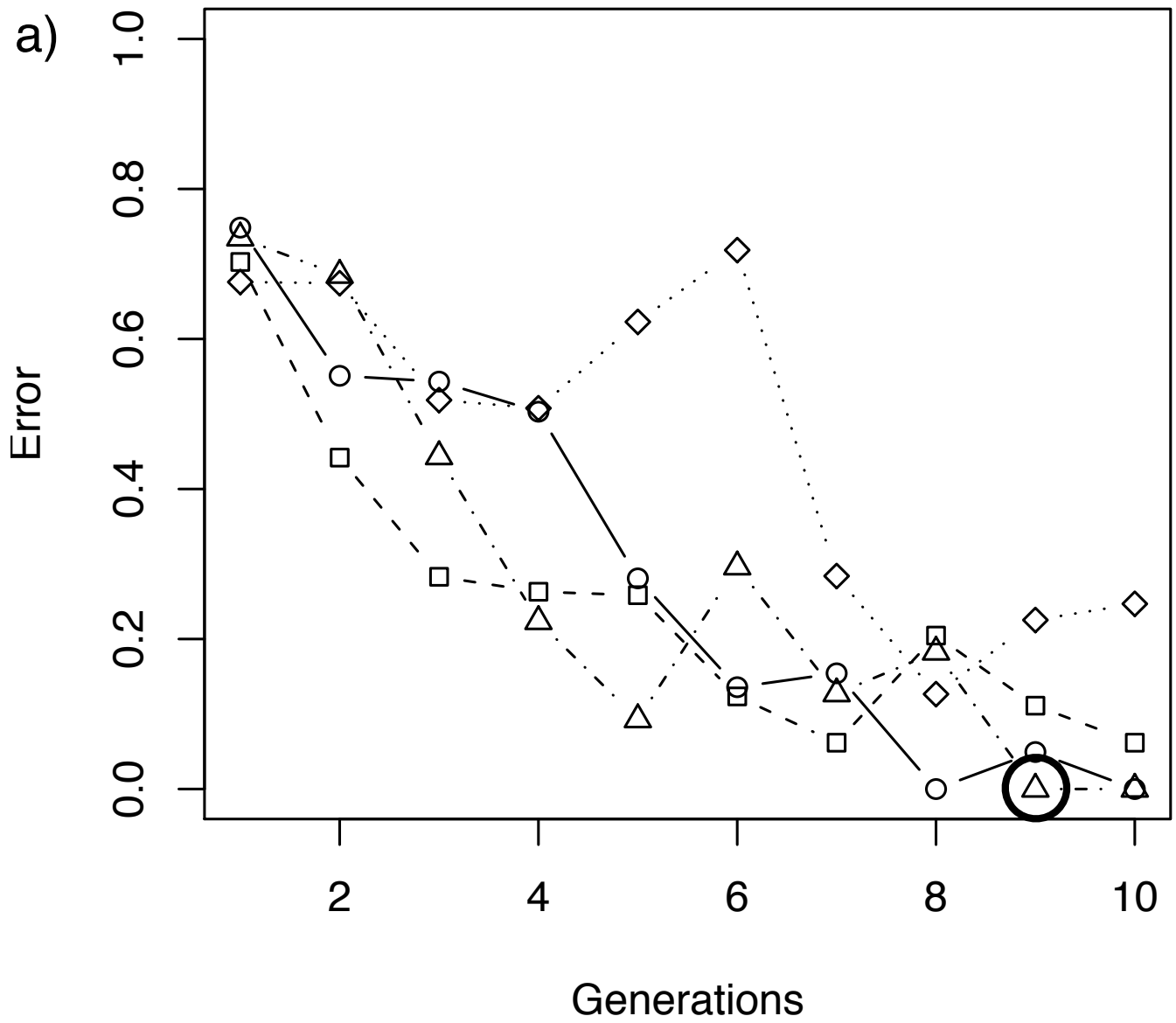


kunige

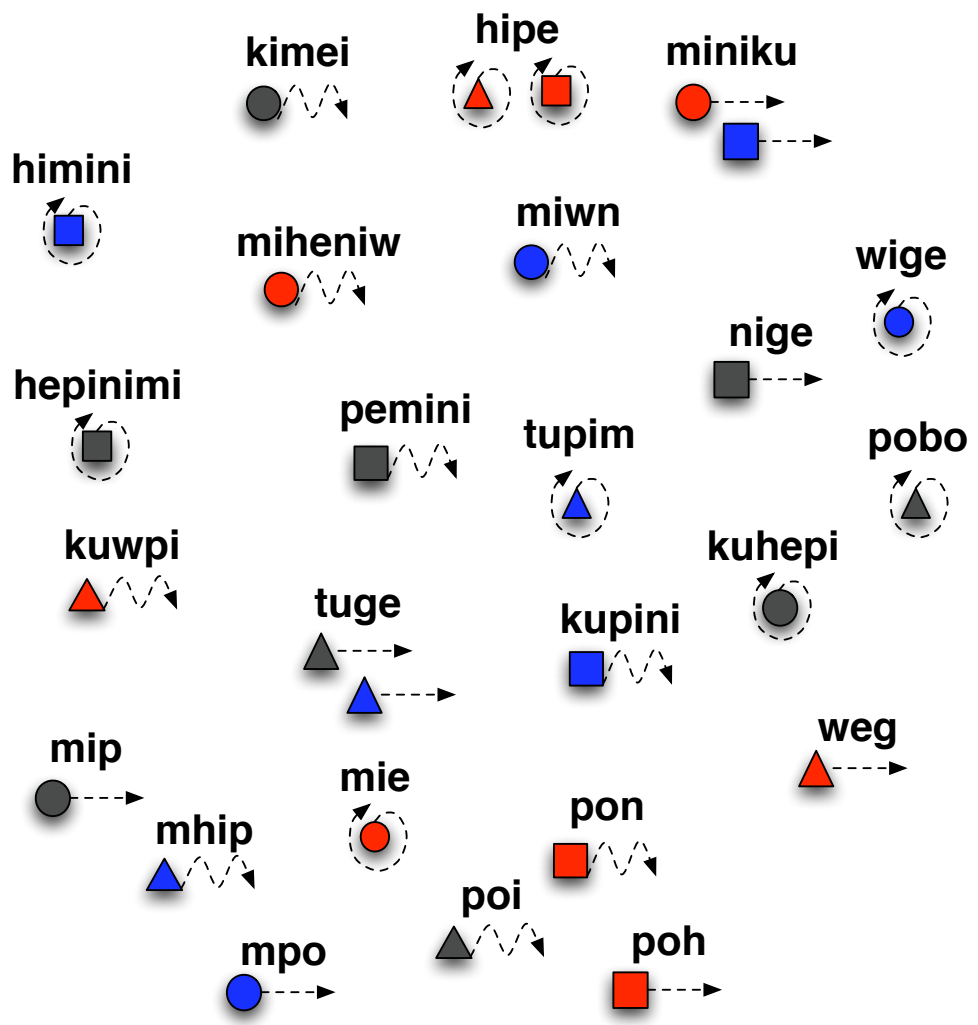
kunige



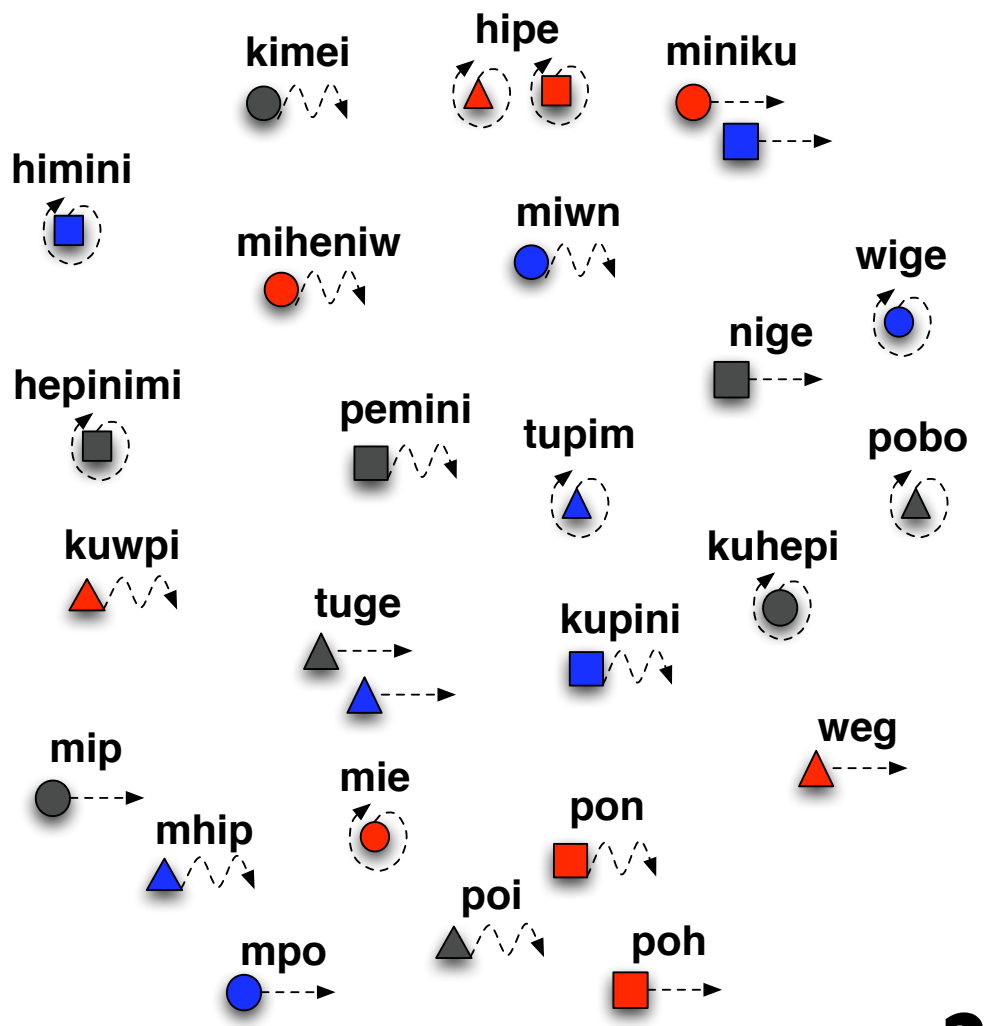
Language becomes easier to learn



After Generation 1:

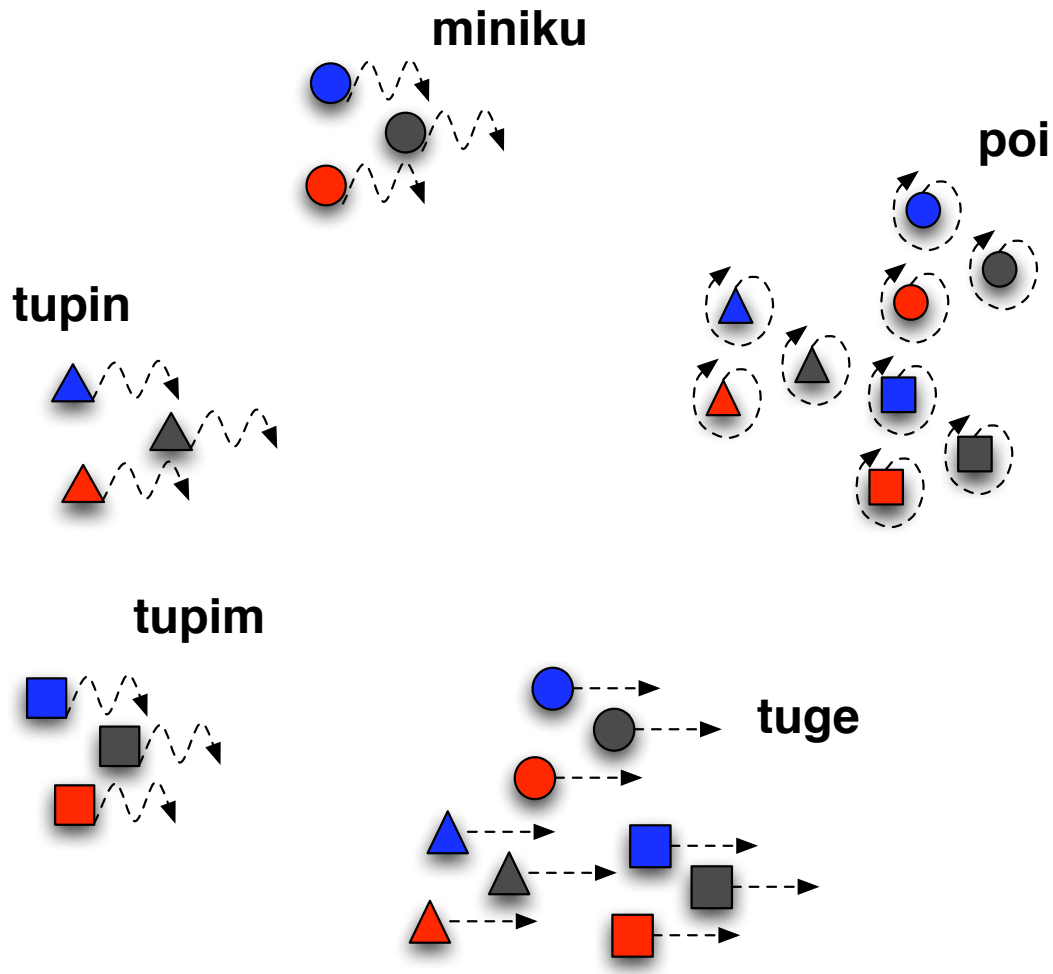


After Generation 1:

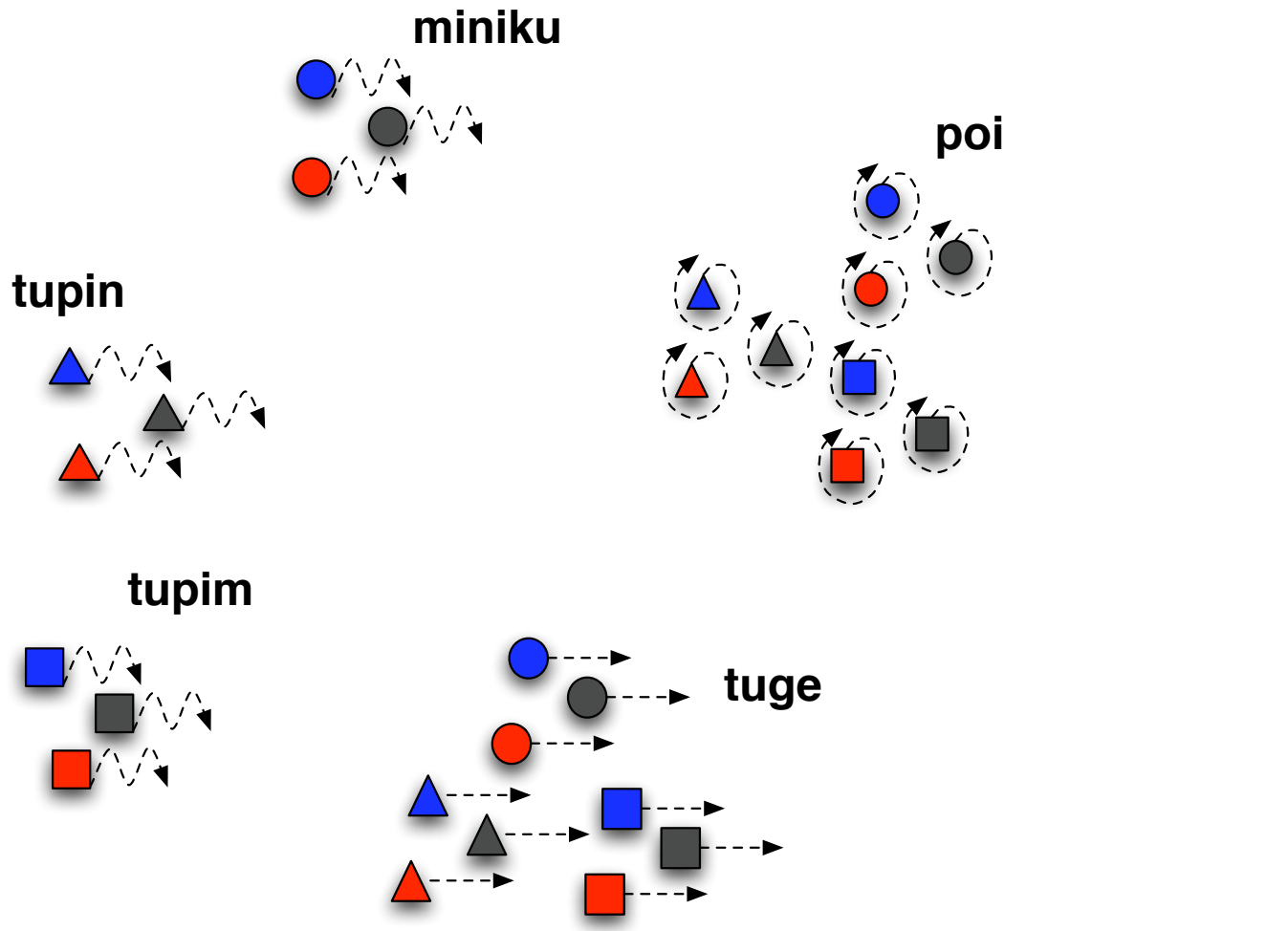


24 words

After Generation 10:



After Generation 10:



5 words

How has language become easier?

How has language become easier?

- Looks like it might be just that there are fewer words.

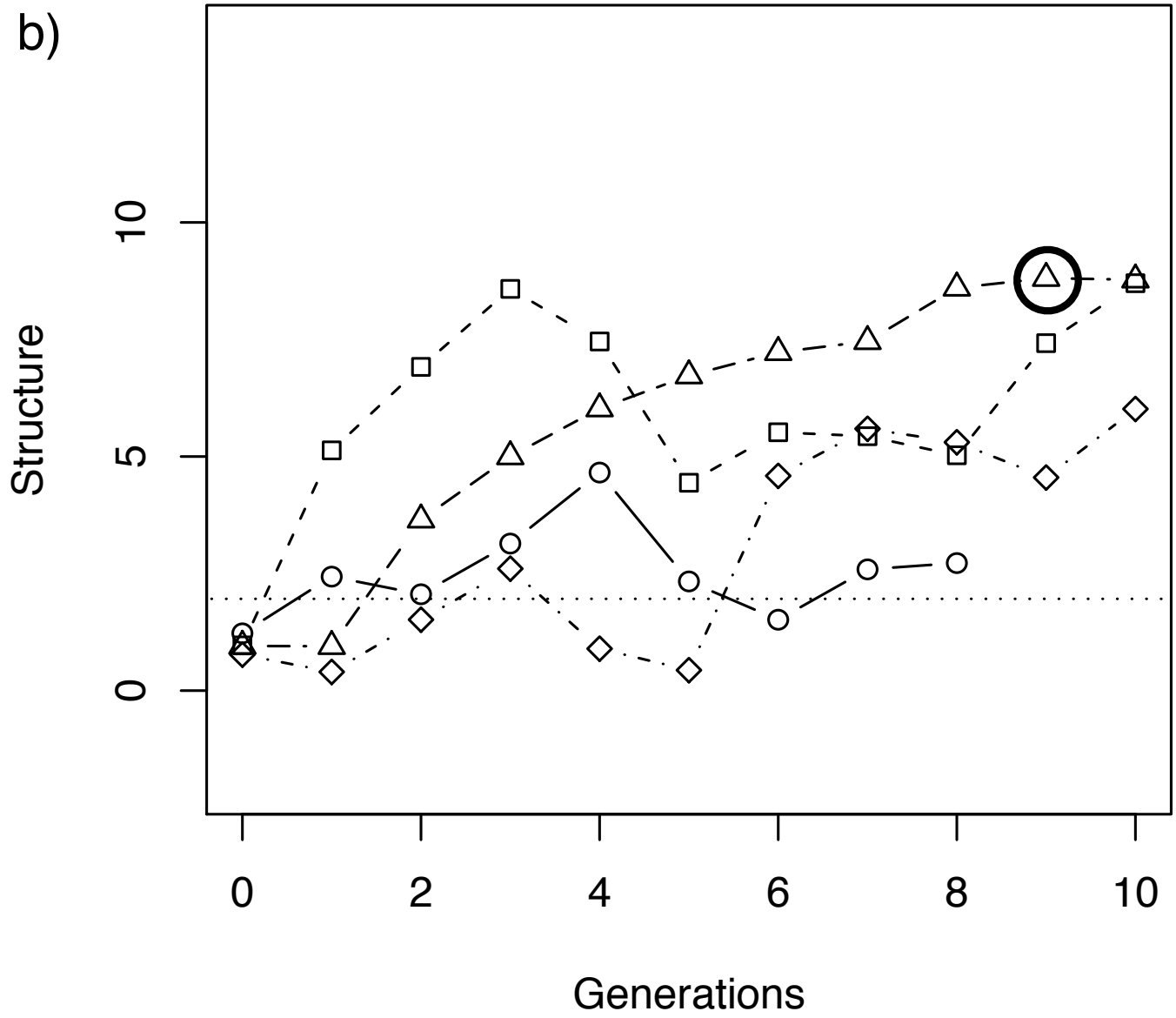
How has language become easier?

- Looks like it might be just that there are fewer words.
- If this were all that was going on, then subjects' performance on unseen items should be random

How has language become easier?

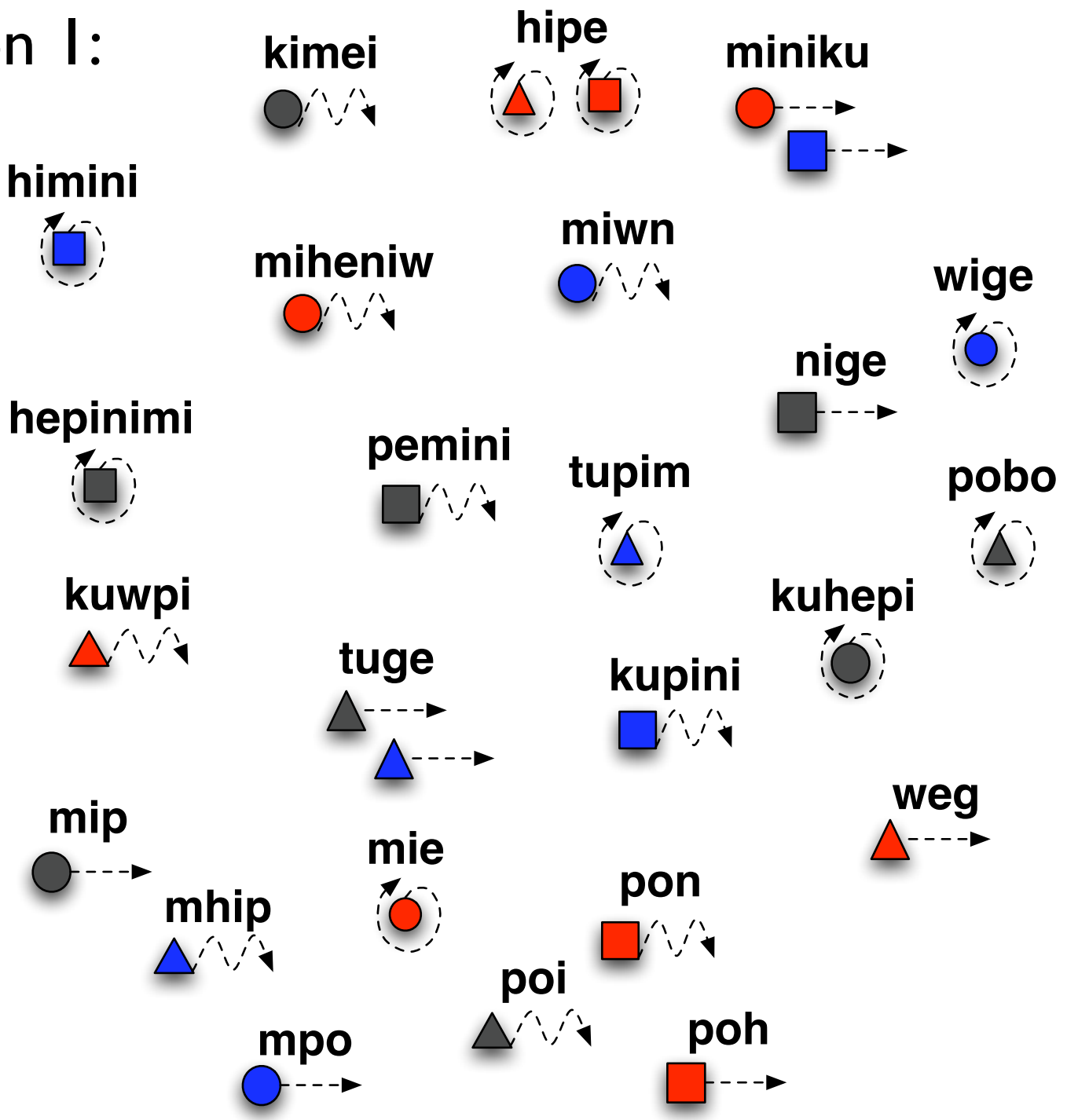
- Looks like it might be just that there are fewer words.
- If this were all that was going on, then subjects' performance on unseen items should be random
- This doesn't appear to be the case...

Language becomes systematic

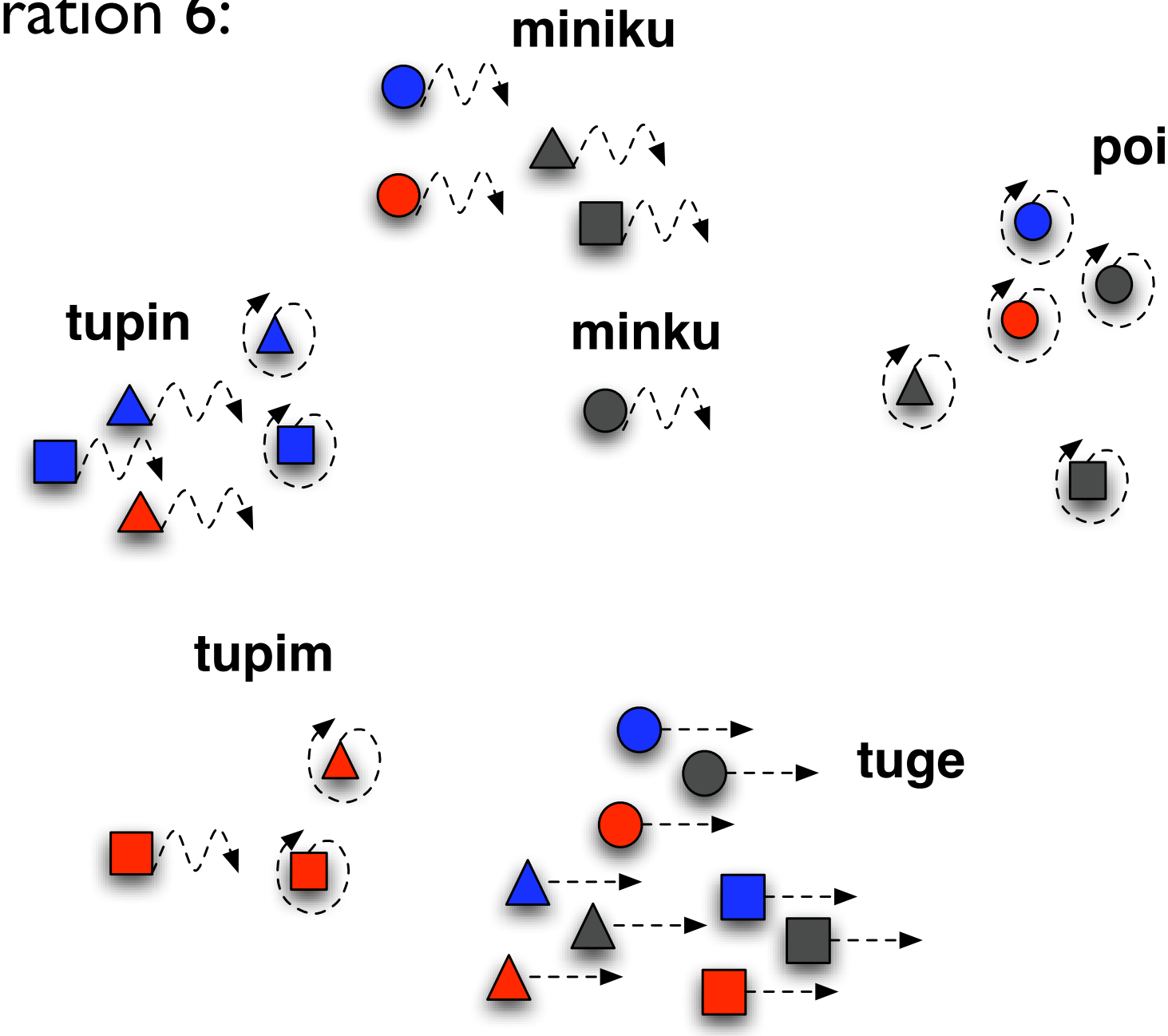


Measure shows the degree to which meaning and form are systematically related.

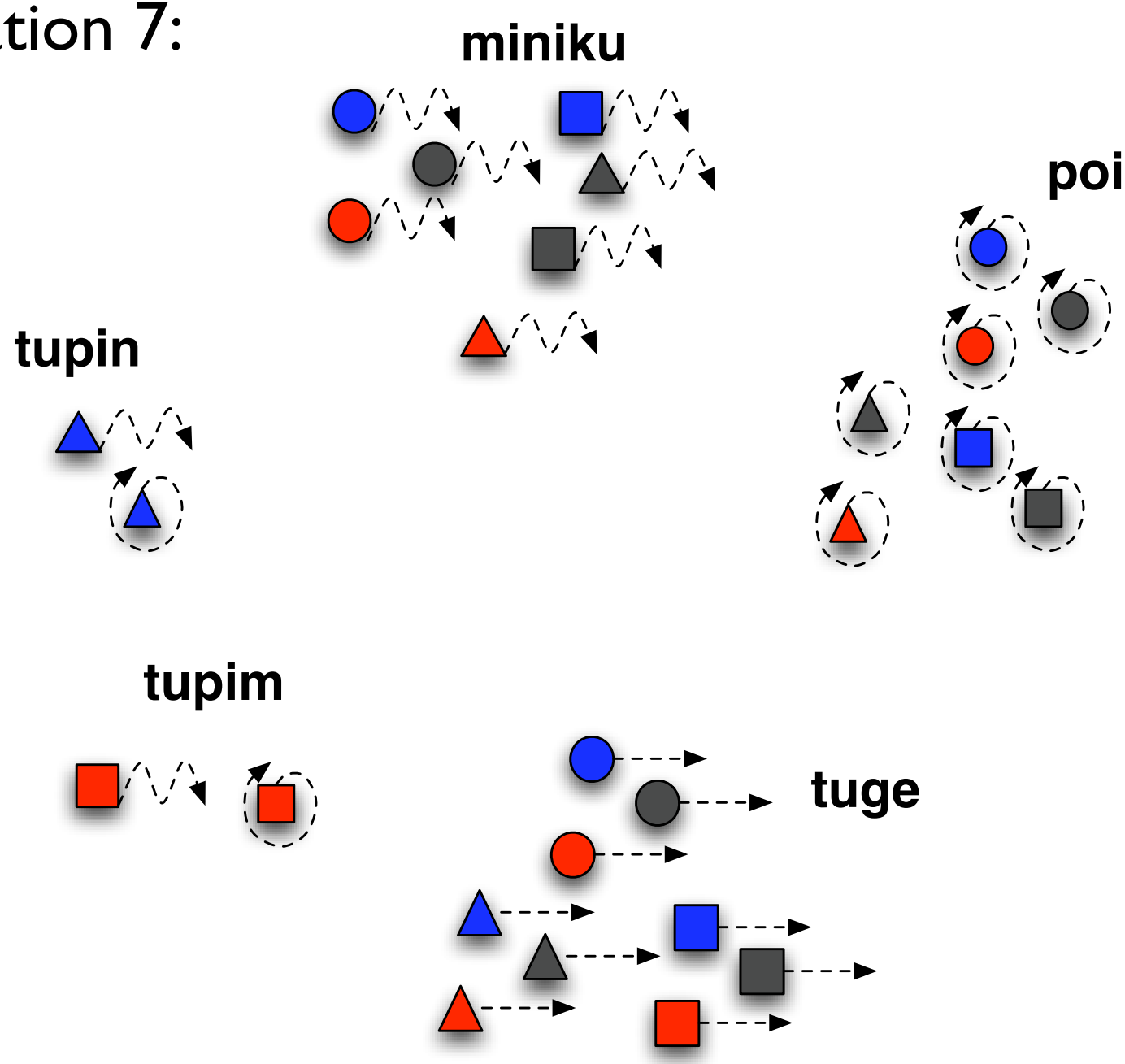
After Generation 1:



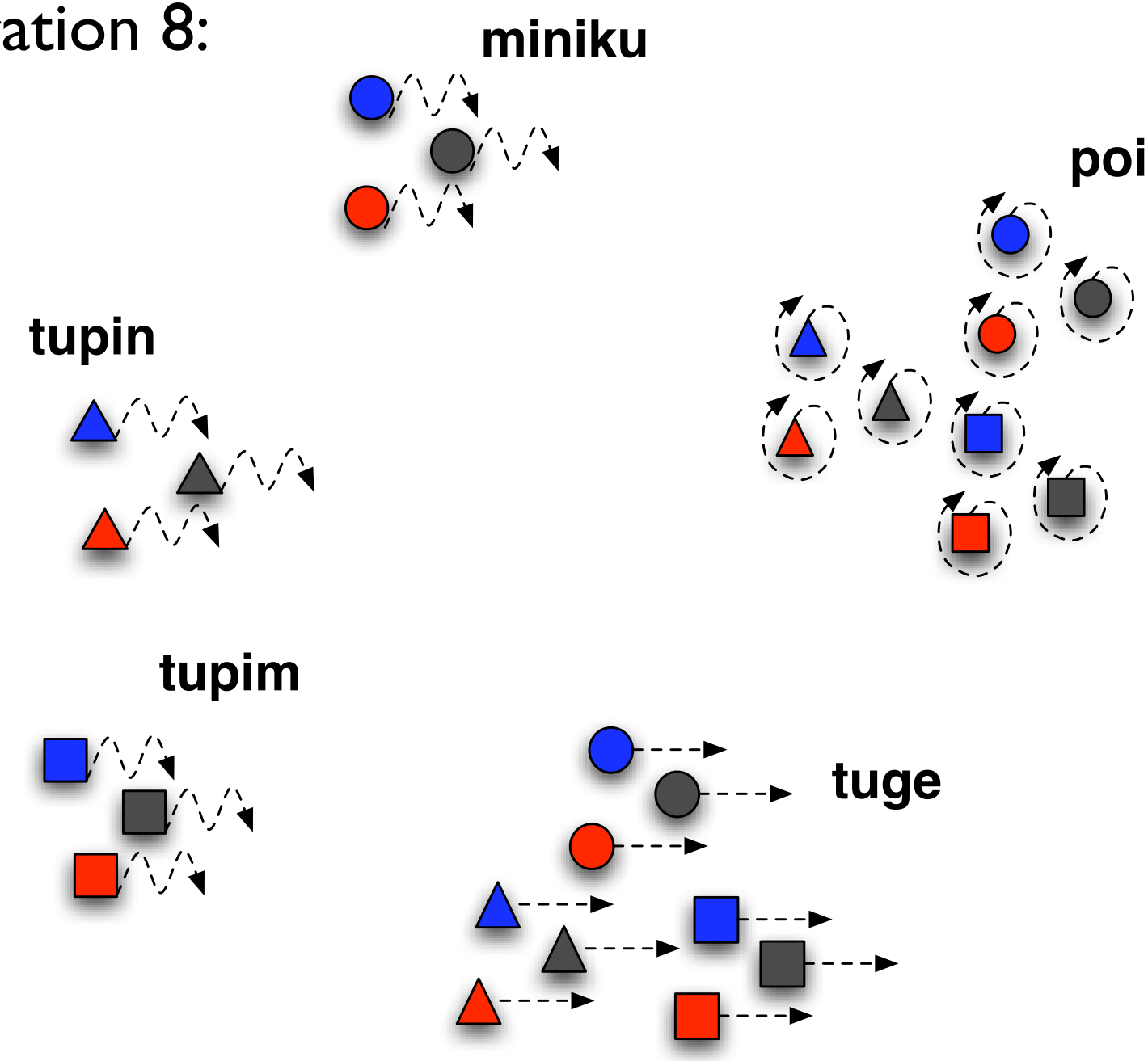
After Generation 6:



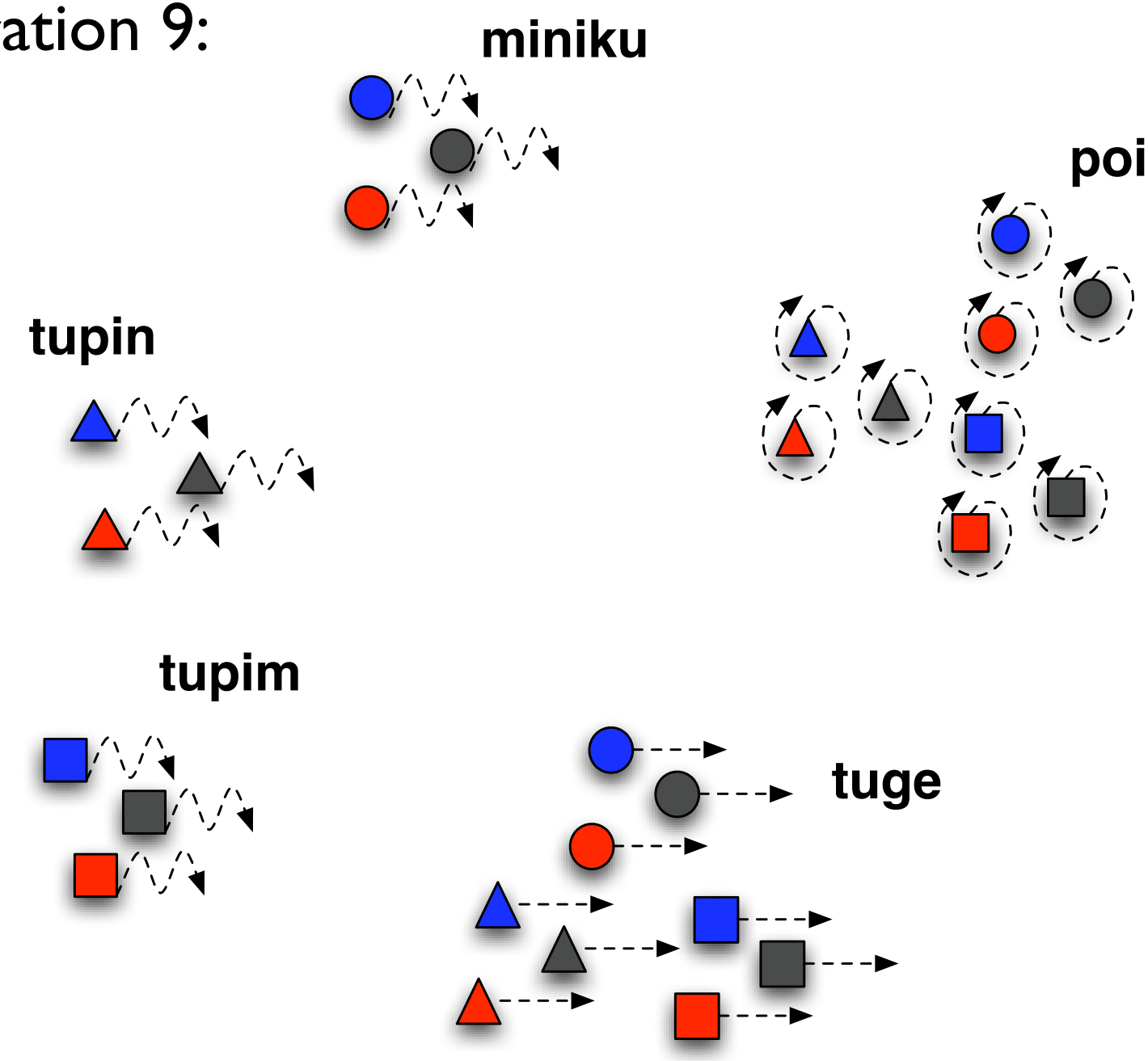
After Generation 7:



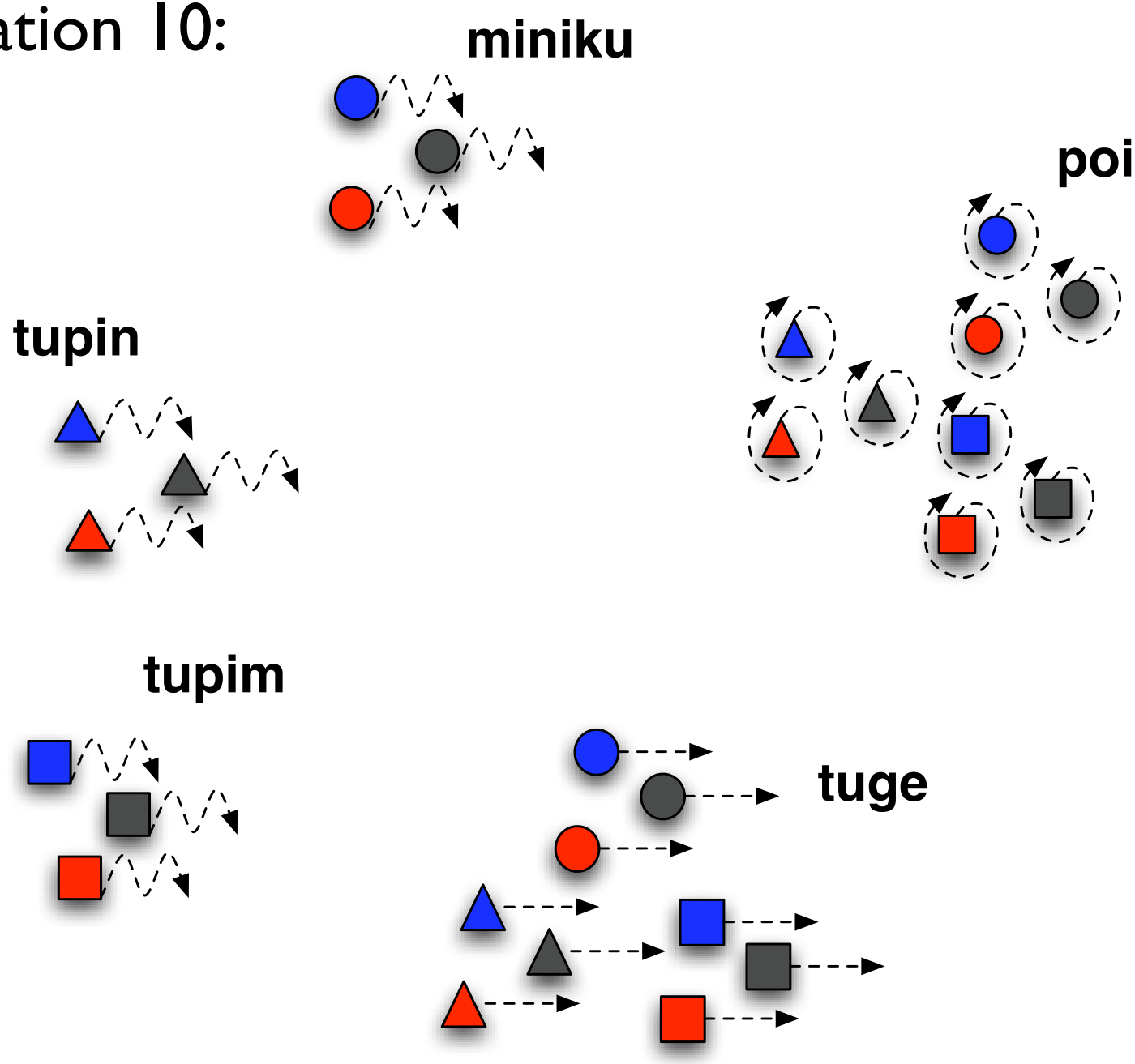
After Generation 8:



After Generation 9:



After Generation 10:



Language adapts to be structured

- Language adapts
  - **Subjects are not aware of this** (they aren't even aware they are being shown unseen items!)
  - Systematic underspecification is an *adaptation by language to aid its own survival*
- Cumulative cultural adaptation without intention

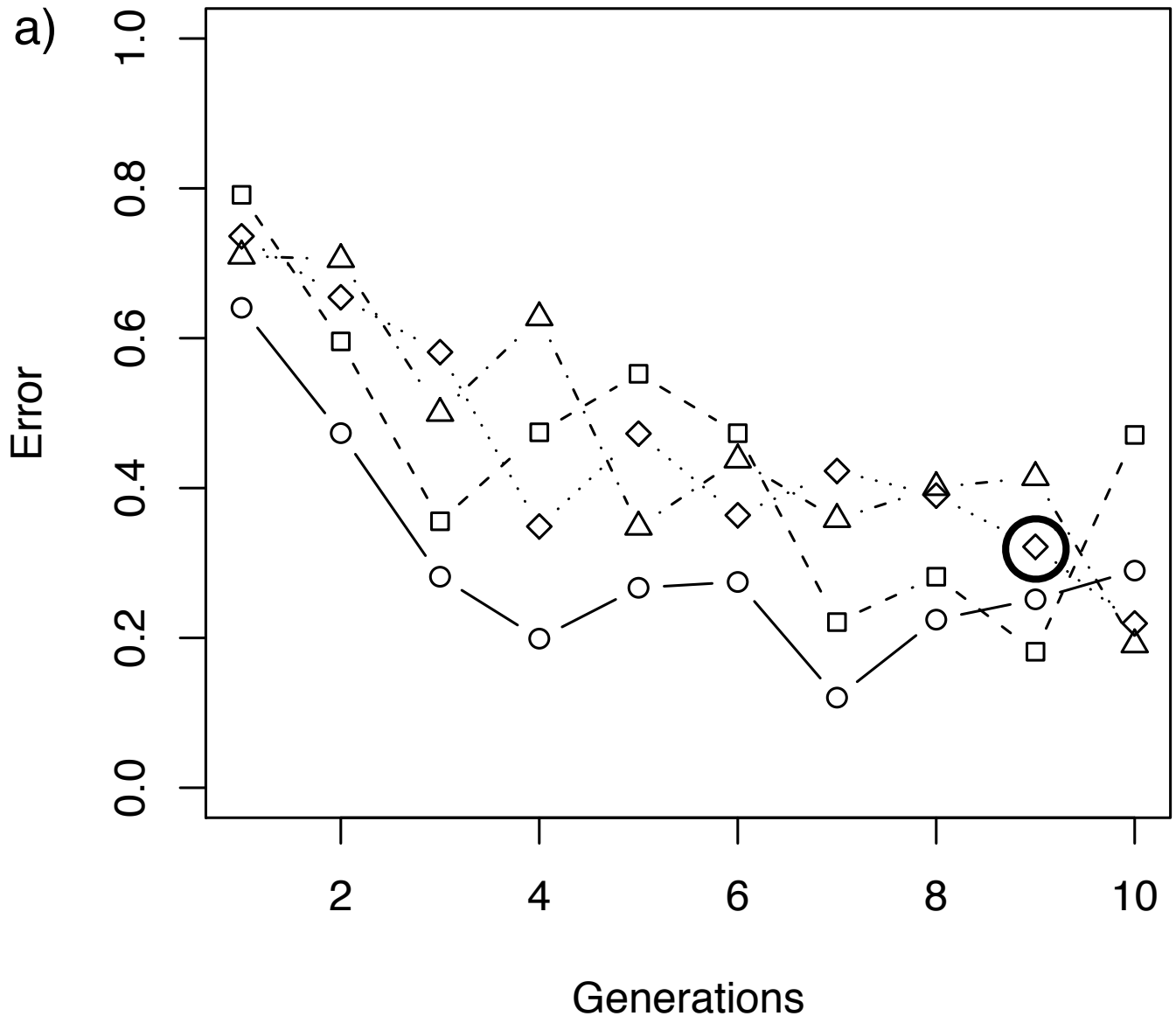
More interesting structure?

- In reality language exhibits compositional structure (e.g. morphology, syntax) that makes it learnable *and* expressive
- There's no pressure for expressivity in the experiment

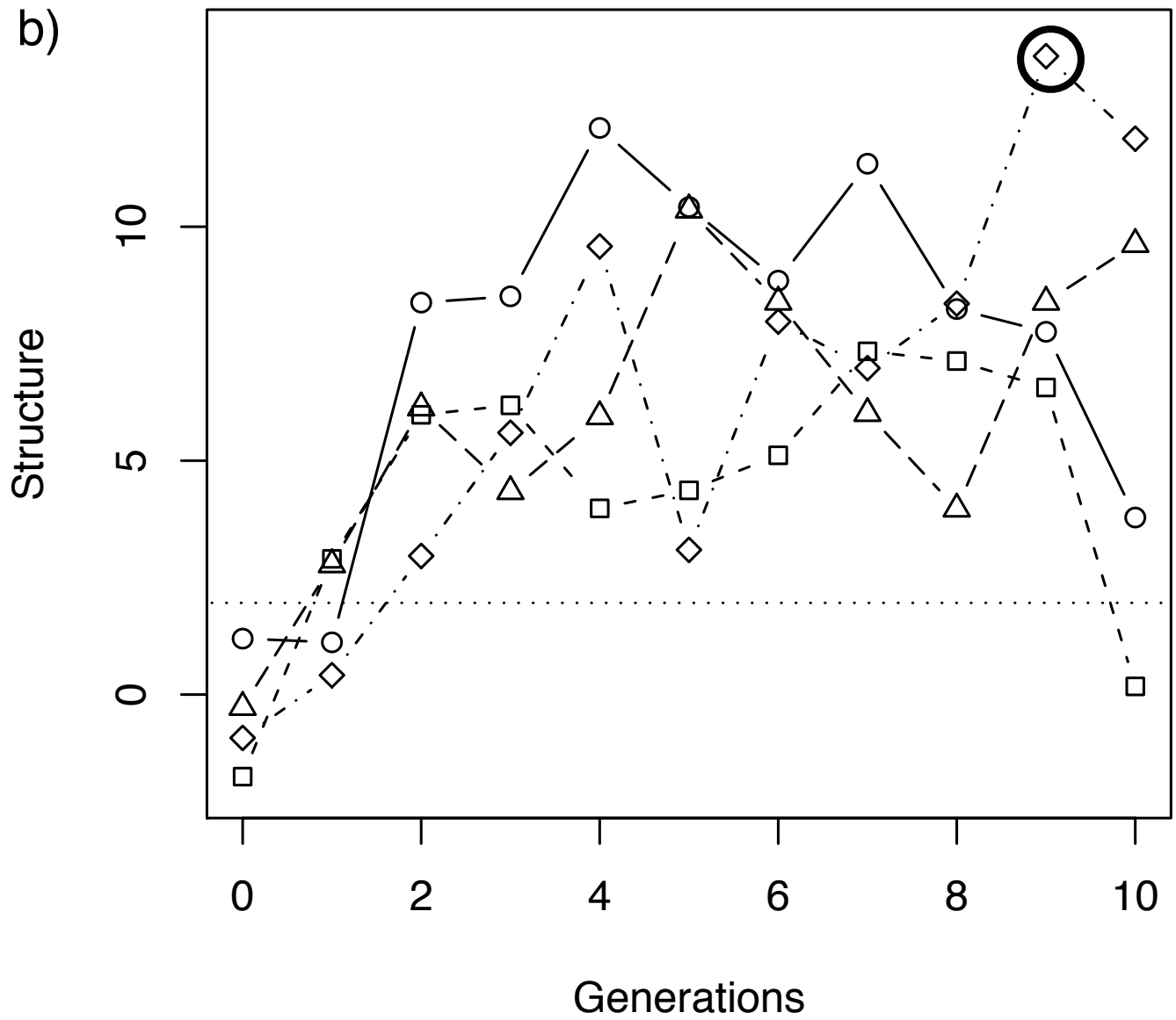
More interesting structure?

- In reality language exhibits compositional structure (e.g. morphology, syntax) that makes it learnable *and* expressive
- There's no pressure for expressivity in the experiment
- Simple modification: filter out all ambiguous items from SEEN set before subjects see them

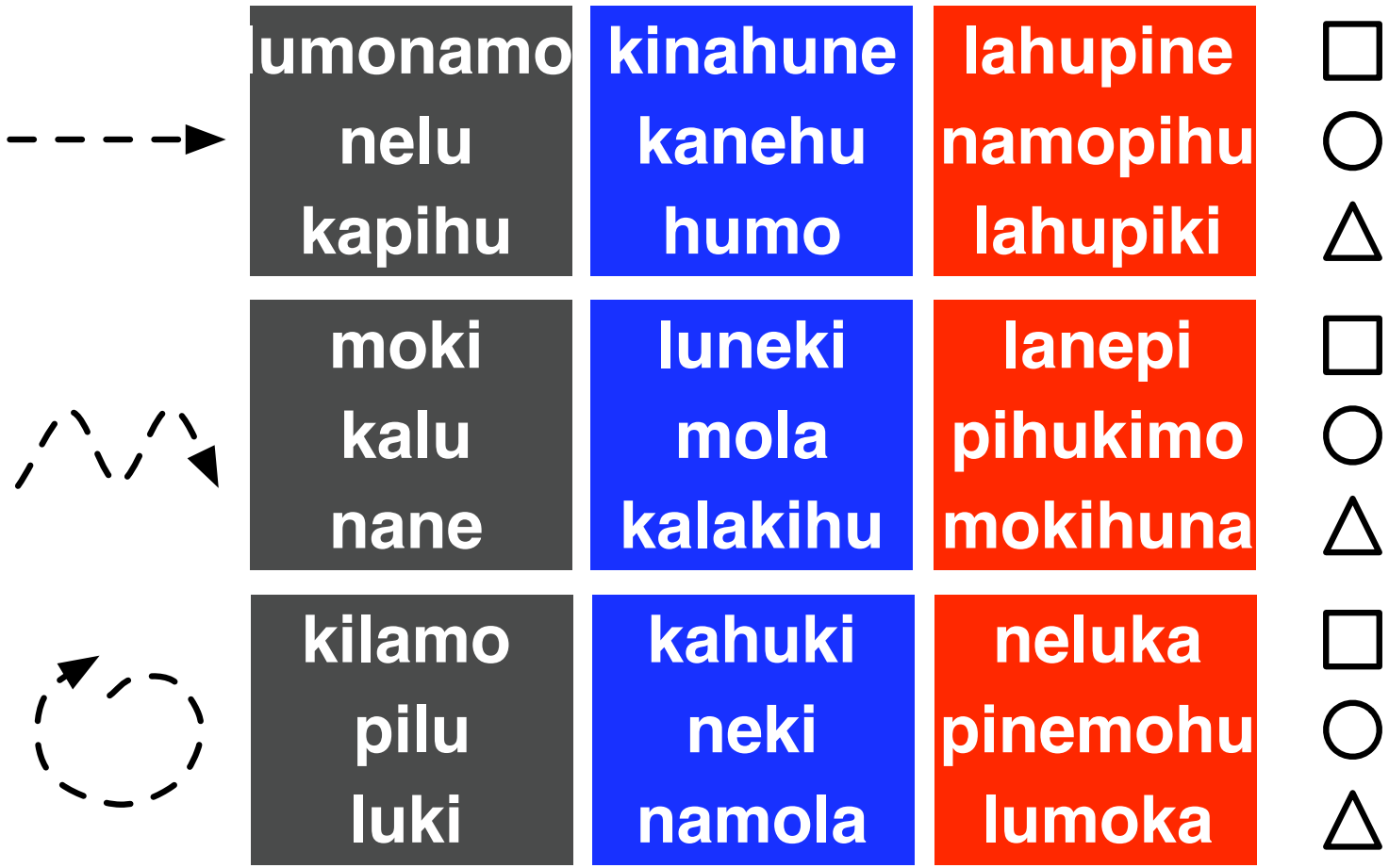
Language becomes easier to learn



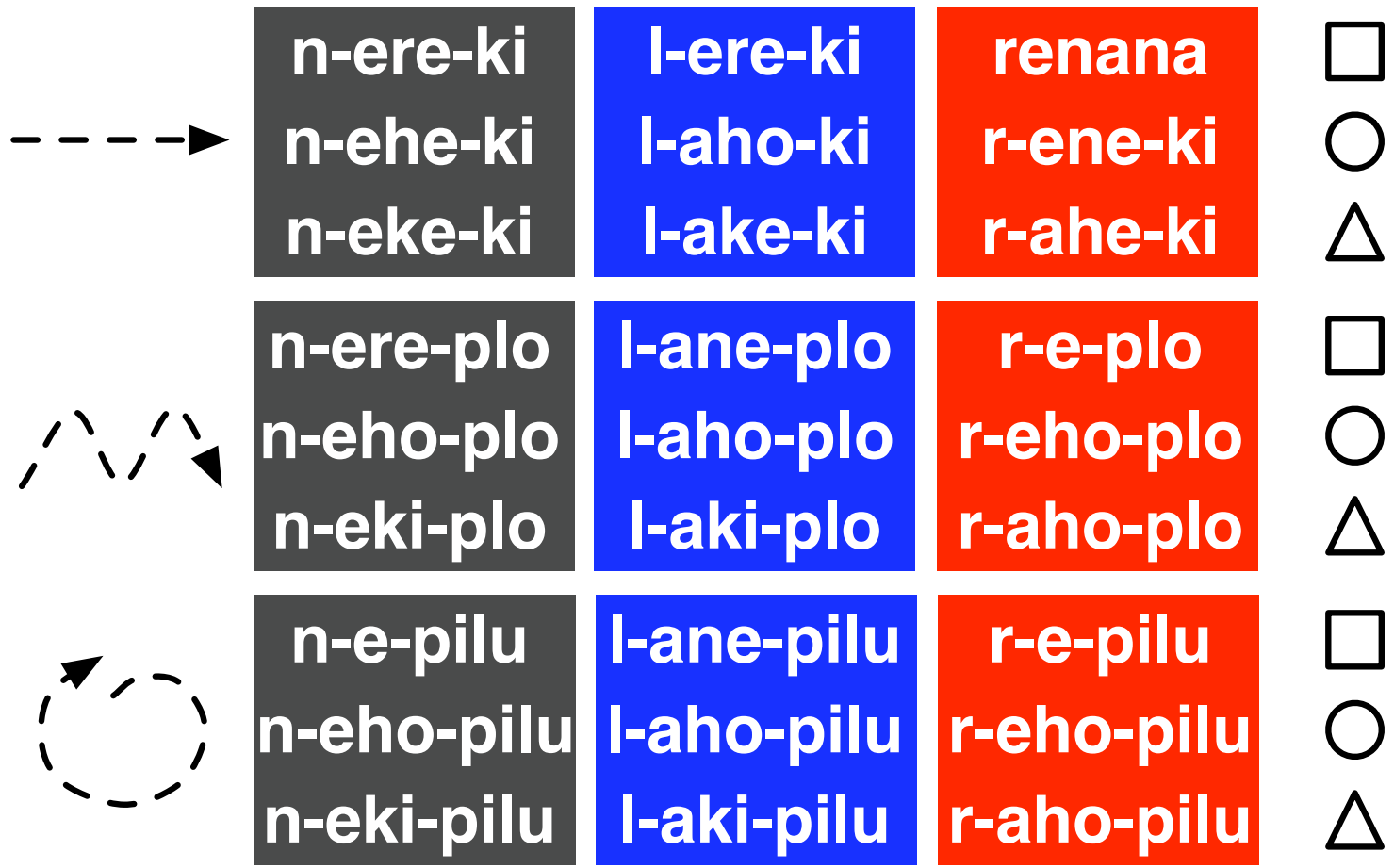
Language becomes systematic



Example initial language



Example final language
(10 "generations" later)



Adaptation again

Adaptation again

- Language adapts to the transmission "bottleneck"

Adaptation again

- Language adapts to the transmission "bottleneck"
- It must be learned even though:
  - only a sub-sample is seen by learners
  - ambiguous signals are filtered out

Adaptation again

- Language adapts to the transmission "bottleneck"
- It must be learned even though:
  - only a sub-sample is seen by learners
  - ambiguous signals are filtered out
- Morphological/syntactic structure is a solution to this problem. Mirrors proposed holistic to compositional protolanguage transition.

Adaptation again

- Language adapts to the transmission "bottleneck"
- It must be learned even though:
  - only a sub-sample is seen by learners
  - ambiguous signals are filtered out
- Morphological/syntactic structure is a solution to this problem. Mirrors proposed holistic to compositional protolanguage transition.
- Note: subjects cannot be aware of the filtering, but language structure is very different
  - Demonstrates that adaptation is *non-intentional*
  - Culture gives us *design without a designer*

Motivations for innateness: complex design

- "Evolutionary theory offers clear criteria for when a trait should be attributed to natural selection: complex design for some function, and *the absence of alternative processes capable of explaining such complexity*. Human language meets these criteria." (Pinker & Bloom 1990)

Motivations for innateness: complex design

- "Evolutionary theory offers clear criteria for when a trait should be attributed to natural selection: complex design for some function, and *the absence of alternative processes capable of explaining such complexity*. Human language meets these criteria." (Pinker & Bloom 1990)
- But iterated learning provides just such a process

Motivations for innateness: poverty of the stimulus

Motivations for innateness: poverty of the stimulus

- Different types of induction (Christiansen & Chater 2008):

Motivations for innateness: poverty of the stimulus

- Different types of induction (Christiansen & Chater 2008):
  - Natural induction: "right" answer defined by outside forces - potentially difficult problem

Motivations for innateness: poverty of the stimulus

- Different types of induction (Christiansen & Chater 2008):
  - Natural induction: "right" answer defined by outside forces - potentially difficult problem
  - Cultural induction: problem determined by the very biases that learners themselves have. Any biases you have are guaranteed to be helpful because target is shaped by similar learners

Motivations for innateness: poverty of the stimulus

- Different types of induction (Christiansen & Chater 2008):
  - Natural induction: "right" answer defined by outside forces - potentially difficult problem
  - Cultural induction: problem determined by the very biases that learners themselves have. Any biases you have are guaranteed to be helpful because target is shaped by similar learners
- The less data learners have, the better the situation gets because data increasingly reflects biases (Kirby et al 2007)
- "The poverty of the stimulus solves the poverty of the stimulus" (Zuidema 2003)

What role for biological evolution?

What role for biological evolution?

- Not likely to be the evolution of strongly-constraining domain-specific structural knowledge (Smith & Kirby 2008)

What role for biological evolution?

- Not likely to be the evolution of strongly-constraining domain-specific structural knowledge (Smith & Kirby 2008)
- Huge role for biological evolution to explain the origins of transmissible signal-meaning mappings:
  - "Vocal" learning of complex sequential signals
  - Inference of complex meanings

What role for biological evolution?

- Not likely to be the evolution of strongly-constraining domain-specific structural knowledge (Smith & Kirby 2008)
- Huge role for biological evolution to explain the origins of transmissible signal-meaning mappings:
  - "Vocal" learning of complex sequential signals
  - Inference of complex meanings
- Ongoing co-evolutionary pressures to maintain and refine suite of cognitive abilities that support cultural transmission (cf. Lachlan's *cultural trap*, and Deacon's *redistributed selection*)

Conclusions

Conclusions

- Language arises from three interacting complex adaptive systems: learning, culture & evolution

Conclusions

- Language arises from three interacting complex adaptive systems: learning, culture & evolution
- The most poorly understood of these is cultural evolution, yet it has much explanatory power

Conclusions

- Language arises from three interacting complex adaptive systems: learning, culture & evolution
- The most poorly understood of these is cultural evolution, yet it has much explanatory power
- To understand this better, we can move from models of learning to models of iterated learning
  - I've given just one example. Much work to be done!

Conclusions

- Language arises from three interacting complex adaptive systems: learning, culture & evolution
- The most poorly understood of these is cultural evolution, yet it has much explanatory power
- To understand this better, we can move from models of learning to models of iterated learning
  - I've given just one example. Much work to be done!
- Ultimately, changes our perspective on the biological evolution of our species capacity for language