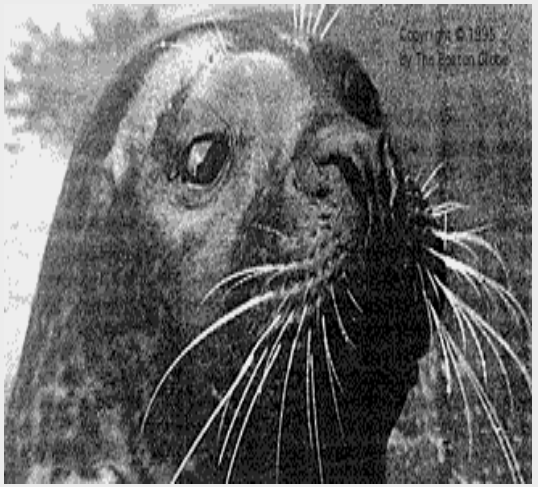


Evolution of Language, 2014
L1: Language in Nature

Jelle Zuidema
Faculty of Humanities / ILLC, UvA

```
(defun all-twigs (tree)
  "returns a list with all twigs of tree, bottom-up, depth-first"
  (if (listp tree)
      (append (list tree) (mapcan 'all-twigs (rest tree)))))

(defun all-prunes (tree internal depth)
  "returns all prunes of tree (with cuts at root if internal)"
  (cond    ((atom tree)          (list tree))
           ((< depth 1)          (list (root-label tree)))
           ((depth-1 tree)       (if internal (list (root-label tree) tree) (list tree)))
        (t
         (let ((rootlabel (root-label tree))
               (left-prunes (all-prunes (left-branch tree) t (- depth 1)))
               (right-prunes (all-prunes (right-branch tree) t (- depth 1)))
               (result (if internal (list (root-label tree)) nil)))
           (dolist (left-variant left-prunes result)
             (dolist (right-variant right-prunes)
                   (let ((new-prune (list rootlabel left-variant right-variant)))
                         (if (valid-subtree new-prune)
                             (push new-prune result)))))))))
```

Plan

- Some background about human language and animal communication
- Nurture or nature?
- Vocal learning
- Artificial Language Learning
  - Segments & Rules project

Some facts about language: symbolicism

- Word meanings are conventional and arbitrary: even e.g. onomatopoeia: cock-a-doodle-do, cocorico, kukeleku
- Adult speakers know >10000 words; children learn these words over the course of just a few years.
- Languages are transmitted culturally, and slowly change over the course of a number of generations, giving rise to an enormous variety of over > 6000 languages.

Some facts about language: phonemic coding

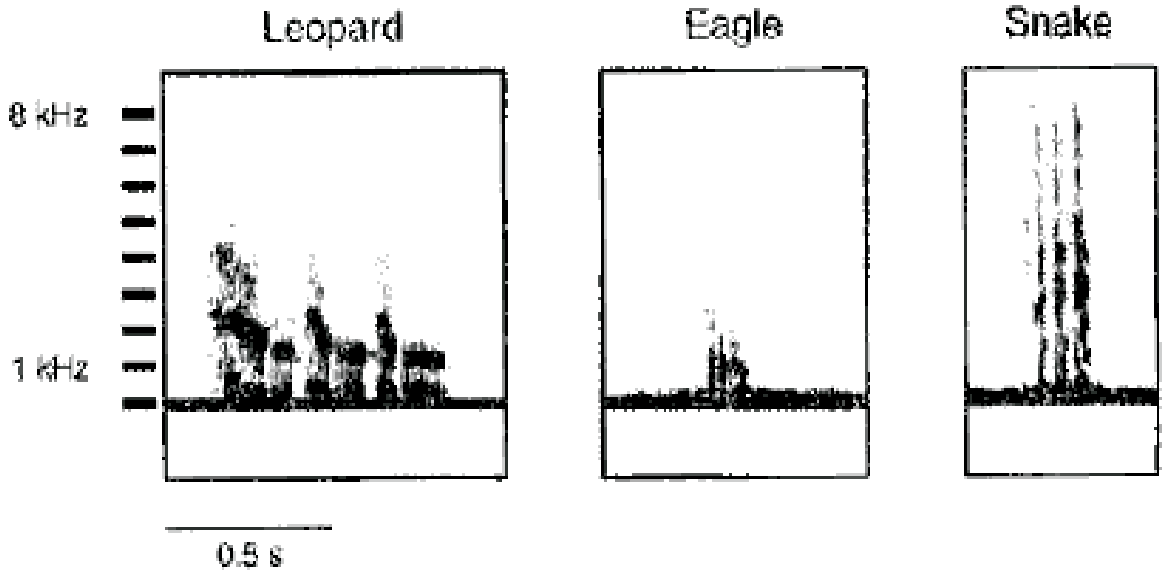
- Words are built-up from (meaningless) basic speech sounds, the phonemes;
- Phonemes are defined as the minimal difference in sound that corresponds to a difference in meaning. E.g. minimal pairs:

  bed bad bet bat /e/, /a/, /d/, /t/

- Phonemes are different in every language (and dialect), but phonemic coding is universal.

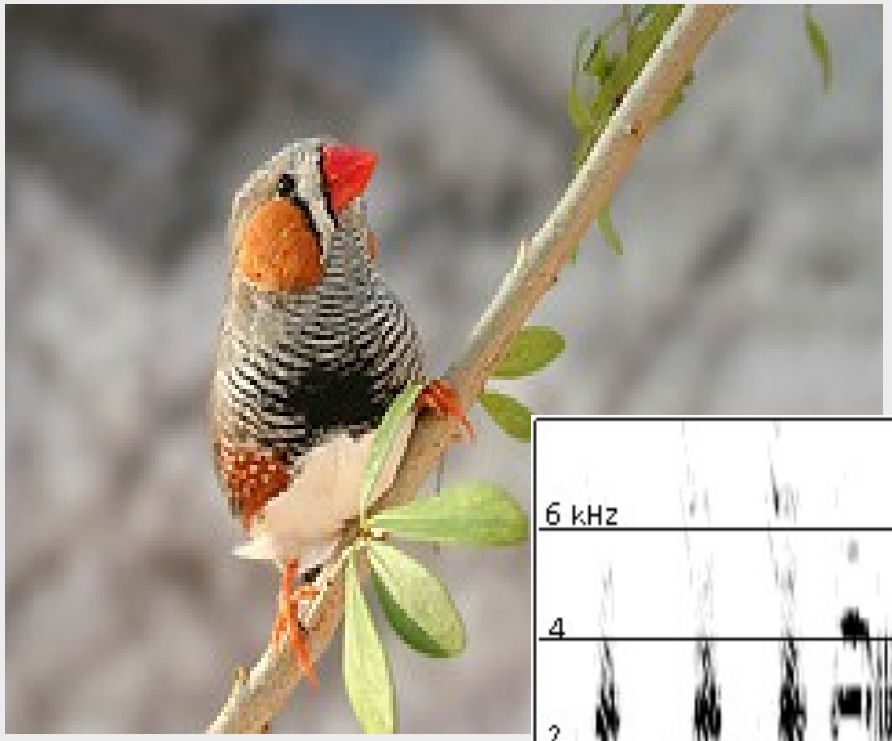
Arbitrariness: Vervet monkey alarm calls

(Struhsaker'67; Seyfarth, Cheney & Marler'80; squirrels: Sherman'77)

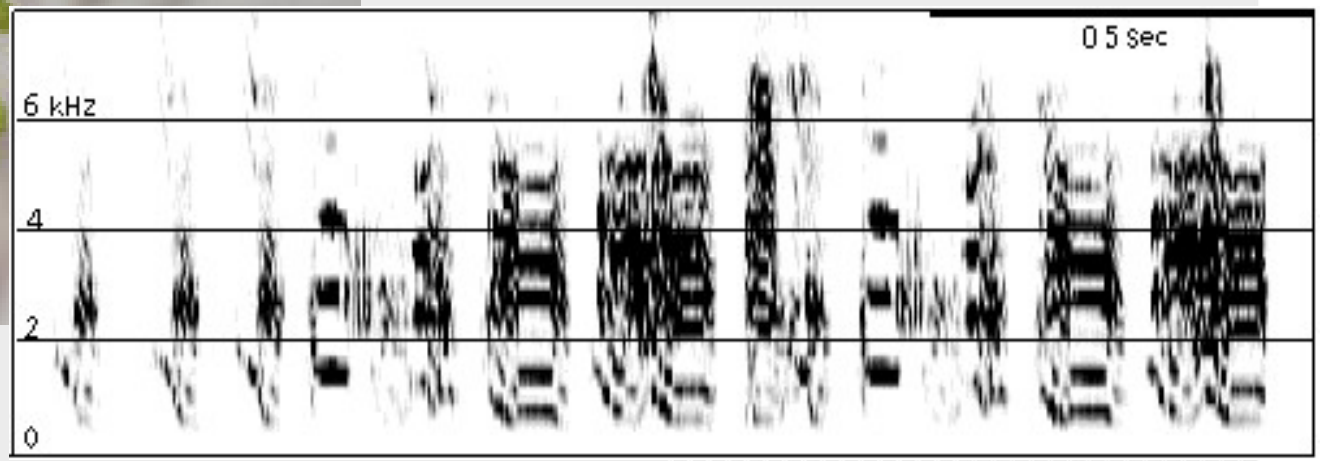


“chirp” “grunt” “chutter”

Cultural transmission, vocal learning, discreteness: songbirds

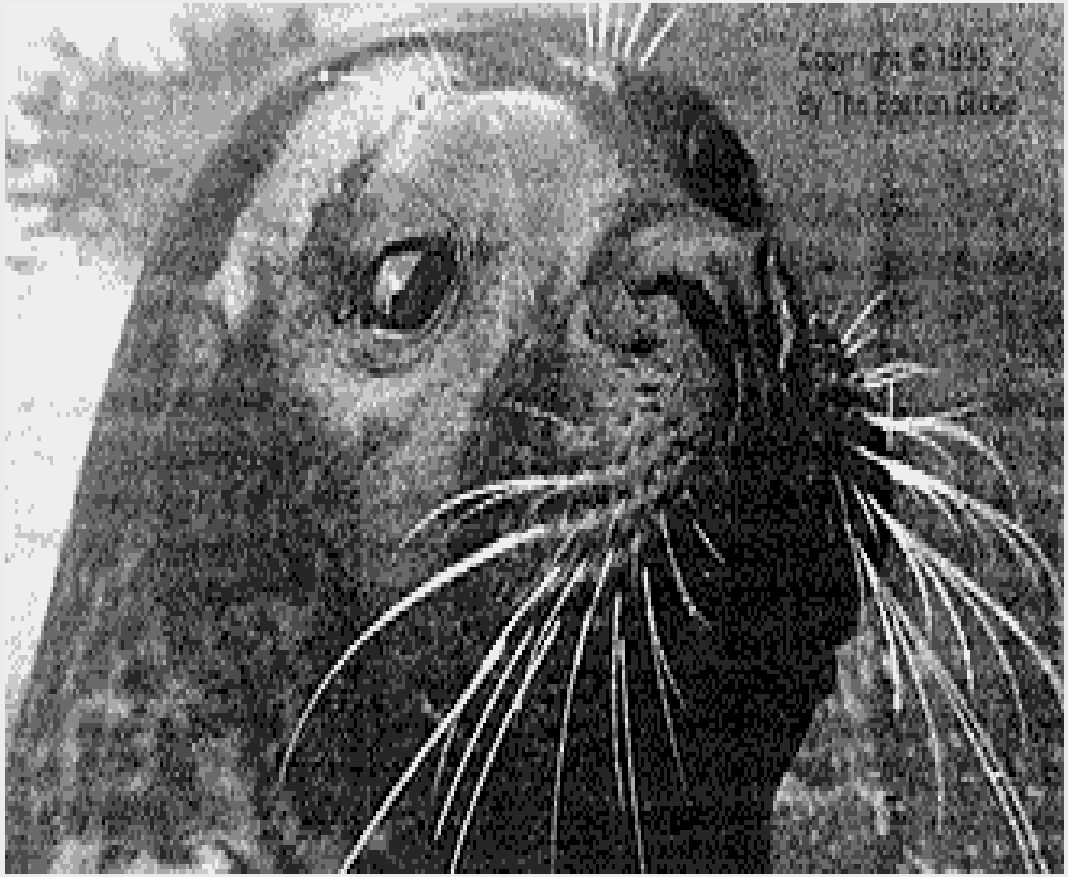


zebrafinch



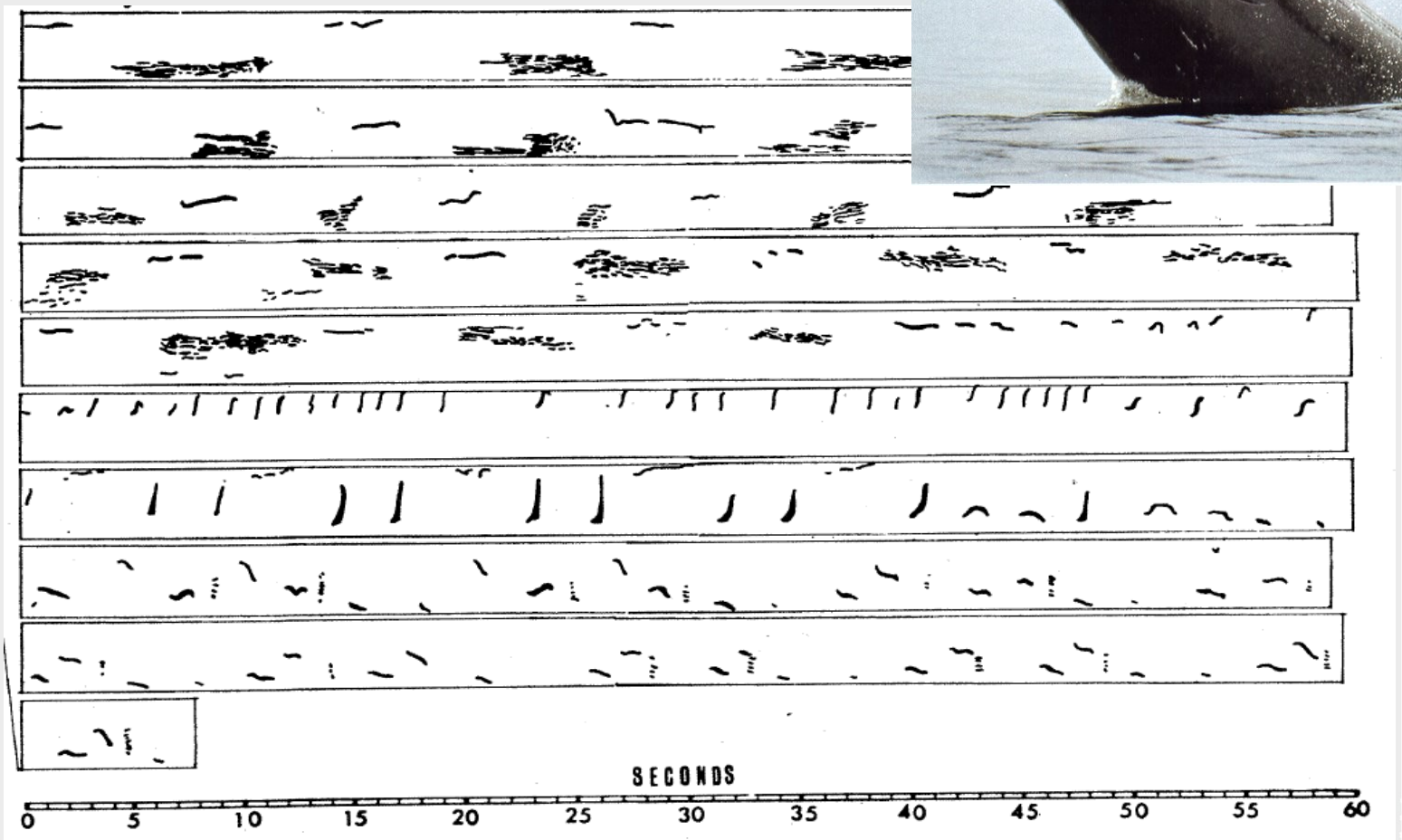
Doupe & Kuhl 1999

Vocal learning: seals



Hoover, the talking seal

Humpback whales



(Watlington'63; Payne & McVay'71)

Some facts about language: compositional semantics

Sentences are built-up from meaningful words.

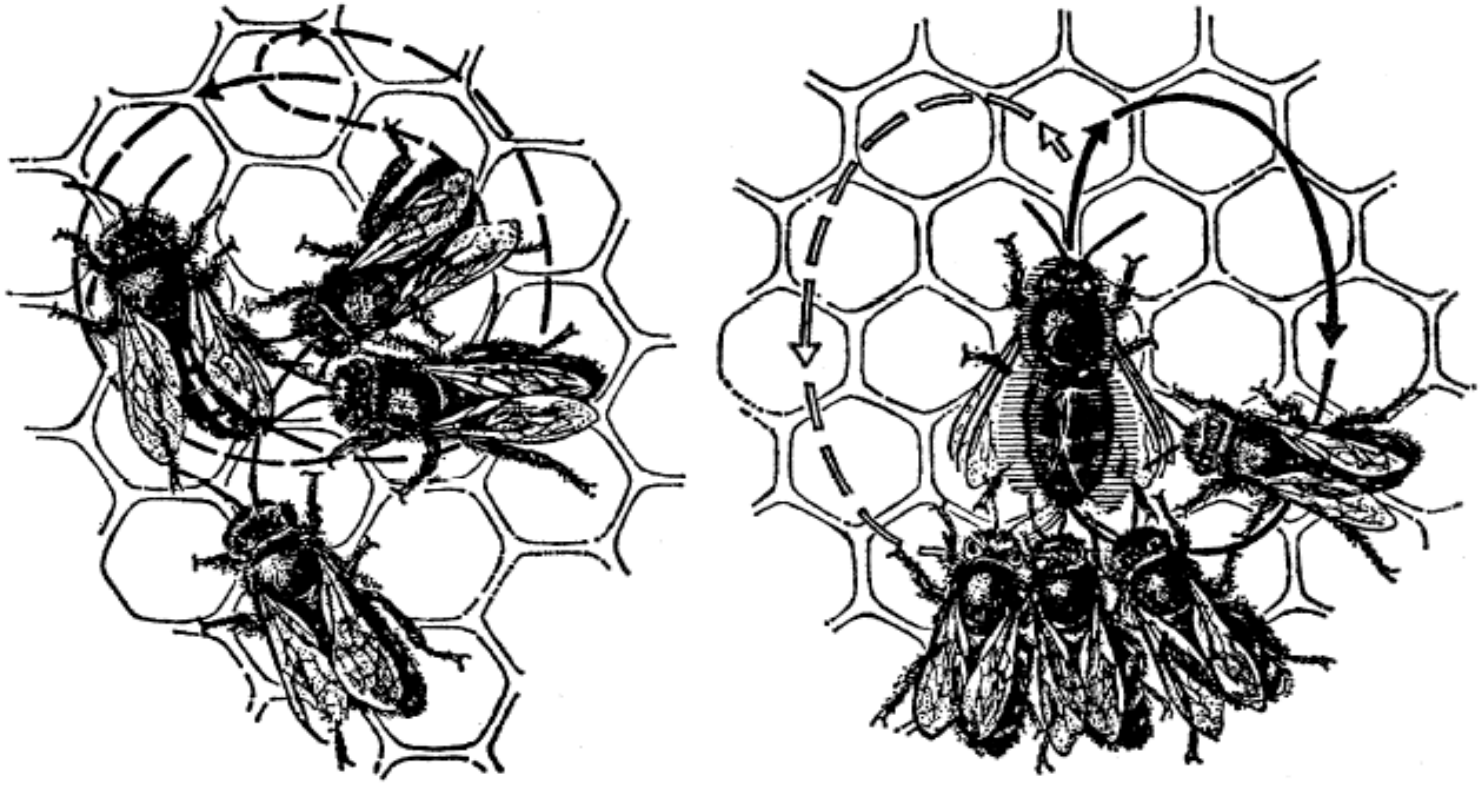
Words can be built from meaningful morphemes.
E.g.: “he walk-s”, disproof, disallow, rearrange

The meaning of a larger whole is determined by the meaning of the words and the way they are put together (compositionality),

i.e. word order and morphological marking and agreement: e.g. “In vino veritas est.”

Bee dance

(von Frisch'65, '74)



a. Round dance

b. Tail-wagging dance

Fig. 1. Running curve of the bee (a) during round dance and (b) during tail-wagging dance. Bees that follow the dancer take in information.

von Frisch'74, *Science* 4152:664

(Zuberbühler'02 An. Beh. 63)

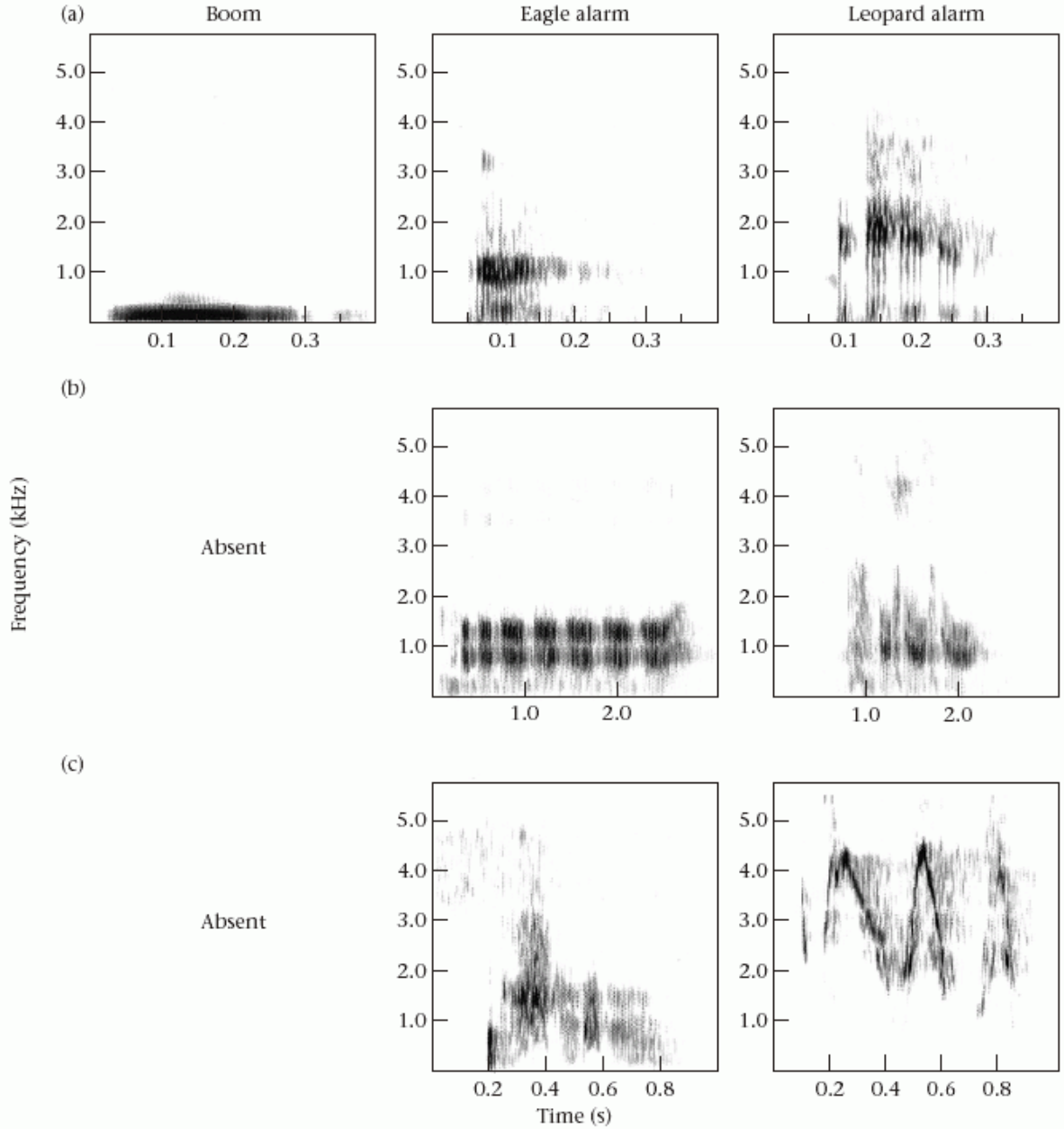
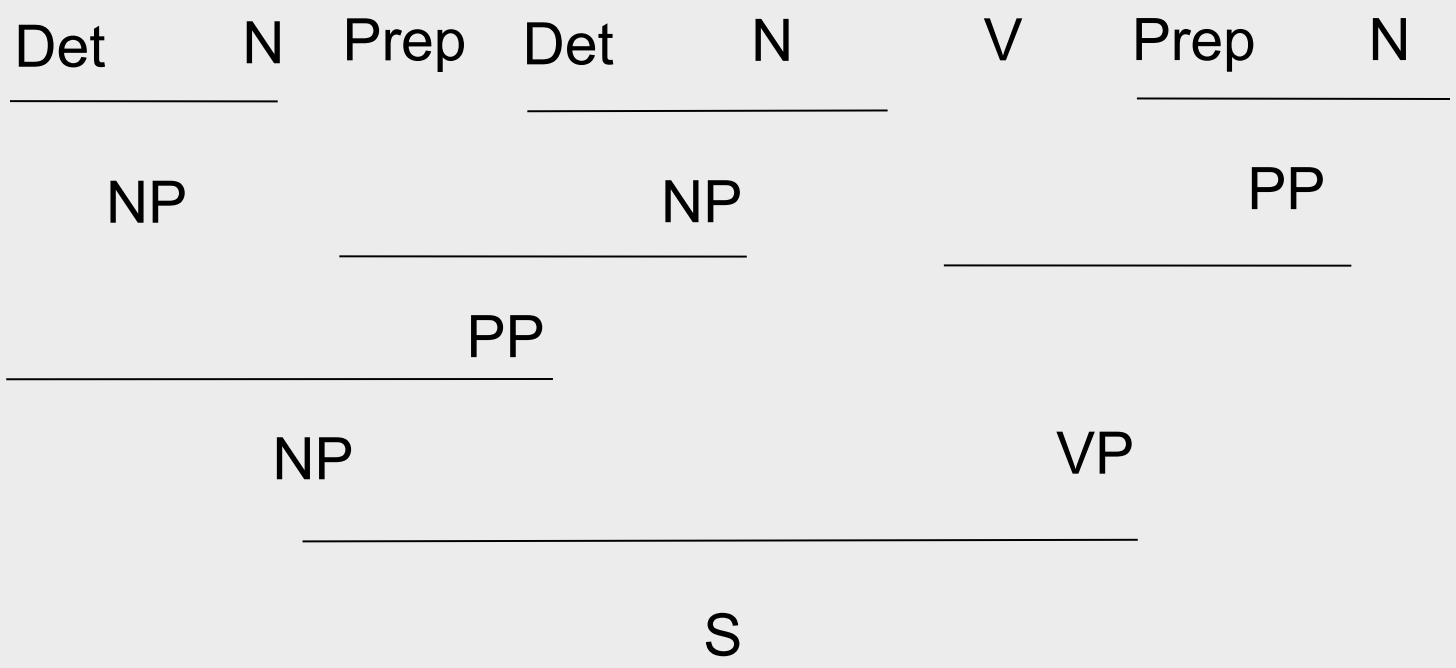


Figure 1. Spectrographic illustrations of the vocalizations used in this study. (a) Male Campbell's monkey, (b) male Diana monkey, (c) female Diana monkey. Male Campbell's and male Diana monkey alarm calls were used as playback stimuli (see text). Recordings were digitized at a sampling rate of 44 kHz (16 bits accuracy) and displayed using a 256-point Fourier transformation (Hamming window function) that resulted in wide-band spectrograms (analysis resolution: 700 Hz with 21.5 Hz/0.72 ms accuracy).

Some facts about language: Hierarchical phrase structure

- *The man with the gun is in town*



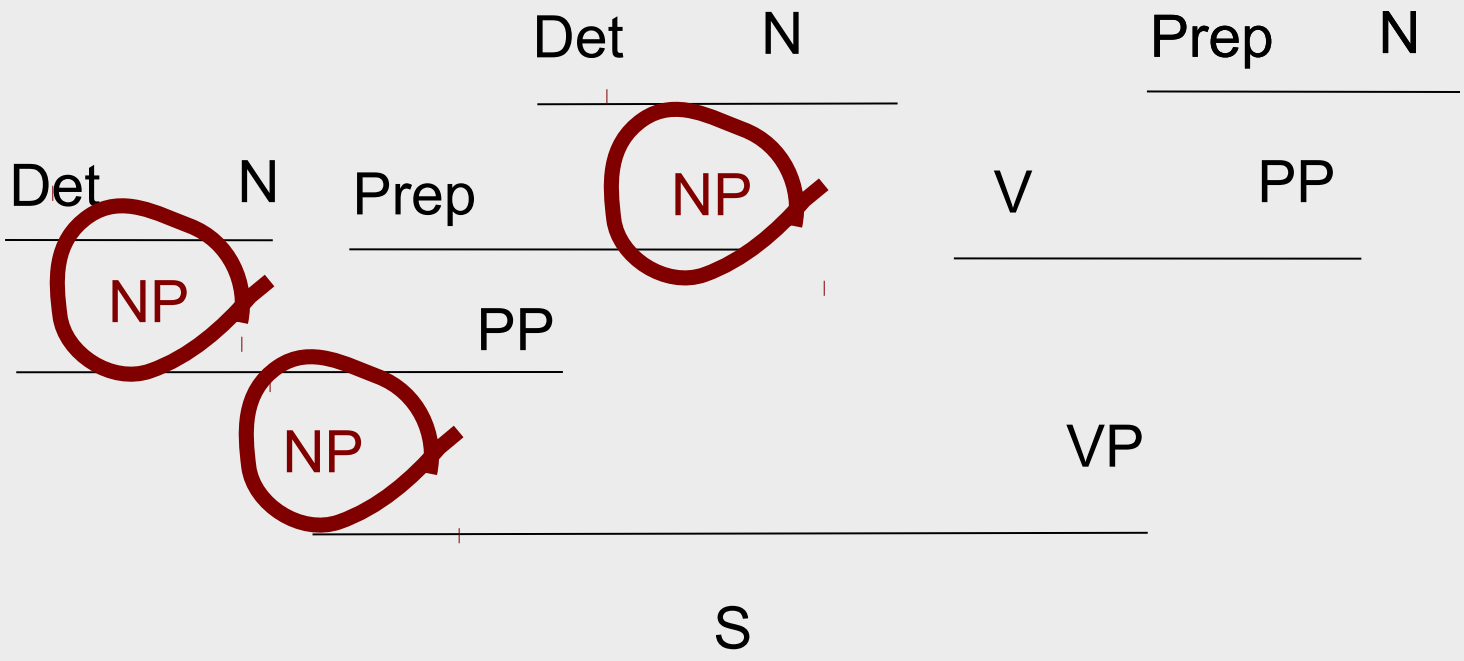
Some facts about language: Recursion

Phrases of category *X* can be embedded in a phrase of the same category *X.*

'the man', 'the gun' and 'the man with the gun' are all noun phrases (NP):

- *The man is in town*
- *The gun fell*

- *The man with the gun is in town*



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)

a. Gilligan behaupte dass Blair das Publikum belügte.

b. Gilligan behaupte dass Campbell Blair das Publikum belügen holf.

c. Gilligan behaupte dass Kelly Campbell Blair das Publikum belügen helfen sah.

(CENTER-EMBEDDING)

a. Gilligan beweert dat Blair het publiek bedroog.

b. Gilligan beweert dat Campbell Blair het publiek hielp bedriegen.

c. Gilligan beweert dat Kelly Campbell Blair het publiek zag helpen bedriegen.

(CROSSING DEPENENCIES)
(Huybrechts, 1983)

a. Gilligan claims that Kelly saw Campbell help Blair deceive the public. (TAIL RECURSION)

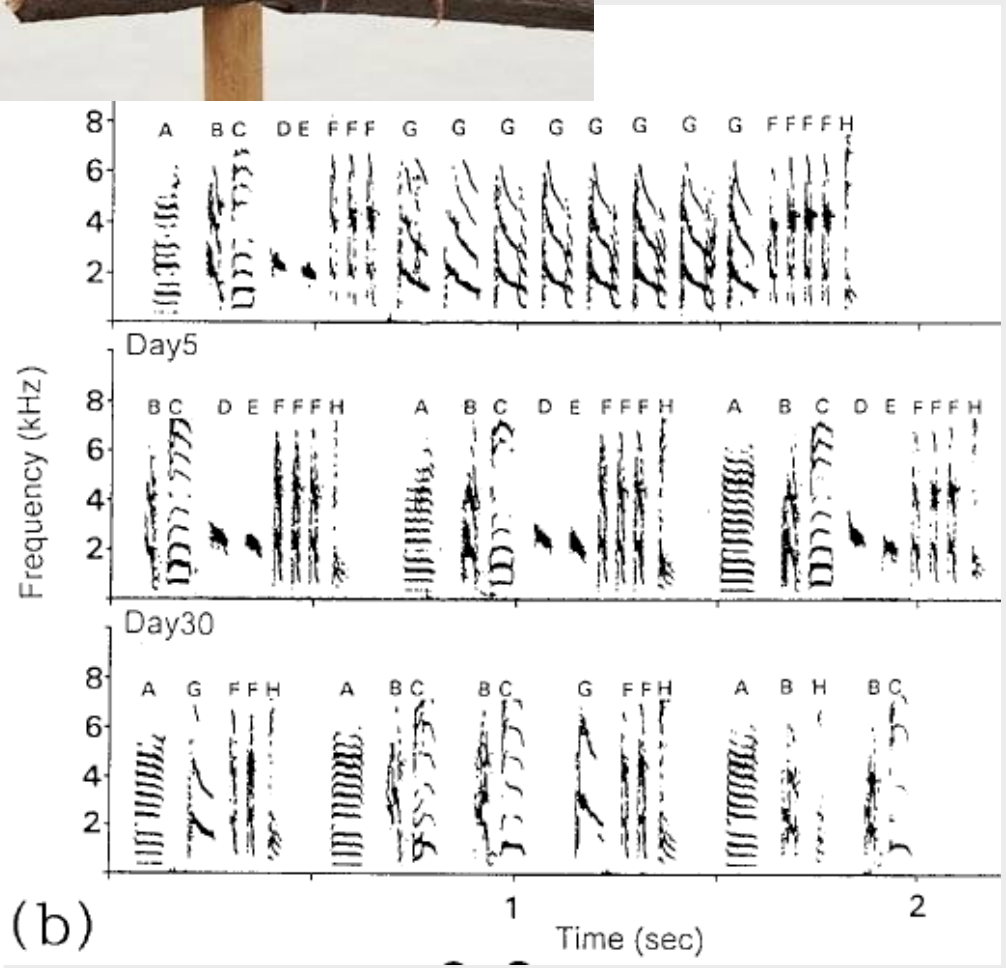
b. Gilligan behaupte dass Kelly Campbell Blair das Publikum belügen helfen sah.
(CENTER EMBEDDING)

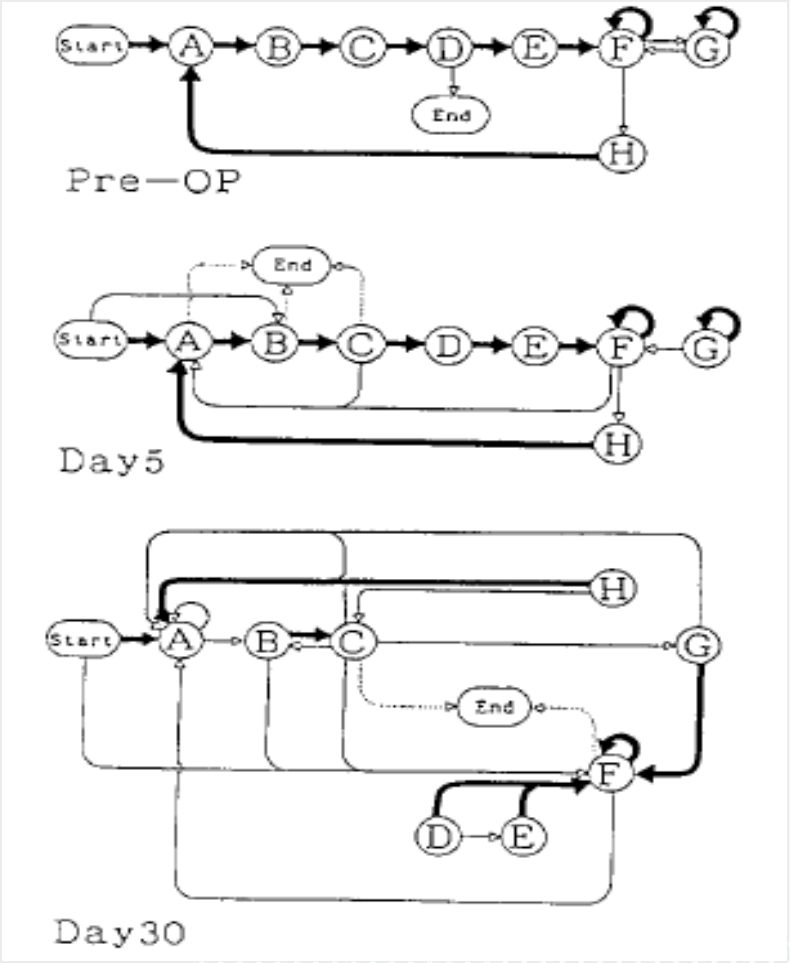
c. Gilligan beweert dat Kelly Campbell Blair het publiek zag helpen bedriegen.
(CROSSING DEPENDENCIES)

Humpback whales

Payne & McVay (1971)

Bird song





(Okanoya & Yamaguchi'97; *J. Neurobio. 33,4)*

Apes: bonobo Kanzi

(Truswell, in prep)

645. (C) *I want Kanzi to grab Rose*.

(Kanzi turns around and grabs Rose on the leg, then walks away.)

581. (C) *Kanzi, tell Rose that you want to go outdoors*.

(Kanzi turns, looks at Rose, and gestures toward the play-yard door.) Rose looks in that direction and says, “You’re supposed to go over there?” (Kanzi heads toward the play-yard door, and Rose follows.)

In all these cases, however, Kanzi’s responses would be identical if he ignored the upstairs clauses, and just responded to the most embedded clause.

(Truswell, in prep)

428. (PC) *Give the water and the doggie to Rose*.

(Kanzi picks up the dog and hands it to Rose.)

526. (PC) *Give the lighter and the shoe to Rose*.

(Kanzi hands Rose the lighter, then points to some food in a bowl in the array that he would like to have to eat.)

281. (C) *Give me the milk and the lighter*.

(Kanzi does so.)

Kanzi's overall accuracy on the coordination construction is at chance level (25%).

Some facts about language: idiosyncratic constraints

(1) a. a violin which this sonata is hard to play upon
b. *a sonata which this violin is hard to play upon (Steedman'03)

(2) a. Every acorn grew into an oak.
b. Every oak grew out of an acorn.
c. An oak grew out of every acorn.
d. *An acorn grew into every oak. (Gruber, 1965)

(3) a. a book which I hope I will write, and I fear that most people will burn without reading
b. *Three mathematicians in ten derive a lemma and in a hundred prove completeness (Steedman'03)

Some facts about language: incremental processing

The horse raced past the barn fell.
The old man the boat.
The man who whistles tunes pianos.
Time flies like an arrow. Fruit flies like a banana.
The Australian woman saw the famous doctor had been drinking quite a lot.
Before the woman visited the famous doctor had been drinking quite a lot.
While the pilot was flying the horse that had arrived stood over by the fence.

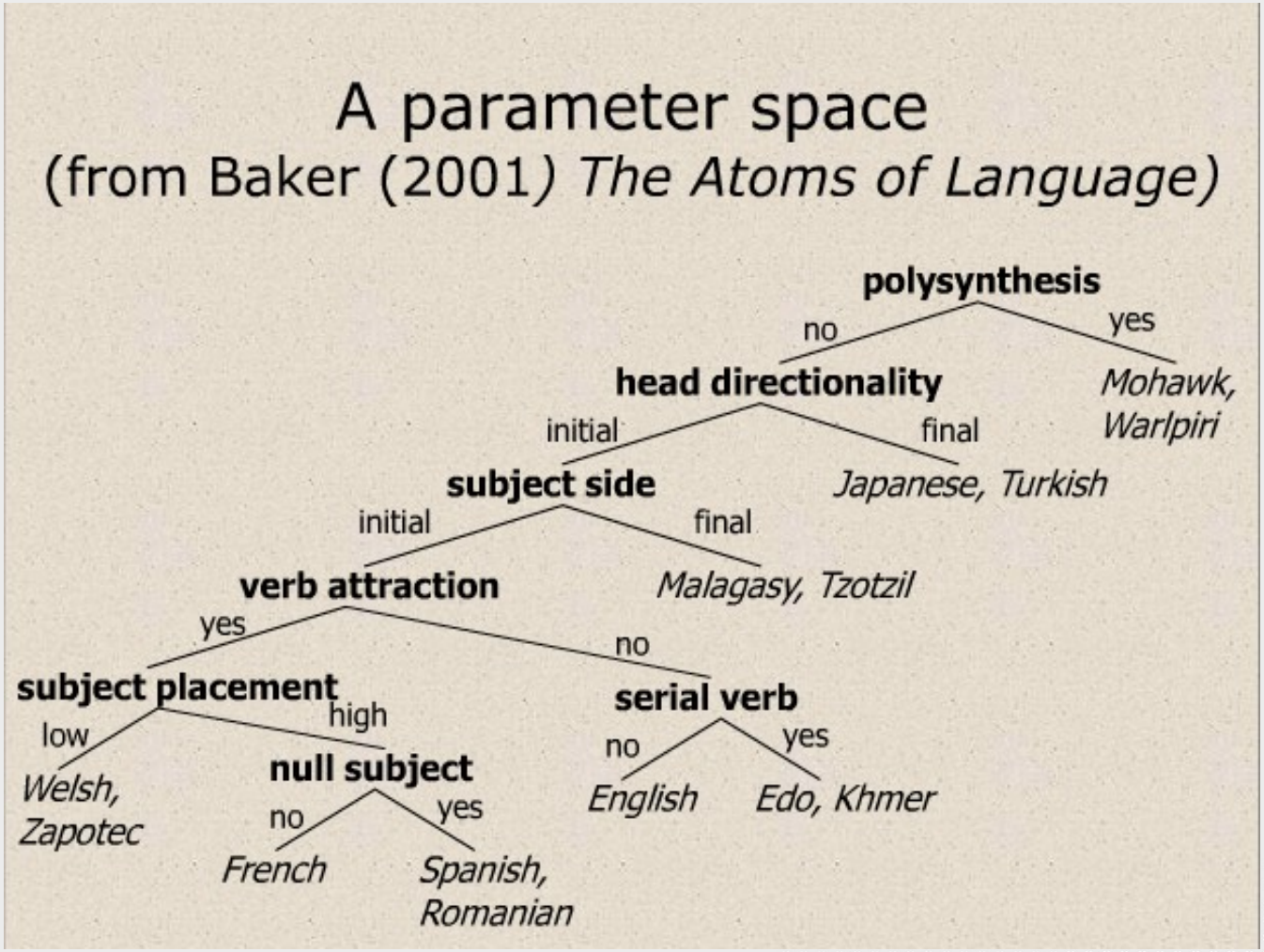
(wikipedia; Sturt, Pickering, Crocker, 1999; Pickering, Traxler, and Crocker,2000)

Human language:

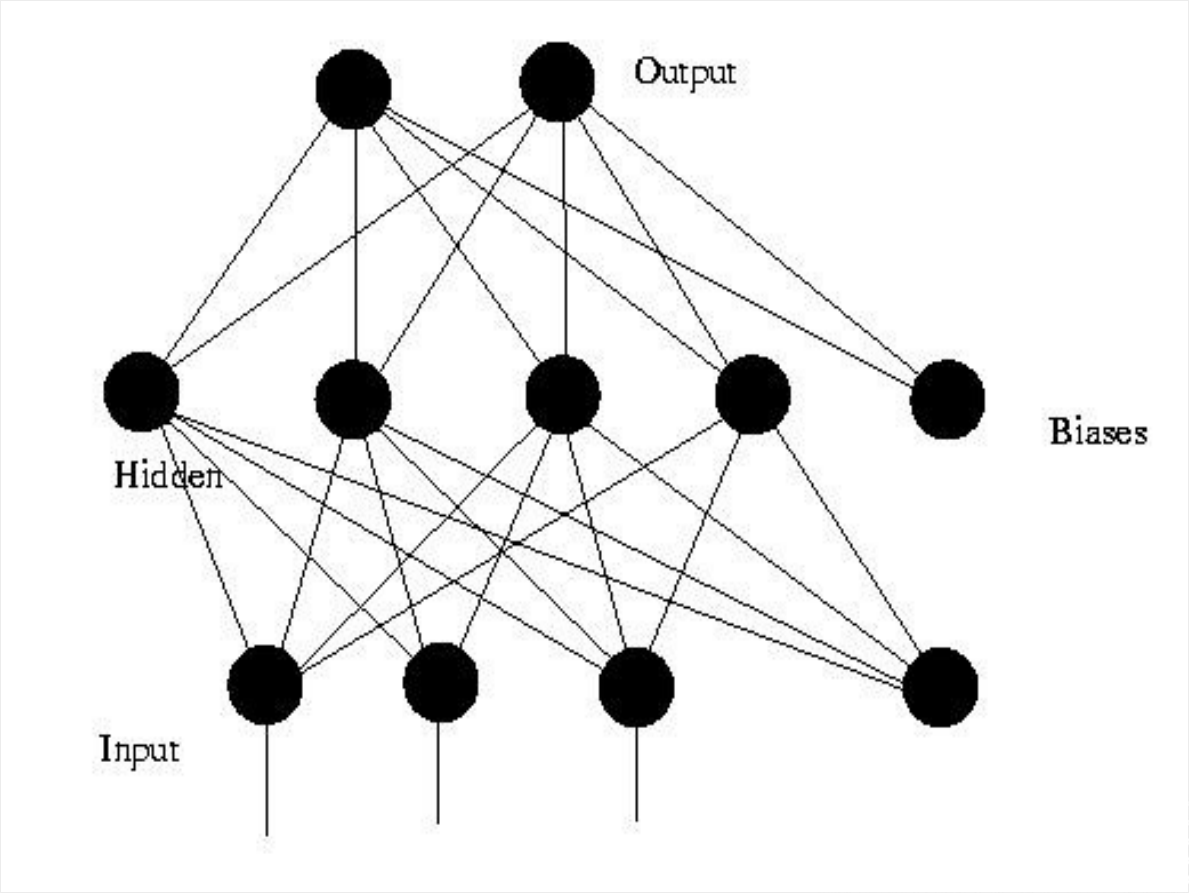
- Is an extremely complex and varied phenomenon;
- Orders of magnitude more complex than any animal communication system discovered so far;
- Requires extensive memory and sophisticated computations to be produced, interpreted and learned.

Is this complexity due to nurture or nature?

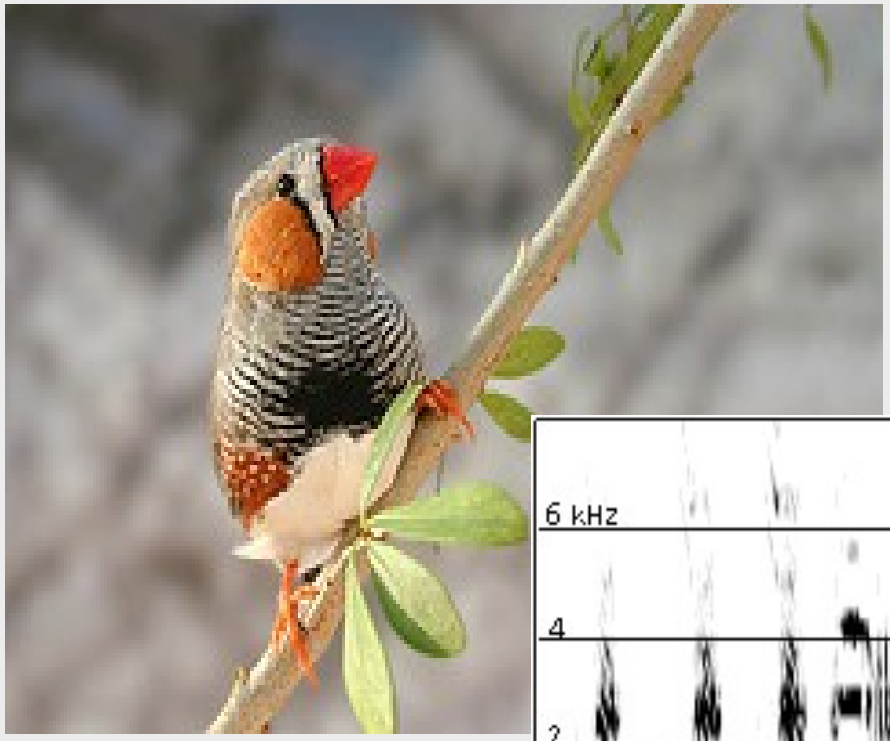
1980s Nativism

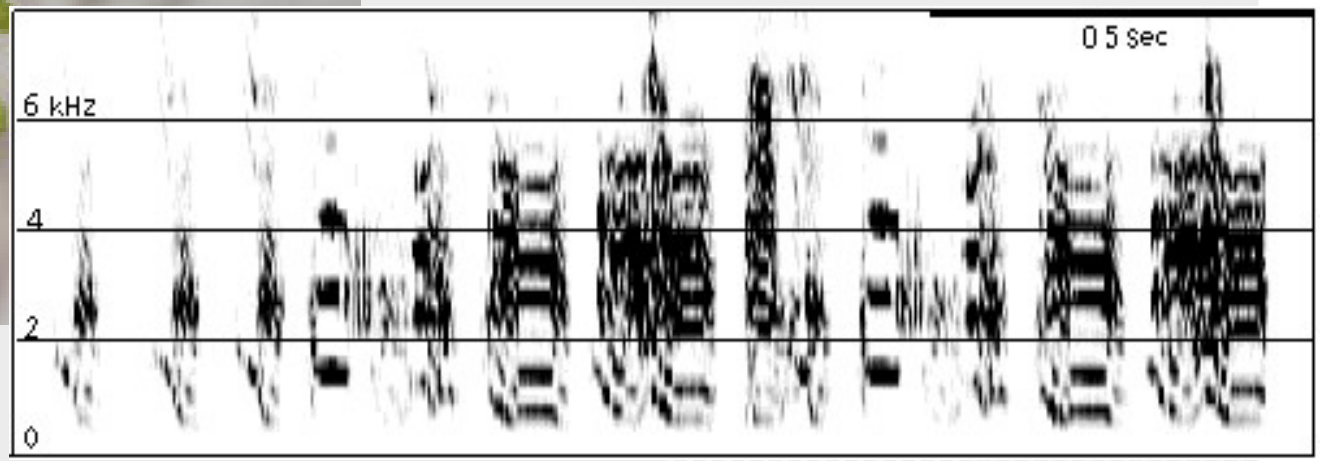


1980s Empiricism

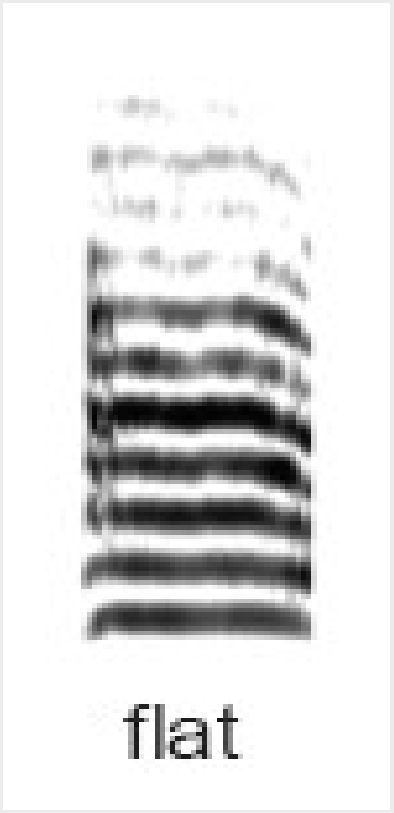


Zebrafinch song



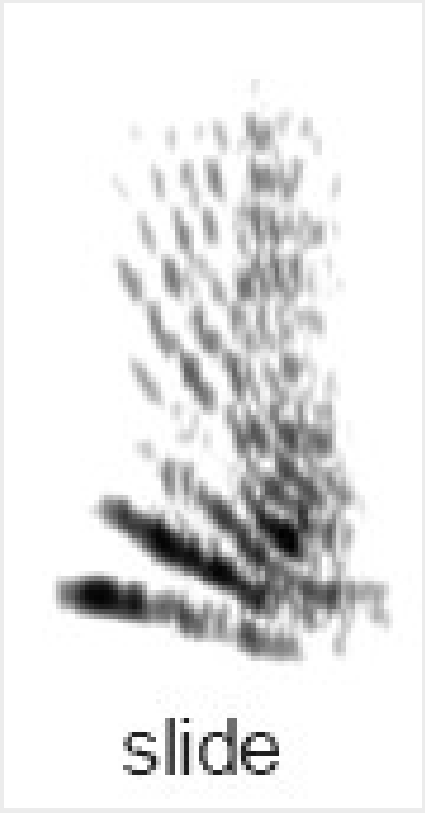


Song typology

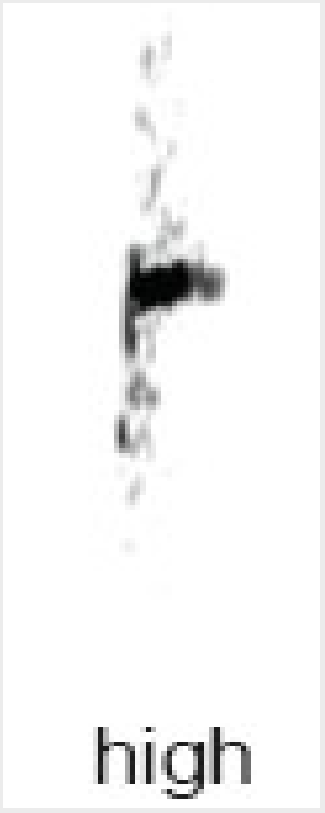


A

B

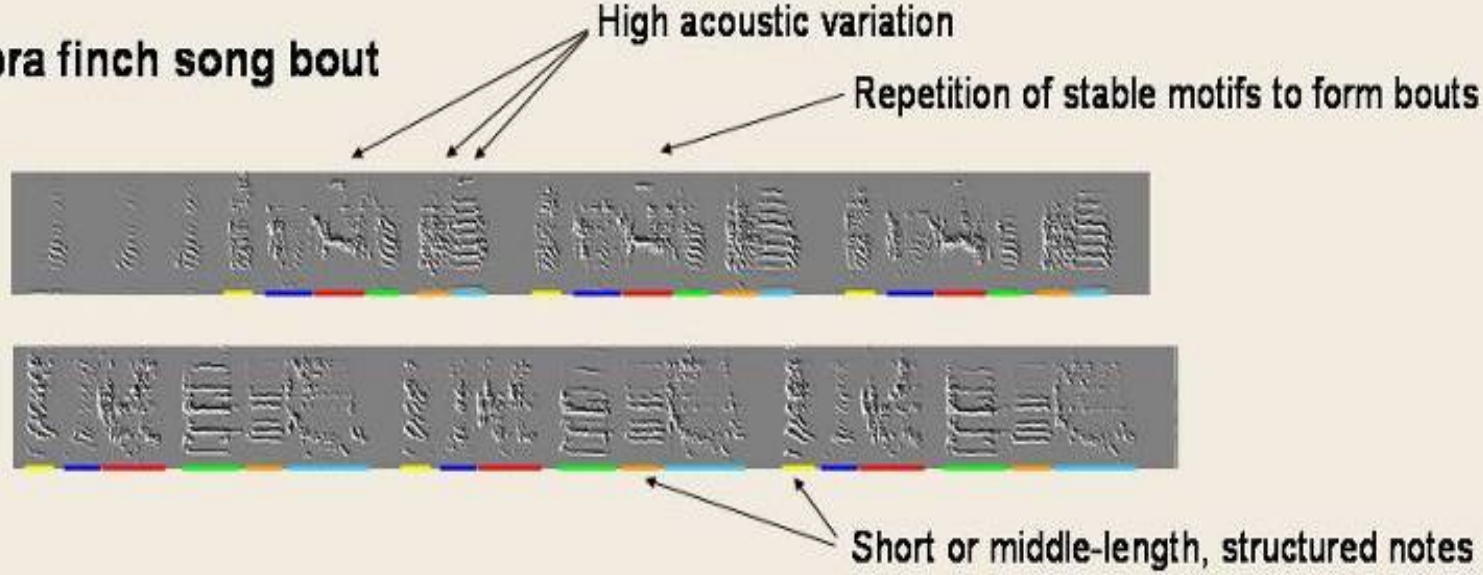


C

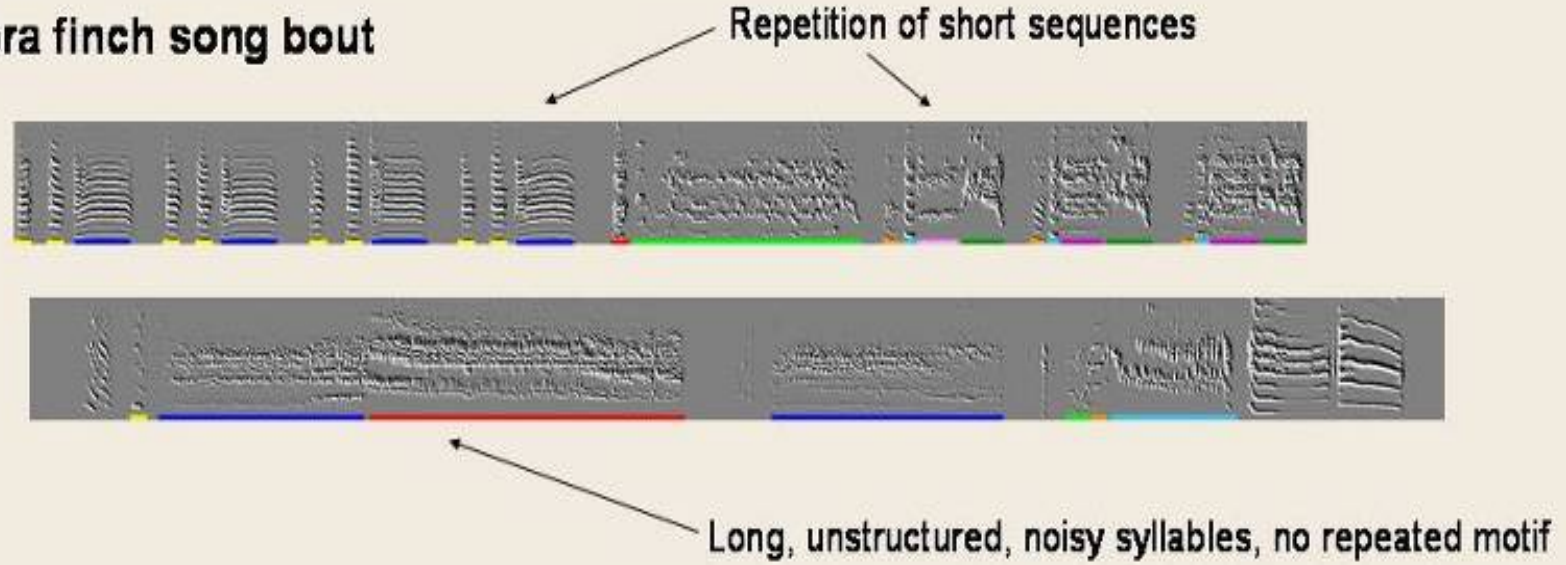


D

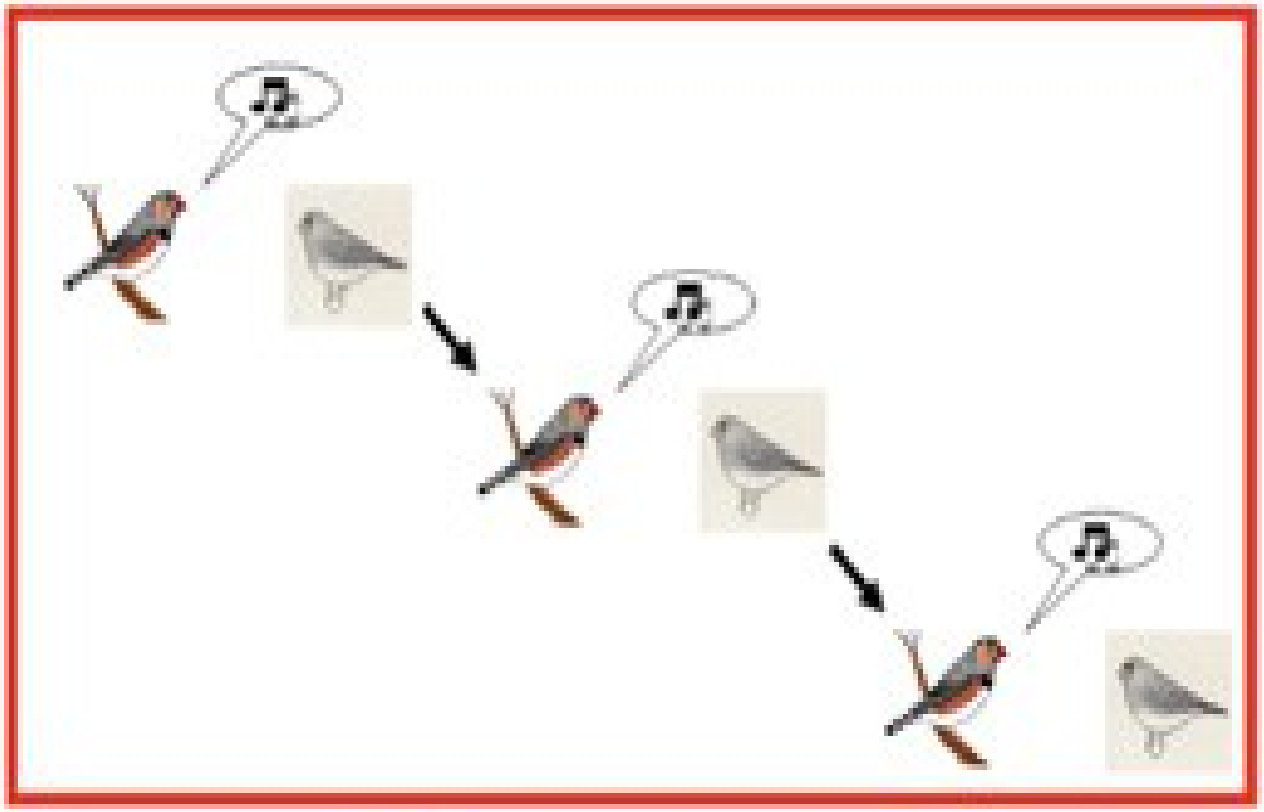
Normal zebra finch song bout



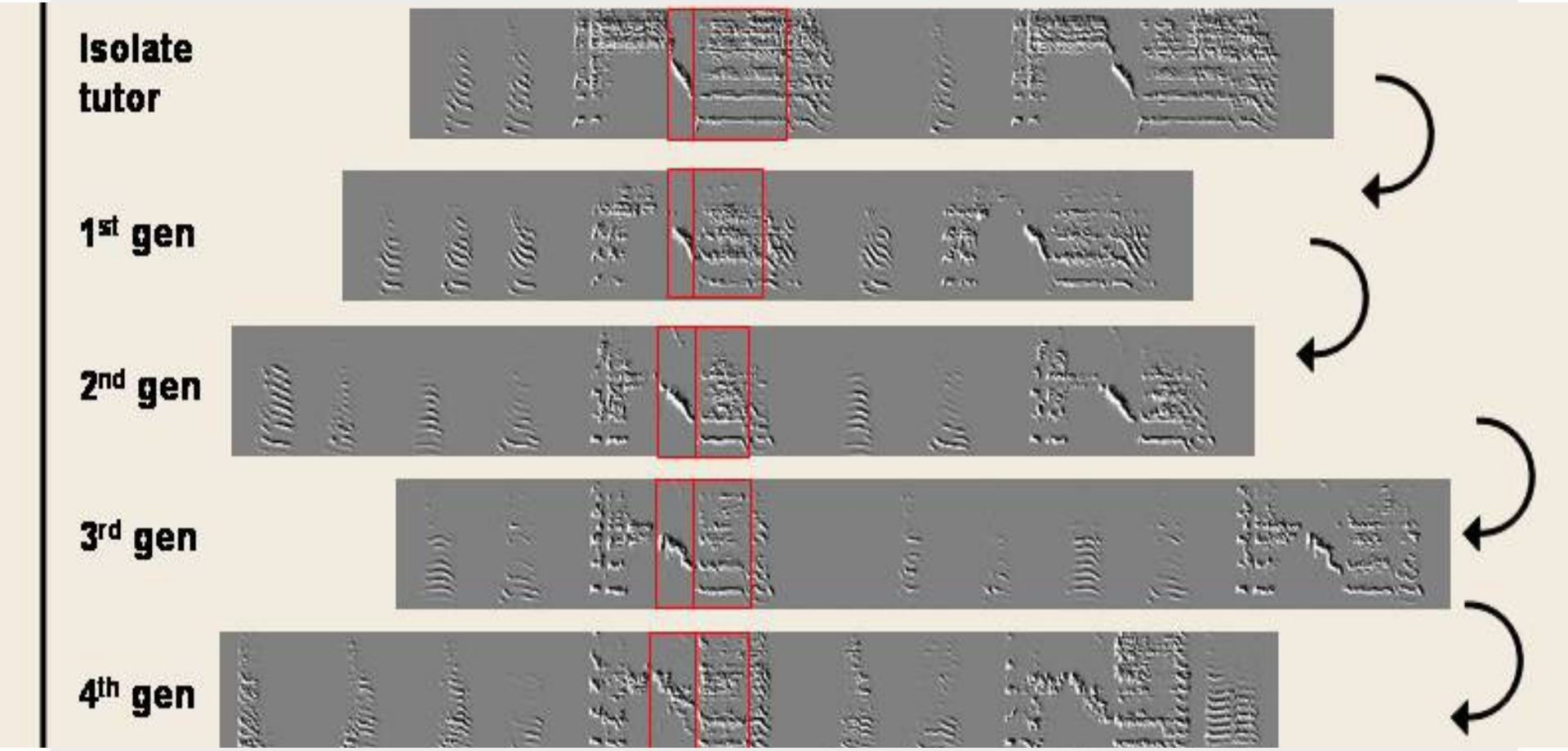
Isolate zebra finch song bout

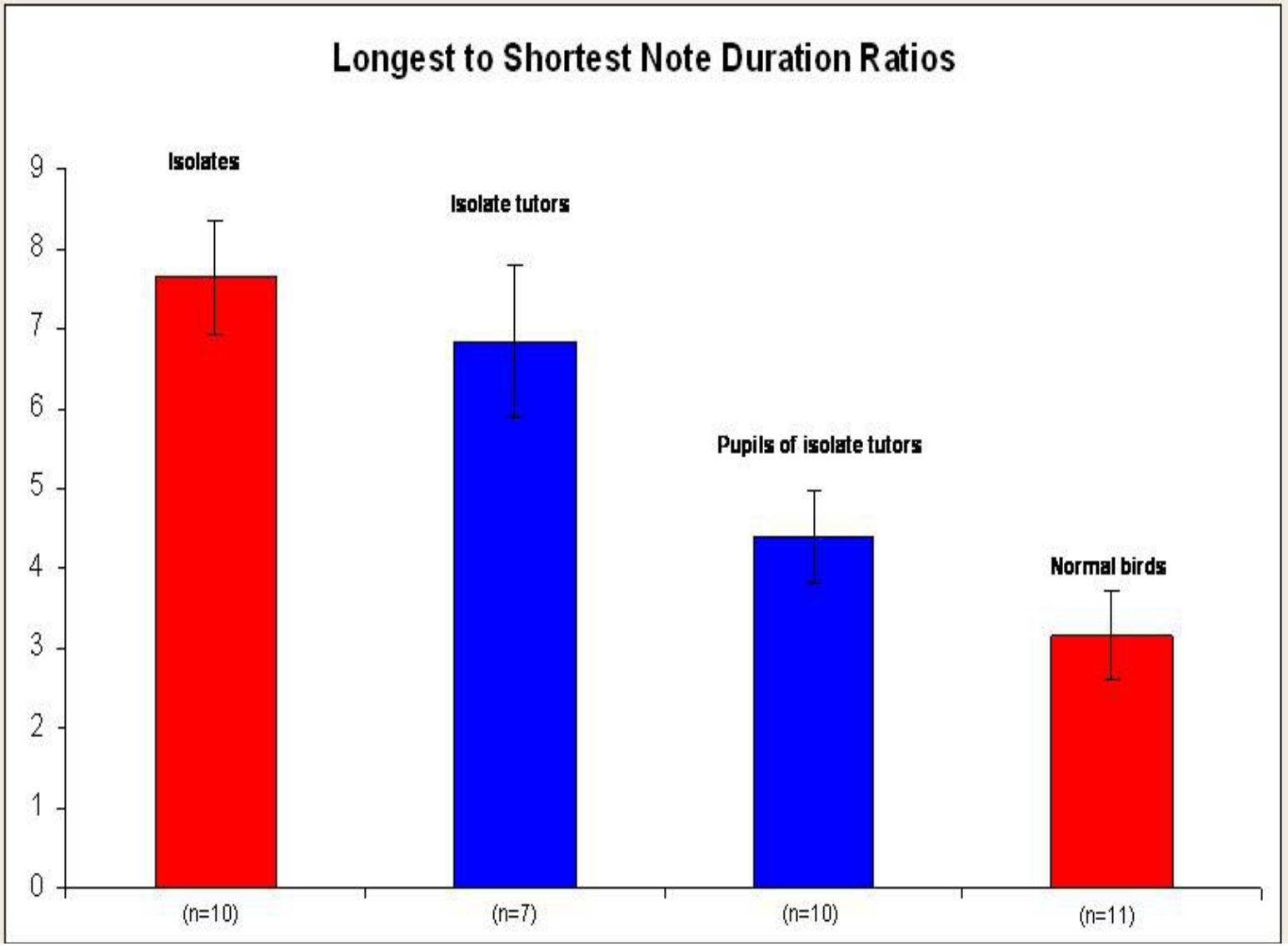


Feher et al. 2009: Iterated Learning



(Kirby, 2002; Zuidema, 2003)


Isolate tutor
1st gen
2nd gen
3rd gen
4th gen


Longest to Shortest Note Duration Ratios
Isolates
Isolate tutors
Pupils of isolate tutors
Normal birds
9
8
7
6
5
4
3
2
1
0
(n=10)
(n=7)
(n=10)
(n=11)

- 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

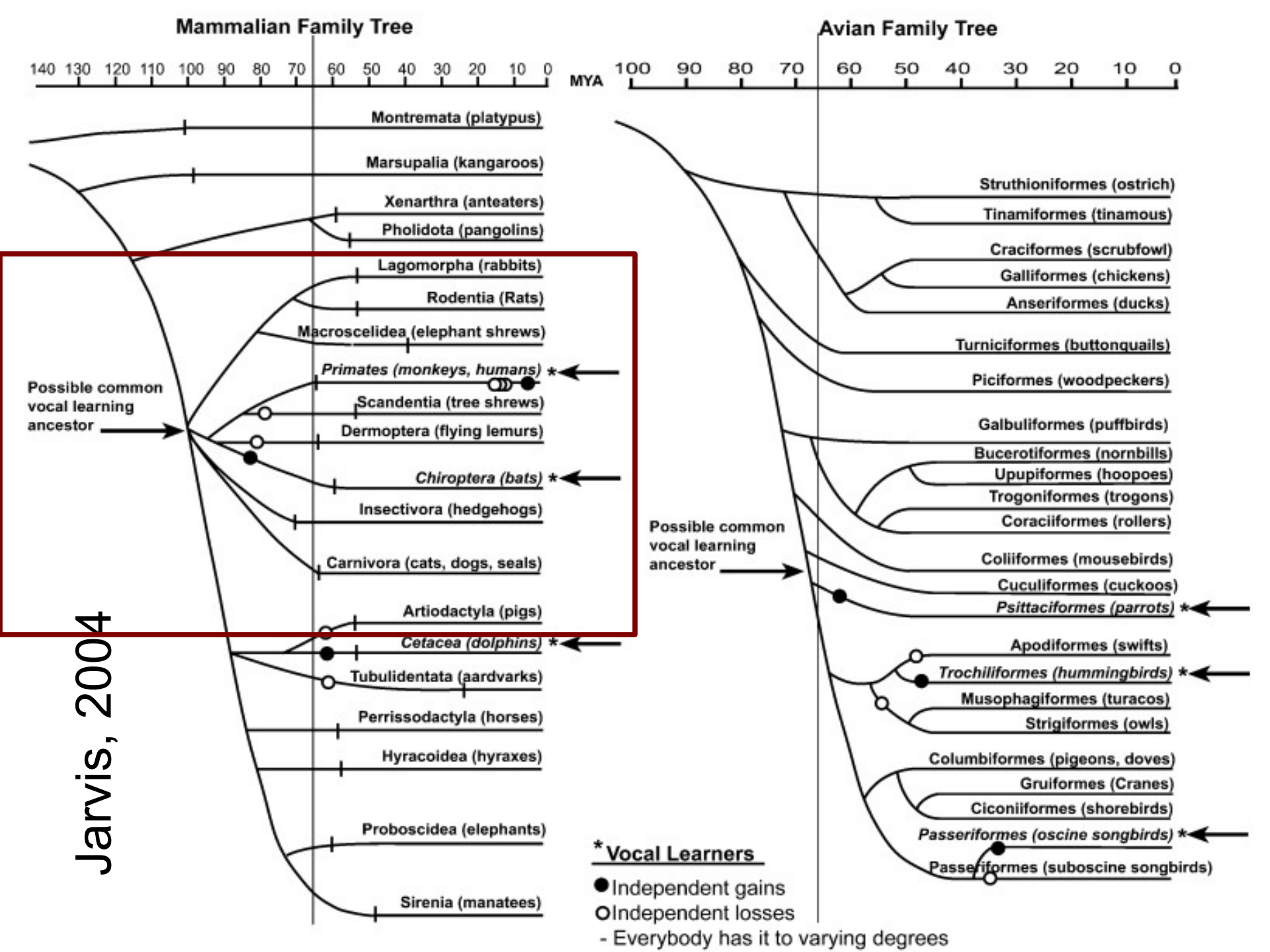
    environmental factors;

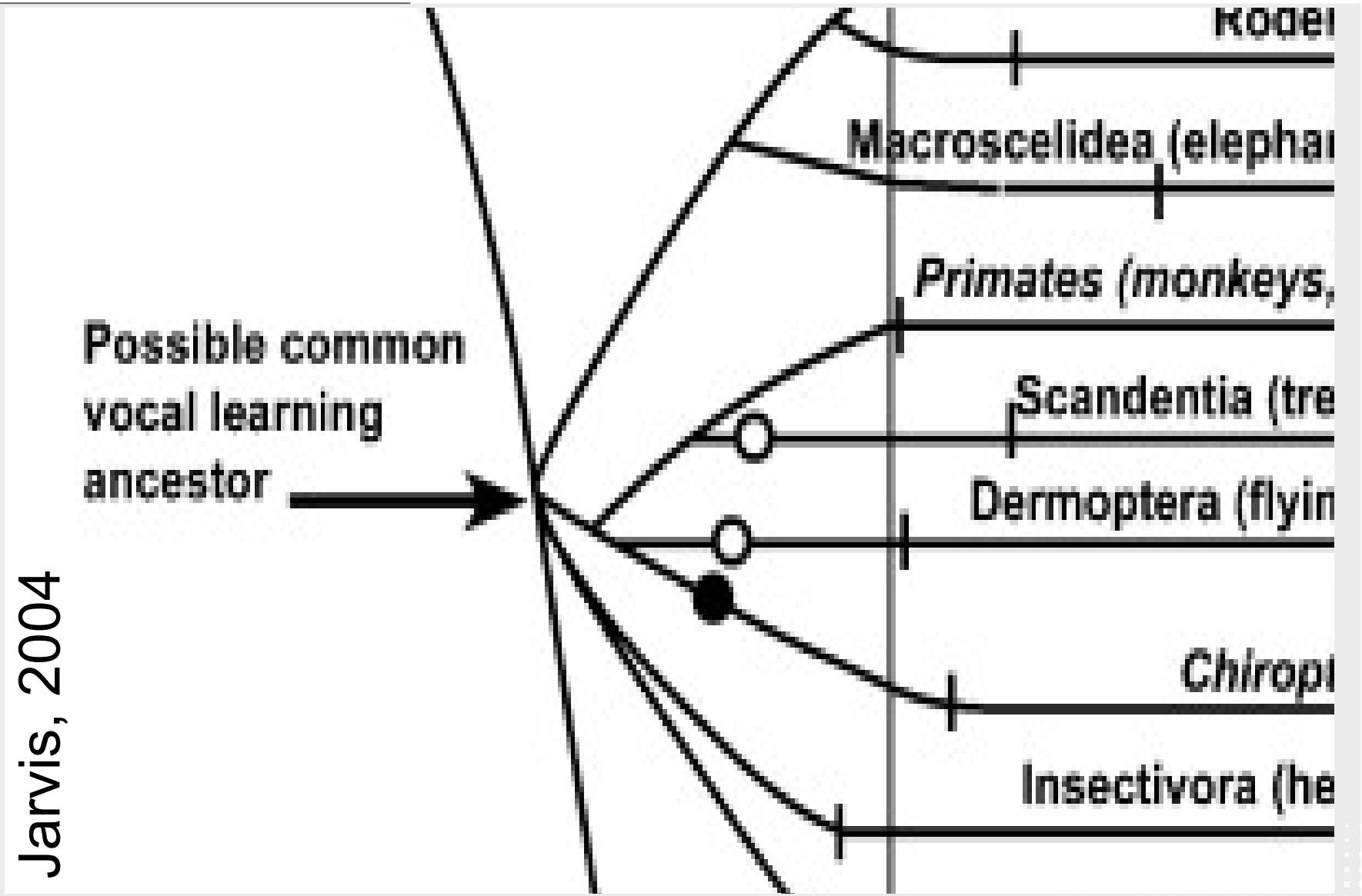
Lessons learned

- 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

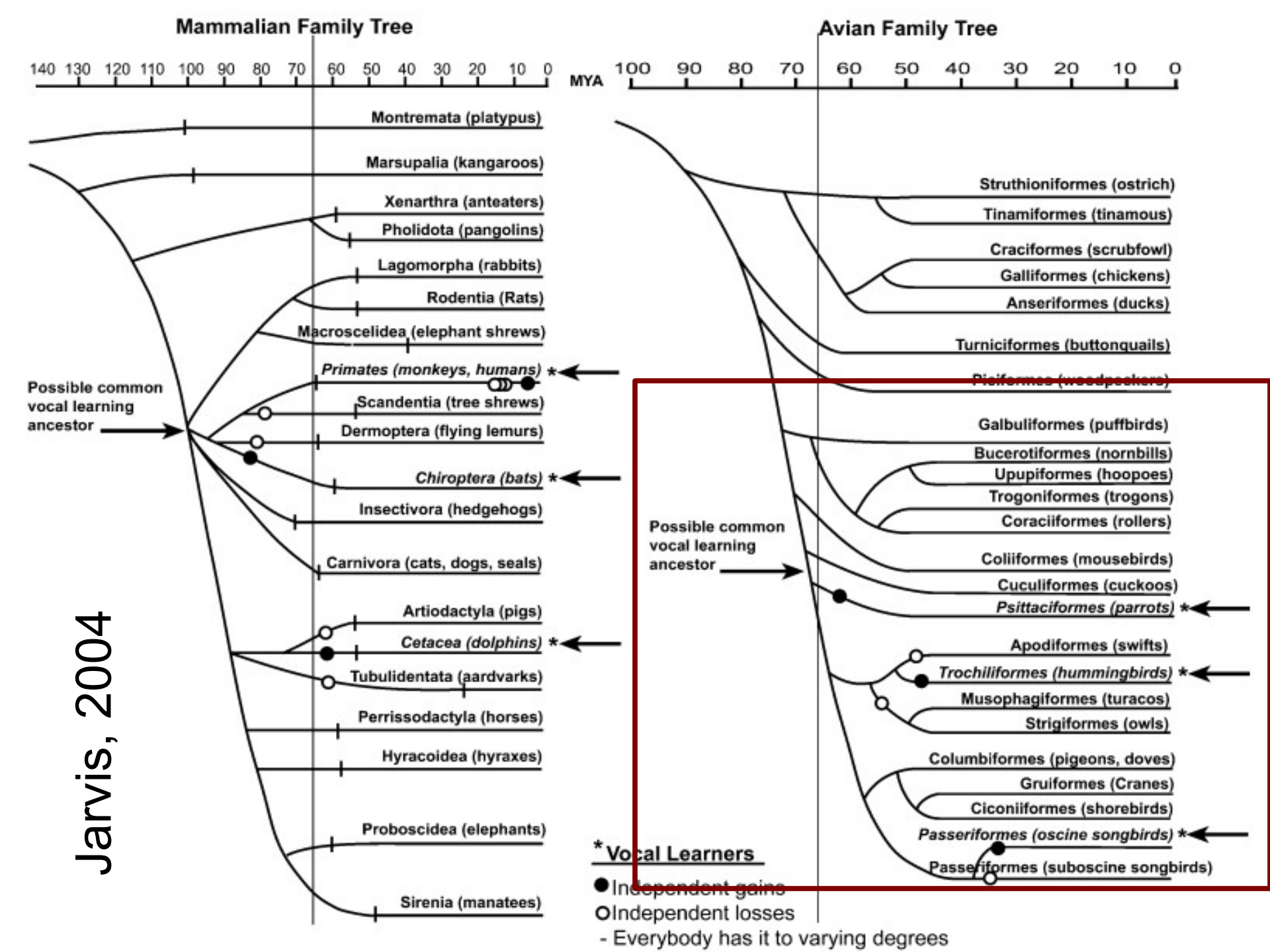
Lessons learned (ctd)

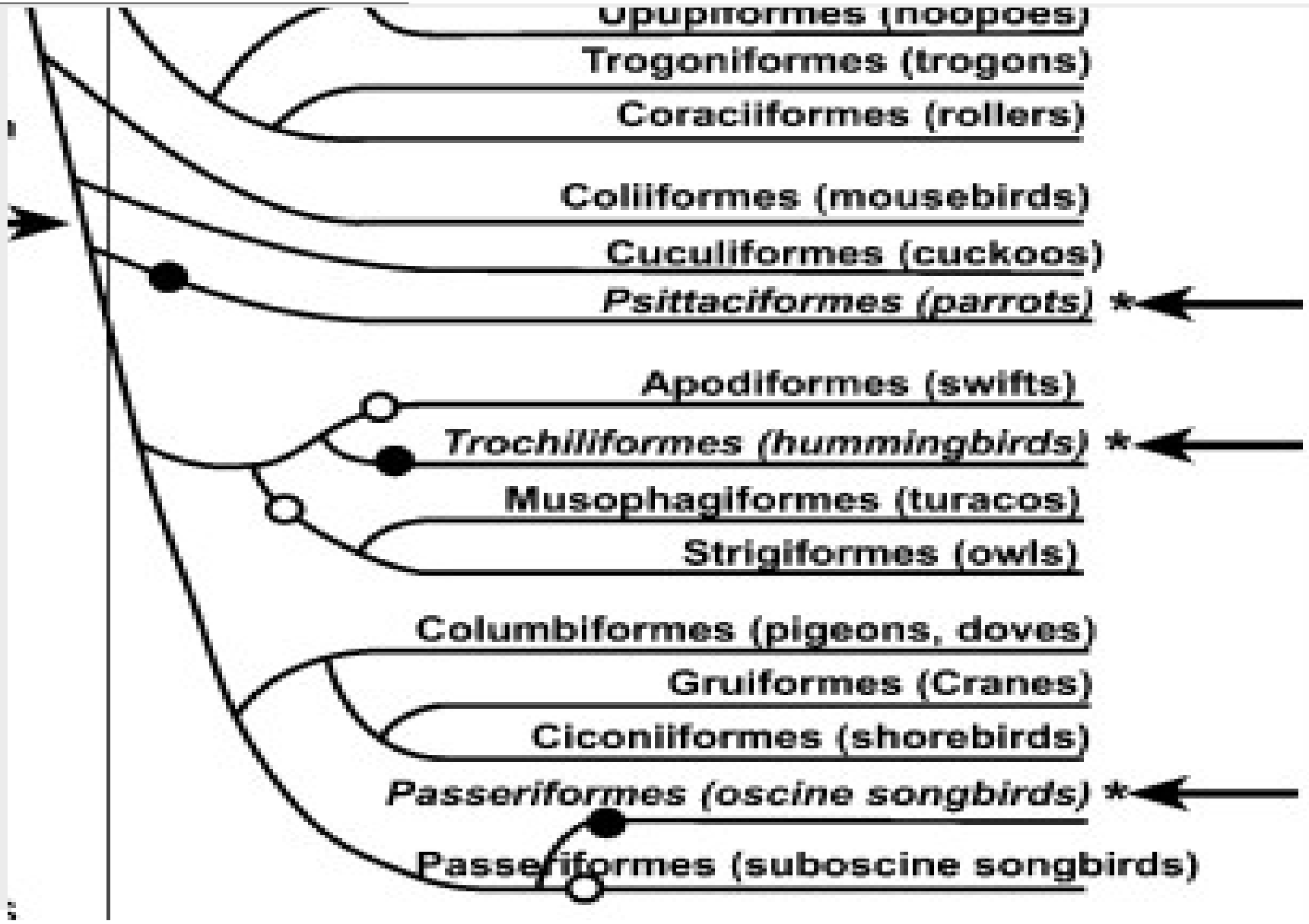
- Cultural transmission in (spoken) language depends on vocal learning – this might be a key ingredient
- We cannot count on simple, one-to-one correspondences between features of language and human learning biases;
- Need to study learning biases in their own right, through **Artifial Language Learning**, and study
  - the complex relation with features of language (modelling!)
  - the possibility that they are uniquely human (experiments!)

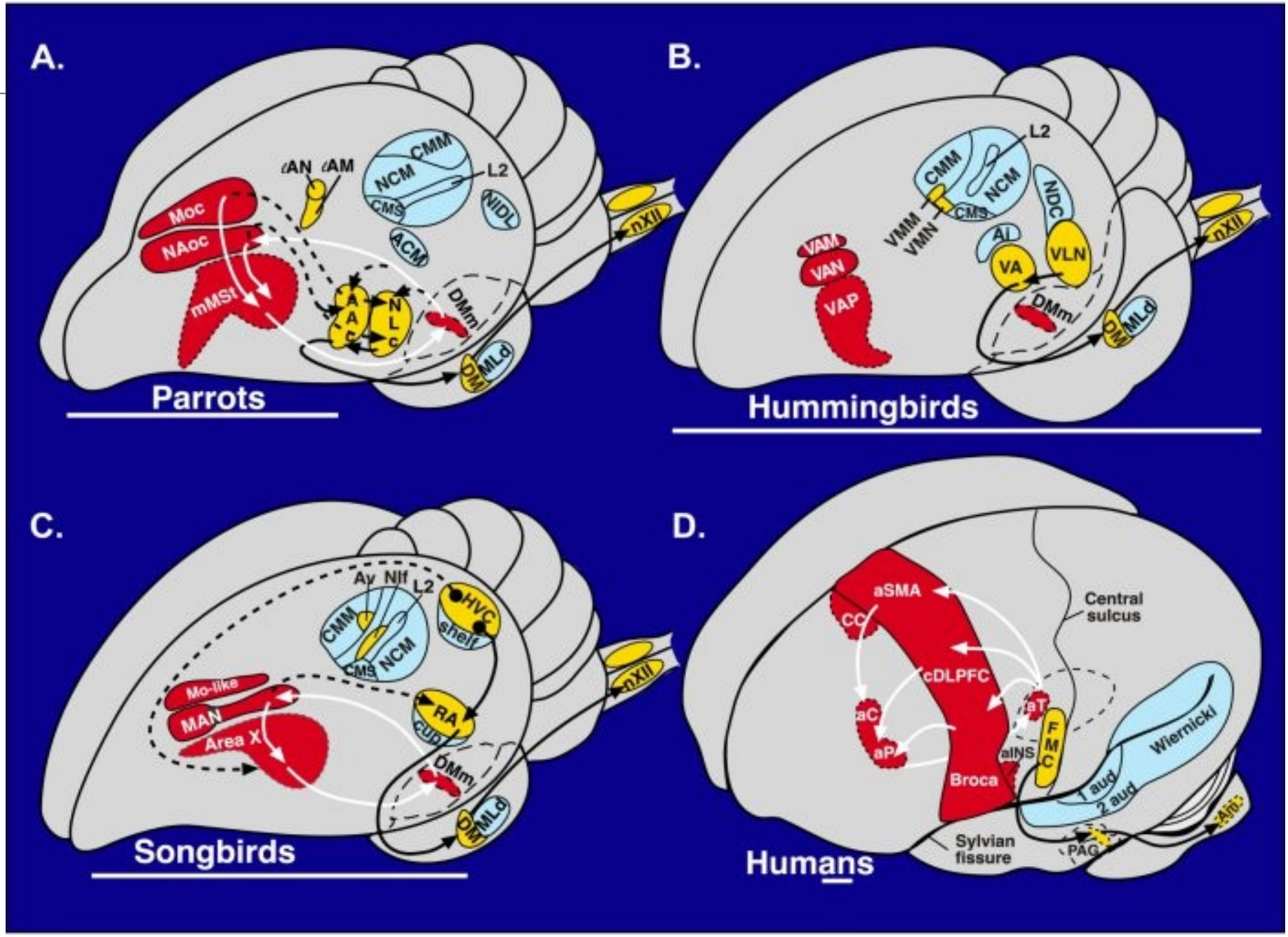

Mammalian Family Tree
140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
MYA
Montremata (platypus)
Marsupalia (kangaroos)
Xenarthra (anteaters)
Pholidota (pangolins)
Lagomorpha (rabbits)
Rodentia (Rats)
Macroscelidea (elephant shrews)
Primates (monkeys, humans) *
Possible common vocal learning ancestor
Scandentia (tree shrews)
Dermoptera (flying lemurs)
Chiroptera (bats) *
Insectivora (hedgehogs)
Carnivora (cats, dogs, seals)
Artiodactyla (pigs)
Cetacea (dolphins) *
Tubulidentata (aardvarks)
Perrissodactyla (horses)
Hyracoidea (hyraxes)
Proboscidea (elephants)
Sirenia (manatees)
Avian Family Tree
100 90 80 70 60 50 40 30 20 10 0
Struthioniformes (ostrich)
Tinamiformes (tinamous)
Craciformes (scrubfowl)
Galliformes (chickens)
Anseriformes (ducks)
Turniciformes (buttonquails)
Piciformes (woodpeckers)
Galbuliformes (puffbirds)
Bucerotiformes (nornbills)
Upupiformes (hoopoes)
Trogoniformes (trogons)
Coraciiformes (rollers)
Possible common vocal learning ancestor
Coliiformes (mousebirds)
Cuculiformes (cuckoos)
Psittaciformes (parrots) *
Apodiformes (swifts)
Trochiliformes (hummingbirds) *
Musophagiformes (turacos)
Strigiformes (owls)
Columbiformes (pigeons, doves)
Gruiformes (Cranes)
Ciconiiformes (shorebirds)
Passeriformes (oscine songbirds) *
Passeriformes (suboscine songbirds)
Jarvis, 2004
*Vocal Learners
Independent gains
Independent losses
- Everybody has it to varying degrees


Possible common
vocal learning
ancestor
Rodei
Macroscelidea (elephar
Primates (monkeys,
Scandentia (tre
Dermoptera (flyin
Chiropi
Insectivora (he

Jarvis, 2004


Mammalian Family Tree
140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
MYA
Montremata (platypus)
Marsupalia (kangaroos)
Xenarthra (anteaters)
Pholidota (pangolins)
Lagomorpha (rabbits)
Rodentia (Rats)
Macroscelidea (elephant shrews)
Primates (monkeys, humans) *
Scandentia (tree shrews)
Dermoptera (flying lemurs)
Chiroptera (bats) *
Insectivora (hedgehogs)
Carnivora (cats, dogs, seals)
Artiodactyla (pigs)
Cetacea (dolphins) *
Tubulidentata (aardvarks)
Perrissodactyla (horses)
Hyracoidea (hyraxes)
Proboscidea (elephants)
Sirenia (manatees)
Possible common vocal learning ancestor
Avian Family Tree
100 90 80 70 60 50 40 30 20 10 0
Struthioniformes (ostrich)
Tinamiformes (tinamous)
Craciformes (scrubfowl)
Galliformes (chickens)
Anseriformes (ducks)
Turniciformes (buttonquails)
Piciformes (woodpeckers)
Galbuliformes (puffbirds)
Bucerotiformes (nornbills)
Upupiformes (hoopoes)
Trogoniformes (trogons)
Coraciiformes (rollers)
Coliiformes (mousebirds)
Cuculiformes (cuckoos)
Psittaciformes (parrots) *
Apodiformes (swifts)
Trochiliformes (hummingbirds) *
Musophagiformes (turacos)
Strigiformes (owls)
Columbiformes (pigeons, doves)
Gruiformes (Cranes)
Ciconiiformes (shorebirds)
Passeriformes (oscine songbirds) *
Passeriformes (suboscine songbirds)
Possible common vocal learning ancestor
* Vocal Learners
● Independent gains
○ Independent losses
- Everybody has it to varying degrees
Jarvis, 2004


Upupiformes (hoopoes)
Trogoniformes (trogons)
Coraciiformes (rollers)
Coliiformes (mousebirds)
Cuculiformes (cuckoos)
Psittaciformes (parrots) *
Apodiformes (swifts)
Trochiliformes (hummingbirds) *
Musophagiformes (turacos)
Strigiformes (owls)
Columbiformes (pigeons, doves)
Gruiformes (Cranes)
Ciconiiformes (shorebirds)
Passeriformes (oscine songbirds) *
Passeriformes (suboscine songbirds)


A.
AN
AM
CMM
NCM
L2
CMS
NIDL
ACM
Moc
NAoc
mMSt
A
A
C
N
L
C
DMm
DM
MLd
nXII
Parrots
B.
L2
CMM
NCM
CMS
VMM
VMN
NDC
AL
VA
VLN
VAM
VAN
VAP
DMm
DM
MLd
nXII
Hummingbirds
C.
Av
Nif
L2
CMM
NCM
CMS
HVC
shelf
Mo-like
MAN
Area X
RA
cup
DMm
DM
MLd
nXII
Songbirds
D.
aSMA
CC
Central
sulcus
cDLPFC
aC
aP
aINS
Broca
FMC
Wiernicki
1 aud
2 aud
Sylvian
fissure
PAG
Humans

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*

le di di

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*

fi je je

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*

je je di

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*

di le le

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*

Comparative biology of syntax?

- Artificial Language Learning experiments in adults, infants, apes, monkeys and birds
- E.g., Marcus et al. 1999 *Science*
- Marcus et al. 2007, *PsychSci*
  - 7.5 month old children can do this only for speech stimuli; they fail on tones, pictures, timbres, animal sounds
  - Older children can do it in any domain
  - 7.5 month old succeed when first familiarized with speech stimuli
- Hauser et al. '02: monkeys can do it. RETRACTED!

Segments & rules project

- ILLC (UvA), Behavioral Biology (Leiden), Linguistics (Leiden)
- Funded by NWO-GW (2012-2016)
- Three PhD students
- Artificial Language Learning experiments with
  - infants
  - adults
  - zebrafinches
- Computer modelling and analysis of the data

Modelling Artificial Language Learning

pulikiberagatafodupuraki

(Pena et al, 2002)

Modelling Artificial Language Learning

pulikiberagatafodupuraki

Modelling Artificial Language Learning

pulikiberagatafodupuraki

pufoki kibera

Conclusions

- Language is special:
  - Symbolism, cultural transmission
  - Compositional semantics
  - Recursive, Hierarchische Phrasestructure

- Artificial language learning experiments
  - Only simple patterns, eschew full complexity of language
  - Can be applied to adults and infants, humans and animals
  - Have the potential to reveal uniquely human learning biases

- Thanks to colleagues in the Segments & rules project:
  - Carel ten Cate, Michelle Spiering (Leiden, Biology)
  - Claartje Levelt, Andrea Geambasu (Leiden, Linguistics)
  - Raquel Alhama (ILLC, UvA)

- Contact details:
  - staff.science.uva.nl/~jzuidema
  - zuidema@uva.nl

Next week

- Make sure you own the book, and start reading!
- This week: chapters 1 and 4
- Visit the Blackboard forum to post questions and answers;
- I will now and then answer questions, and post secret tips about sections you may skip.