Singapore, NTU, 8 November 2013

Linguistic descriptive bias and the nature of grammar

Rik De Busser
rdbusser@nccu.edu.tw

Part 1:
The problem of bias

Bias

- Bias and linguistic description
- Some illustrations:
  - Selection bias: Dutch causative constructions
  - Confirmation bias: Takivatan Bunun argument alignment
- Implications

Bias

- Is pervasive in research and human cognition
- Is not necessarily harmful
- Can “lead to severe and systematic errors” (Tversky & Kahneman 1982: 3)
- Incidental vs. systematic bias

Bias

- Systematic bias can introduce patterns in the data that are easily interpreted as meaningful
- Bias is not necessarily the result of:
  - Stupidity
  - Negligence
  - Malice
  - Ignorance

Bias

- Received considerable attention in:
  - Psychology
  - Statistics
  - Epidemiology and clinical studies

- How many studies on methodological bias in linguistics are you aware of?

Dutch causatives

- General picture: two causative verbs
  - *doen* ‘do’: direct causation
  - *laten* ‘let’: indirect causation

Verhagen & Kemmer (1997)

Coppen et al. (2007), *ANS*

Dutch causatives

– *Doen* ‘do’: Causer has a tendency to be inanimate (58%)

(1)	*de*	*stralen-de*	*zon*	***doe-t***	*de*	*temperatuur*	*oplop-en*
	the	shine-ADJR	sun	do.PRES-3S	the	temperature	rise-INF

‘The bright sun **makes** the temperature rise.’ (V&K)

– *Laten* ‘let’: Causer is typically animate (99%)

(2)	*de*	*sergeant*	***liet***	*ons*	*door*	*de*	*modder*	*kruip-en*
	the	sergeant	let.PST.S	us.ACC	through	the	mud	crawl-INF

‘The sergeant **had/made** us crawl through the mud.’ (V&K)

Dutch causatives

- The problem: other constructions with causative-like semantics
  - *Maken* ‘make’

(3)

hij	***maakte***	*me*	*nerveus*
3S.NOM	make-PST.S	1S.ACC	nervous

‘He **made** me nervous’ (fv800876)

(4)

...	*ze*	***maakte***	*me*	*ook*	*aan*	*het*	*lachen*
	3S.F.NOM	make-PST.S	1S.ACC	also	at	the.N	laugh-INF

‘she also **made** me laugh.’ (fv800706)

Dutch causatives

- The problem: other constructions with causative-like semantics
  - *Geven* ‘give’

(5)

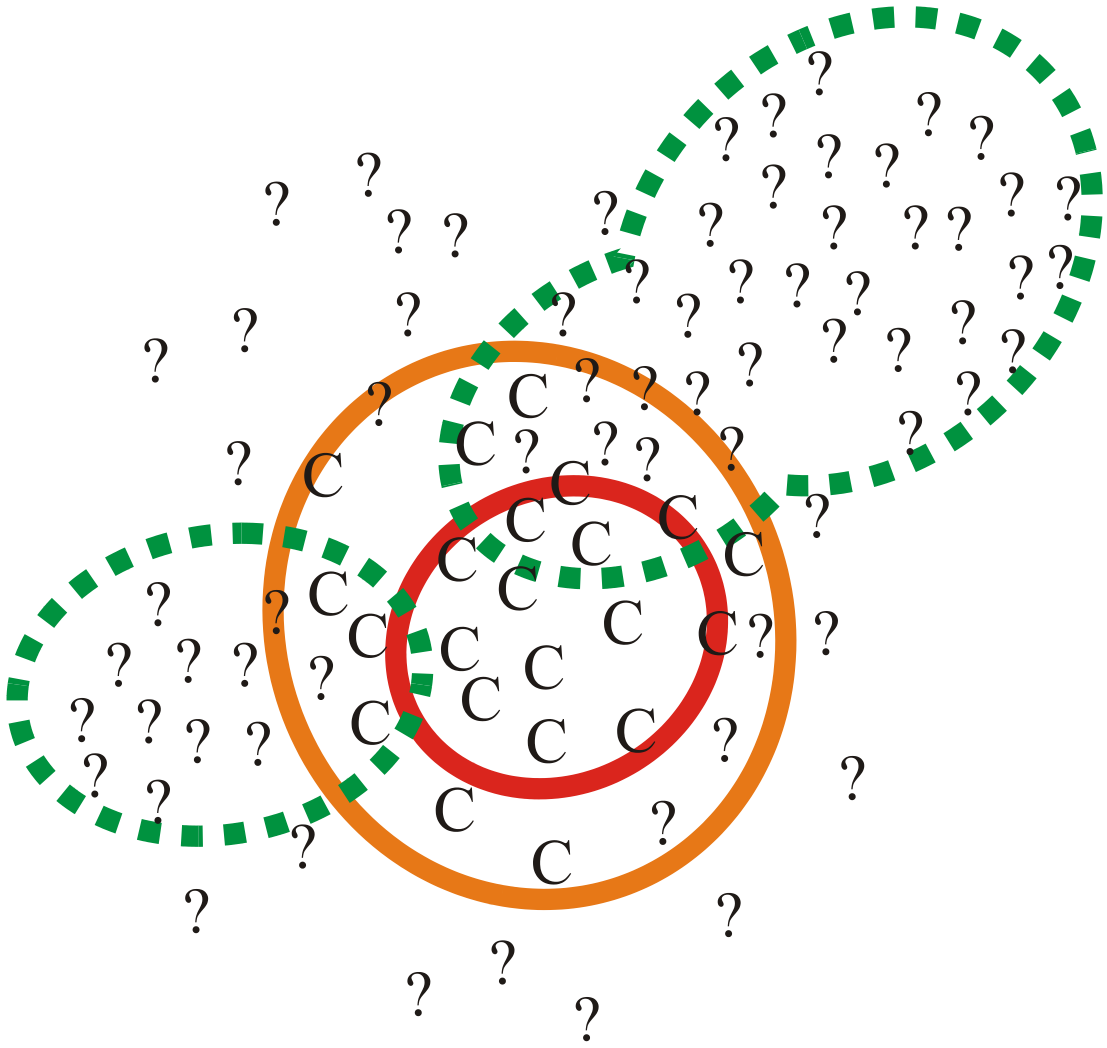
Ø	***geef***	*me*	*gras*	*te*	*eten.*
	give	1S.NOM	grass	PRT	eat.INF

‘… **make** me eat grass.’ (fv800618)

(6)

…	***geef***	*ons*	*iets*	*te*	*doen...*
	give	1P.ACC	something	PRT	do.INF

‘[If You have special wishes,] **let** us know it ...’ (internet)



Dutch causatives

- Why are these ‘prototypical’ causatives more interesting for linguistic description?
- Why are certain instances considered atypical?

Dutch causatives

- Because we believe there is a group of ‘causative’ constructions that is somehow theoretically priviliged
- A priori theoretical bias
  - Retrievability / imaginability

    (Tversky & Kahneman 1982: 11ff)
  - Negative bias
  - Selection bias

Bunun argument structure

- Bunun, Austronesian, Taiwan
  - Takivatan dialect
- Predicate-initial
- Complex verbal morphology
- Philippine-type voice system
  - ‘focus’ (≠ pragmatic focus)
  - Argument alignment system

Bunun argument structure

- Verbal suffixes:
  - "Focus" / role alignment (AF/UF/LF)

(1) *na-ma-tasʔi-**Ø**-ʔak* *busul*
IRR-DYN-build-**AF**-1S.TOP gun
'I make a gun'

(2) ... *na* *pa-tasʔi-**un***
so CAUS.DYN-build-**UF**
'(The thing is broken,) so I want to have it fixed.'

(3) *pa-tasʔi-an*
CAUS.DYN-build-**LF**
'I want to make it so that something stays in a fixed spot'

Bunun argument structure

- Verbal prefixes (I):
  - Participant orientation (BEN/INSTR/…)

(4) ***ki-****saiv-ʔak* *qaimaŋsuð*
BEN-give-1S.TOP thing
‘Somebody has to give me things.’

(5) ***sin-****su-suað* *bunuað*
RES.OBJ-REP-grow plum
‘They had grown plums.’
(Indicates that the plums are already on the tree)

Bunun argument structure

- Verbal prefixes (II):
  - Internal temporal structure

(7) ***ma-****baliv-ʔak* *iðuq* *a* ***min-****puhuq*
DYN-buy-1S.F orange LNK **INCH-**rot
‘I bought meat that had become rotten.’

(8) *nitu* ***ma-****naskal* *sadu-ki* *uskun-an*
NEG **STAT-**happy see-DEF.SIT.PROX together-LO
‘I was not happy to see my companions do it like this.’

Bunun argument structure

- Verbal prefixes (III):
  - Control (internal/external/joint)

(6) ***pa-***tasʔi-un
CAUS.DYN-make-UF
‘I will have it fixed (by someone else).’

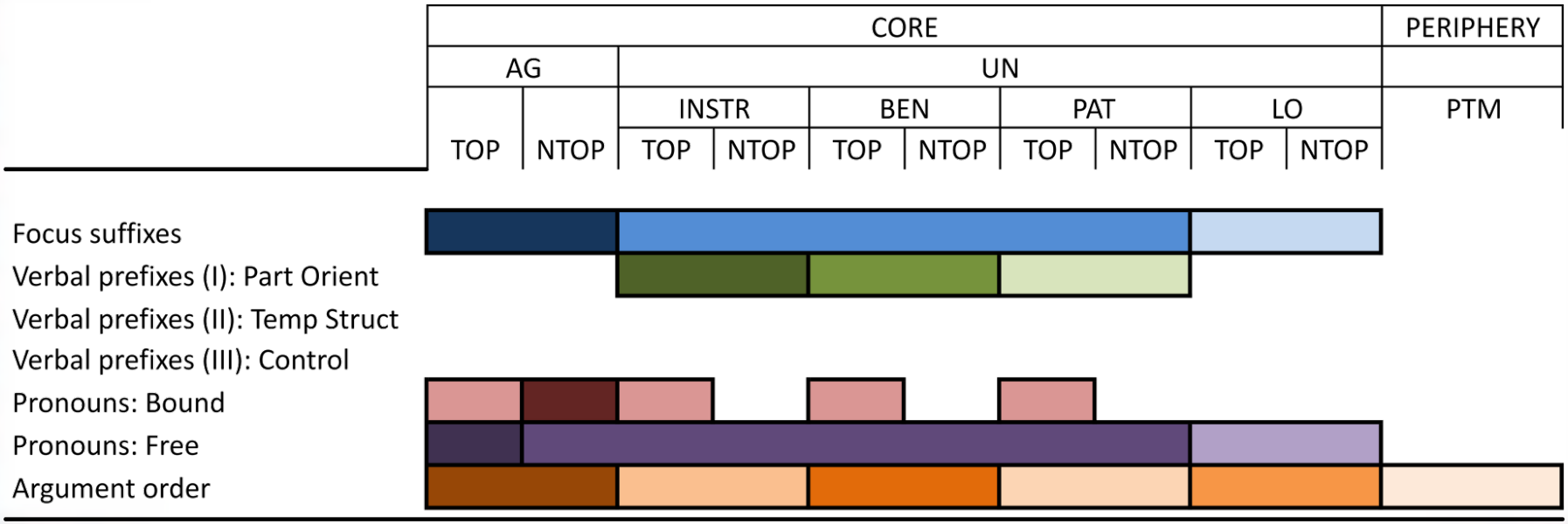
(7) ***ka-****daŋað* *baðbað*
ASSOC.DYN-help have.conversation
‘I’ll help you talk (by speaking in your place).

Bunun argument structure

- Personal pronouns

	Bound		Free	
	Topic (TOP)	Non-topical agent (NTOP.AG)	Neutral (N)	Topical agent (TOP.AG)
1S	*-(ʔ)ak*	*-(ʔ)uk*	*ðaku, nak*	*sak, saikin*
2S	*-(ʔ)as*	—	*suʔu, su*	—
1I	—	—	*mita*	*ʔata, inʔata*
1E	*-(ʔ)am*	—	*ðami, nam*	*ðamu, sam*
2P	*-(ʔ)am*	—	*muʔu, mu*	*amu*

Bunun argument structure



- Different subsystems, different grammatical distinctions

Bunun argument structure

- No single internally consistent argument alignment system
- Transitivity is at best epiphenomenal
- No distinctions corresponding to traditional argument alignment systems (NOM-ACC or ERG-ABS)

Bunun argument structure

- Why do researchers tend to analyse Philippine-type argument alignment as a coherent system?
  - Involving verbal prefixes, infixes, suffixes, reduplication, and nominal morphology
- Why is there a strong inclination to explain systems like this as irregular/unusual ergative alignment?

(e.g. Mithun 1994; Ross 2006)

Bunun argument structure

- A priori theoretical bias
  - Illusory correlation

    (Tversky & Kahneman 1982: 13-14)
  - Positive bias
  - Confirmation bias

Why should we care?

- What if you use this data?
- Method bias:

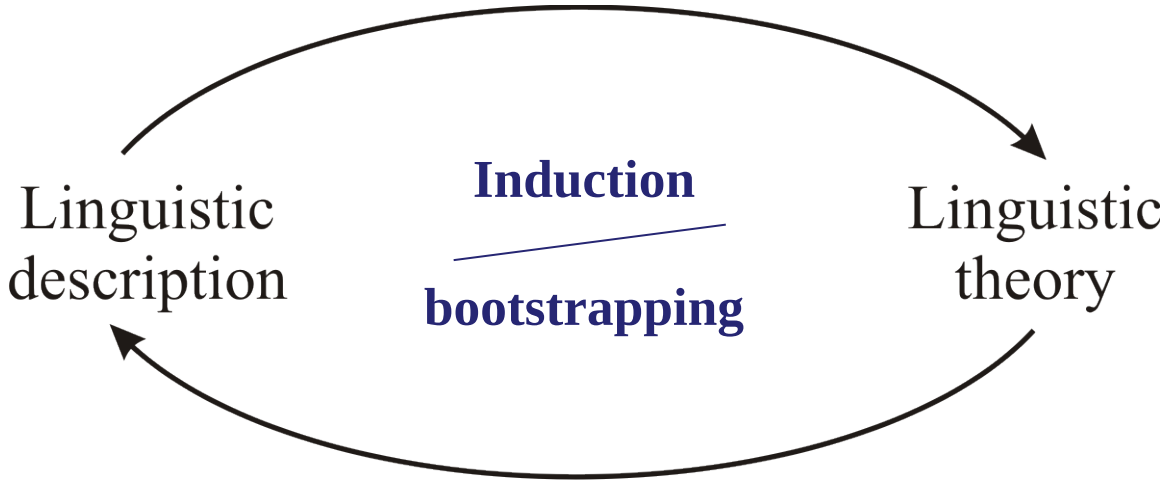
  "Method variance refers to variance that is attributable to the measurement method rather than to the construct of interest."

  (Podsakoff & al 2003 quoting Bagozzi & Yi 1991)

  – How can research based on biased descriptive data avoid drawing biased conclusions?

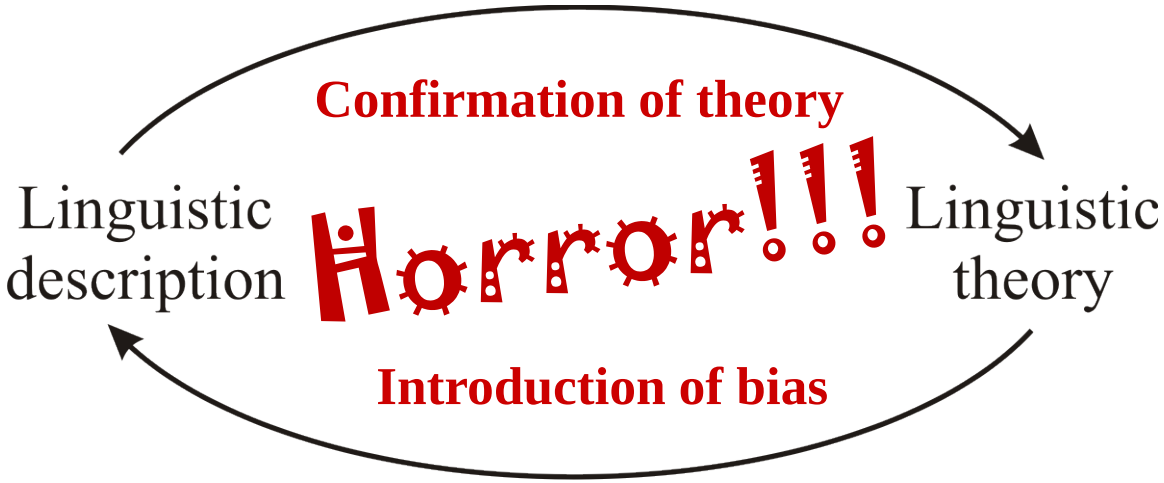
Why should we care?

- Negative effect on comparative research making use of this type of data
- Confirmation of established theories based on method-induced correlations



Why should we care?

- Negative effect on comparative research making use of this type of data
- Confirmation of established theories based on method-induced correlations



What now?

- Don’t panic
  - Bias is unavoidable
  - Bias is natural
    - “people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations”

(Tversky & Kahnemann 1982)

What now?

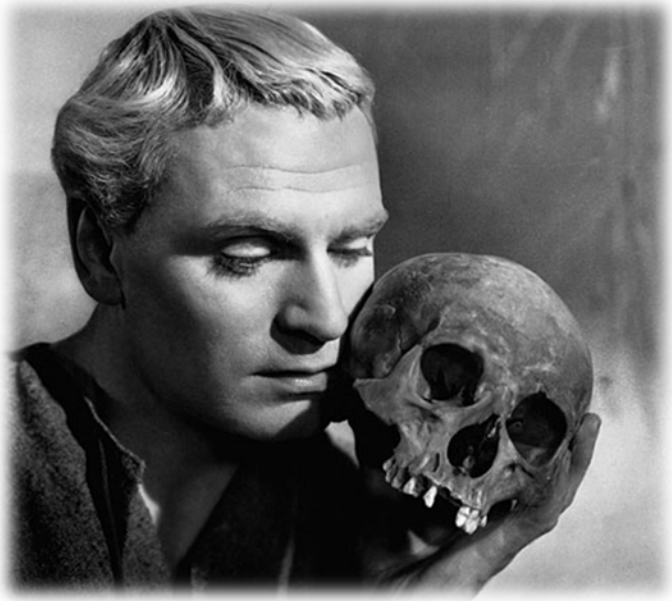
- Awareness and proper appreciation of the problem
- Research into bias and bias reduction in linguistics
- Value of theoretical independence in linguistic description
- Research into incoherence (or even chaos) in linguistic structure

Intermezzo: Preconceptions and grammar

Preconceptions

There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.

Shakespeare, *Hamlet*



Preconceptions

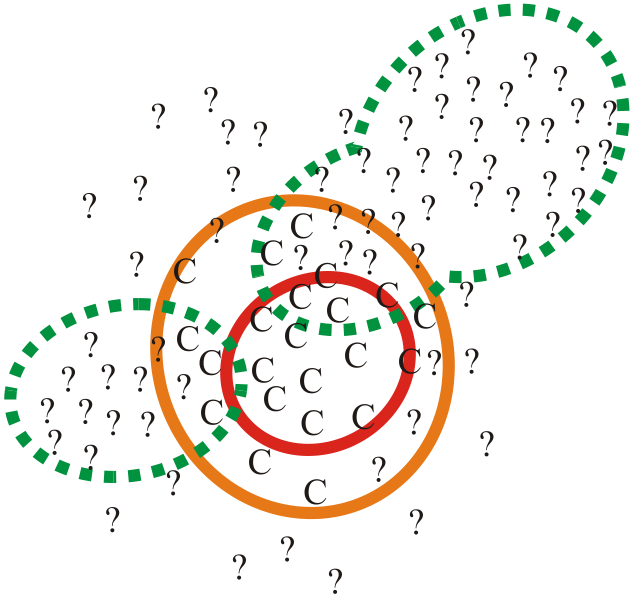
- Where does bias come from?
  - Cognitive / neuro-physiological restrictions
  - Environmentally induced
  - Theory-induced
  - …

Preconceptions

- Influences of theoretically induced bias on our understanding of language:
  1. Selection of evidence
     (e.g. Dutch causatives)
  2. Interpretation of evidence
     (e.g. Bunun pred-arg structure)

Preconceptions

- Questions
  - Are we missing something important?
  - What can we do about it?



Grammatical rules

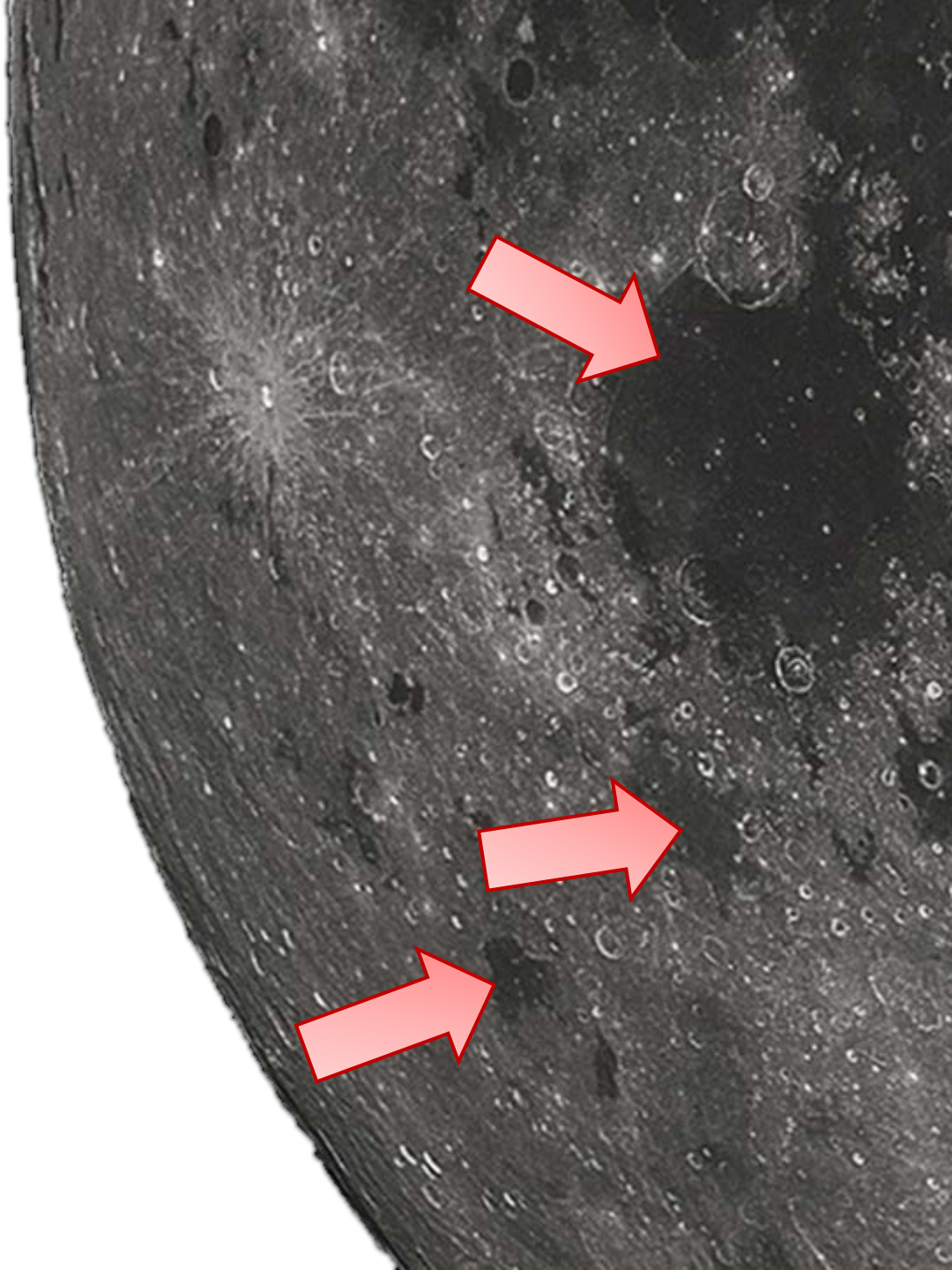
- Traditional view:
  - Grammar = rules about language
  - Language can be modeled in terms of automata theory

Grammatical rules	Language / cognition
discrete	fluidity (semantics, observation, …)
absolute	statistical uncertainty / approximation
synchronic	dynamic nature of language
levels of analysis	interactions between different levels

Selectiveness

- Possible solution: attractor network
- ‘basins of attraction’ (Bybee 2013)
  - “… represent the convergence of cognitive and communicative factors that frequently occur together in human experience.”
  - “ Like the craters on the moon, the basins are nearer or farther from one another according to semantic similarity. Some basins are larger: these represent more contexts of use …”

- What about the empty spaces between the basins?



Selectiveness

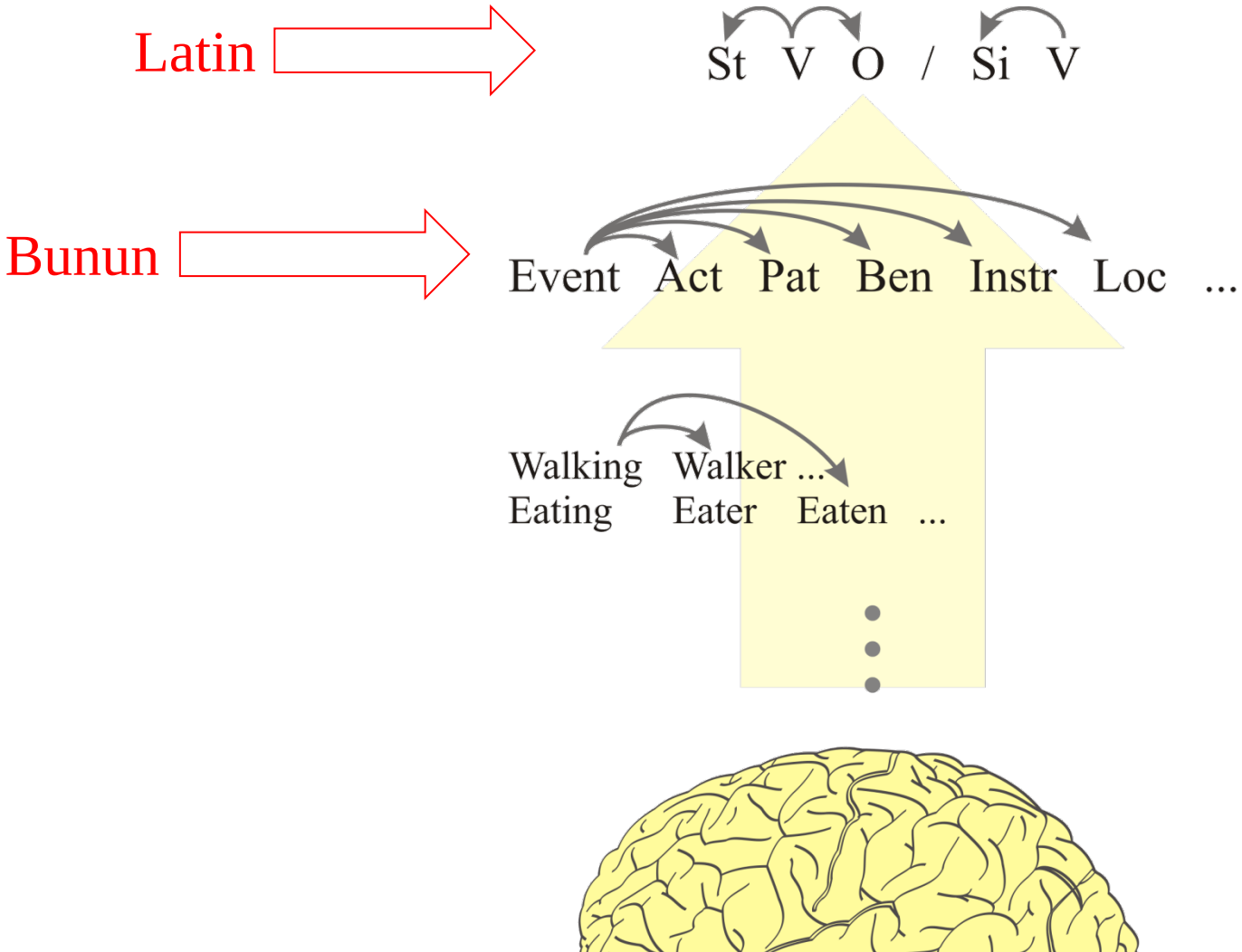
- The empty spaces
  - Does language only consist of craters?
  - Are uncommon or unnoticed phenomena less important to grammar?
  - To what extent is the distinction between relevant and irrelevant linguistic phenomena determined by theoretical considerations?
- Cf. “junk” DNA *(Pennisi 2012)*

Part 2:
A new model of grammatical structure

Construction grammar

- Language is a collection of constructions that represent habitualized linguistic behaviours
  - Usage-based
  - Emergence
  - Non-compositional model (cf. Croft 2001)
  - There are no real grammatical rules

Construction grammar



Construction grammar

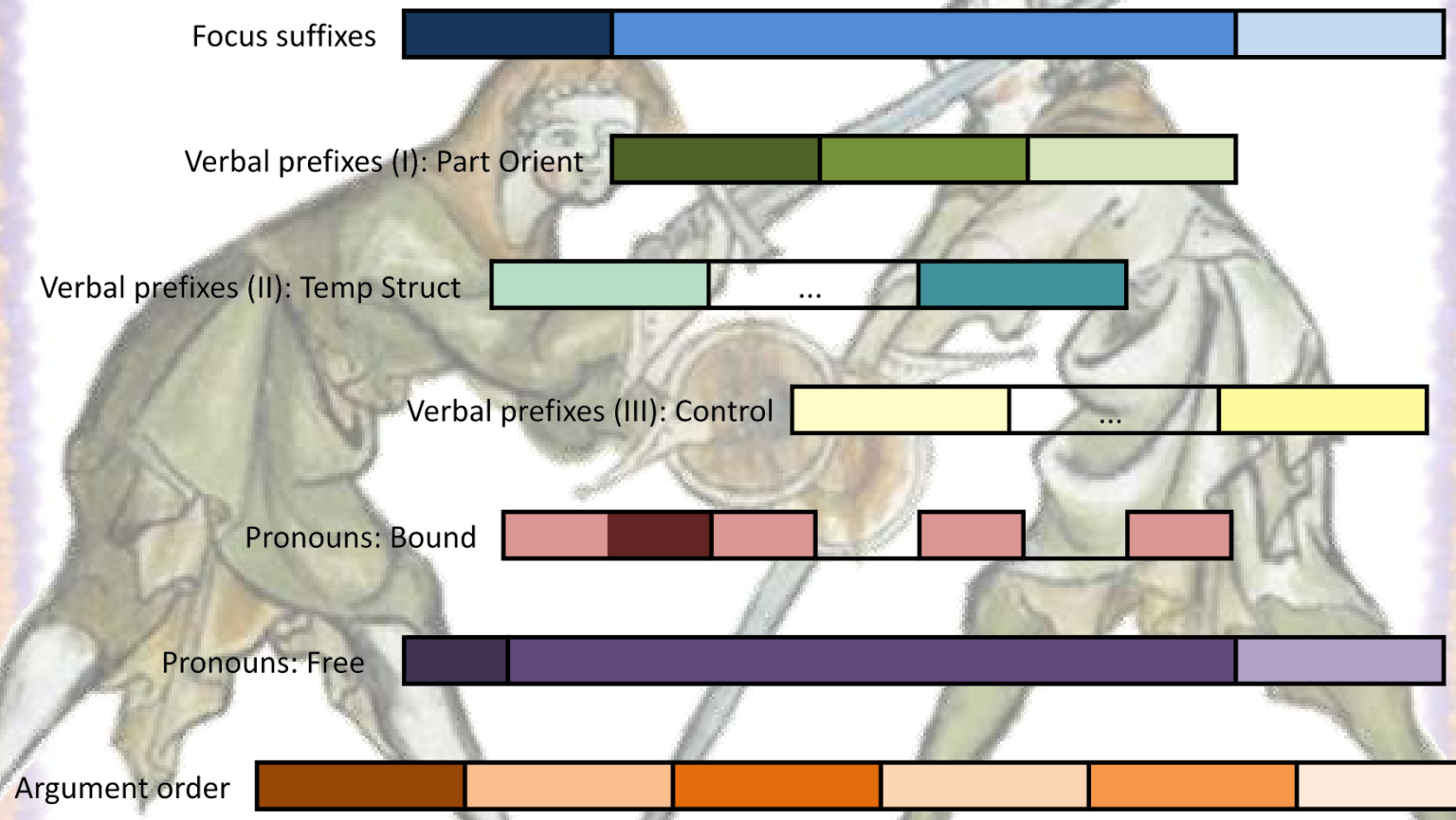
- Problem:
  - In its simplest form, construction grammar presupposes the existence of a single coherent system that generates and stores constructions
  - We saw that Takivatan predicate-argument structure consists of multiple subsystems that are partly clashing with each other.

Language as a complex adaptive system

- E.g. Beckner & al. 2009, Steels 2011

Grammatical rules	Language	CAS
discrete	fluidity (semantics, observation, …)	fluid
absolute	statistical uncertainty / approximation	probabilistic
synchronic	dynamic nature of language	inherently dynamic
levels of analysis	interactions between different levels	interaction between levels

Language as a complex adaptive system



Language as a complex adaptive system

"Language as a CAS involves the following key features: The system consists of **multiple agents** (the speakers in the speech community) **interacting** with one another. The system is **adaptive**, that is, speakers' behavior is based on their past interactions, and current and past interactions together feed forward into future behavior. A speaker's **behavior is the consequence of competing factors** ranging from perceptual constraints to social motivations. The structures of language **emerge** from interrelated patterns of experience, social interaction, and cognitive mechanisms."

(Beckner & al. 2009: 1-2)

Language as a complex adaptive system

- A CAS ≠
  - A complicated system
  - … with lots of participants
  - … that adapts to the environment

(It is much more specific than that)

- A CAS
  - Does not need to have (conscious) agents

Language as a complex adaptive system

- A CAS ≈
  - Decentralized: no central control
  - Connectivity: inter-relationships between elements
  - Dependence on initial conditions
  - Co-evolution: elements change behavior based on interactions with other elements and the environment
  - Emergence: order from randomness
  - Disequilibrium
  - Paradoxical: combination of order and chaos

(Chan 2001)

Language as a complex adaptive system

- … and some other properties
  - Multiplicity: Many interacting parts
  - Non-linear behavior
  - Hierarchical organisation
  - Modularity / specialisation: emergence of subsystems that fulfil specialised functions

Language as a complex adaptive system

- By itself, defining language as a complex adaptive system is not very informative!
  - We need to understand what it does and how it does it
  - We need a formal description of the system

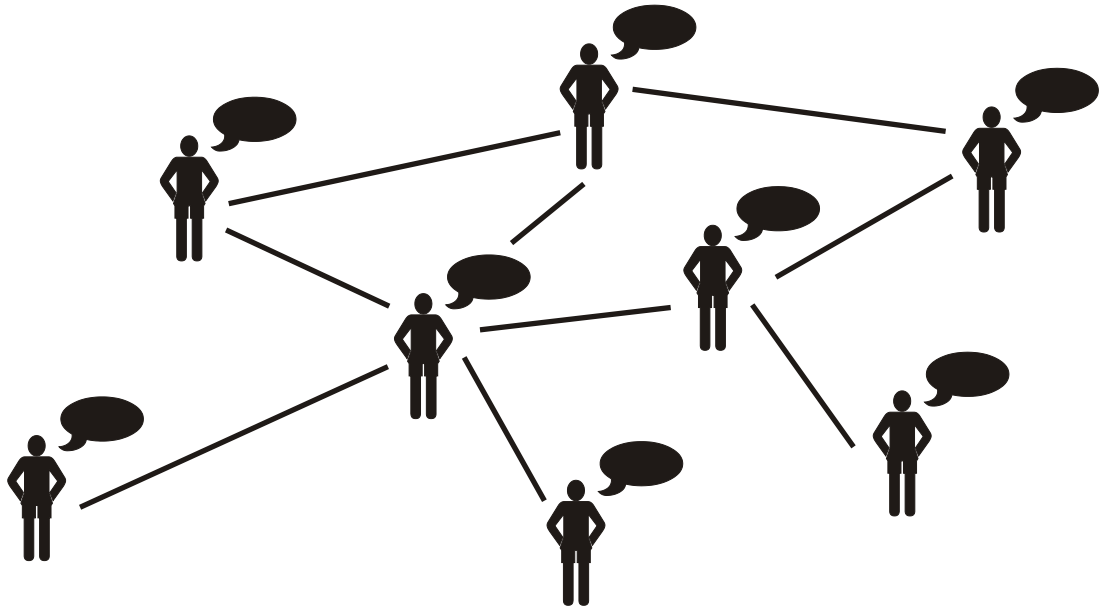
Questions

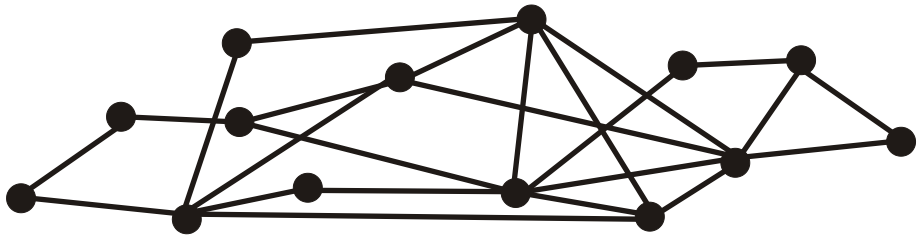
- What are the nodes in this system?
- What is the topology of the abstract space in which this system exists?
- What is the internal structure of this system?
- How does it evolve?

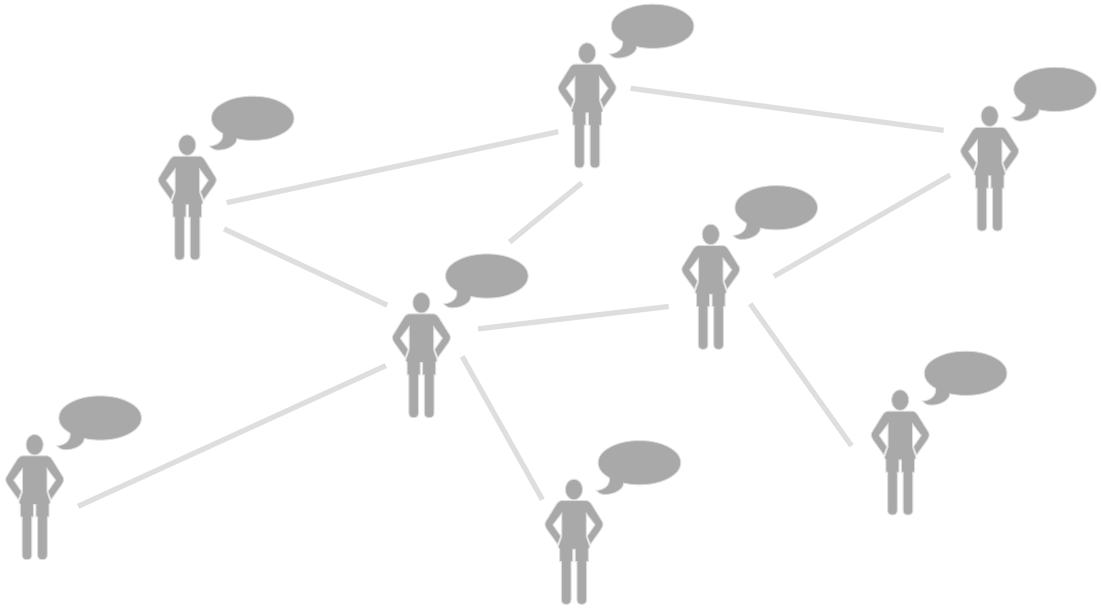
- (What kind of complexity measure?)

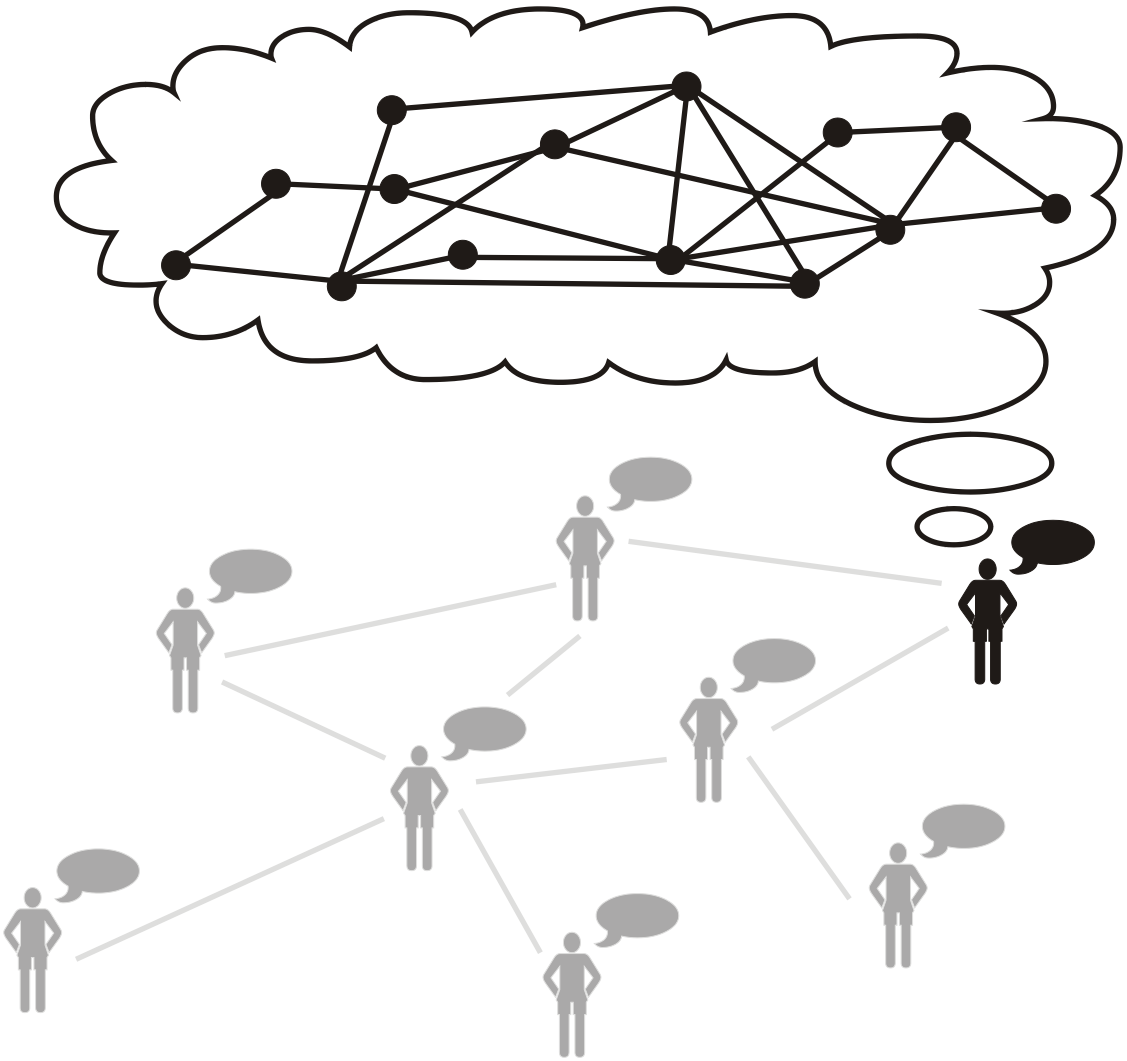
An initial hypothesis

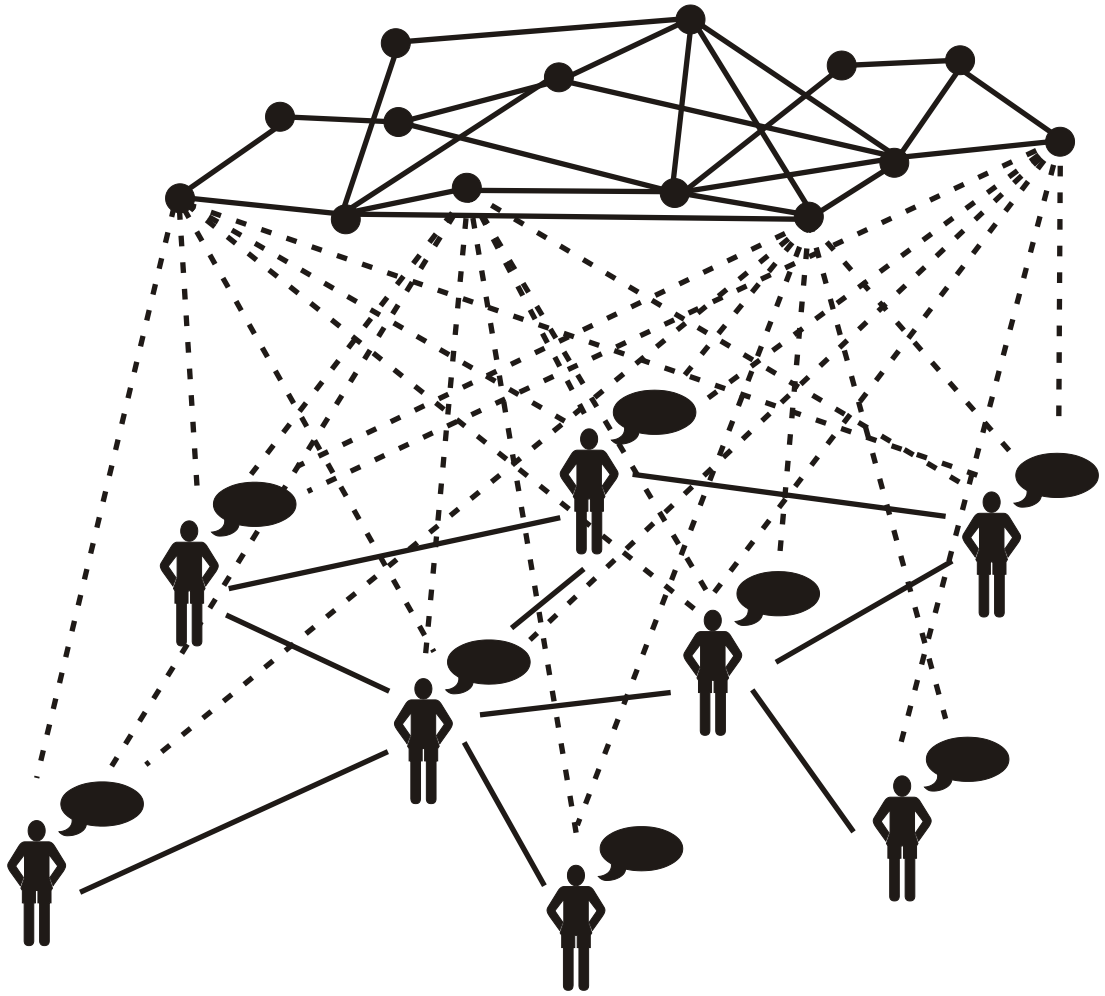
- Grammar is a hierarchical complex adaptive system in an abstract n-dimensional fitness landscape
  - Hierarchical modularity
  - Multi-dimensional abstract space
  - Nodes can be linguistic entities, not speakers





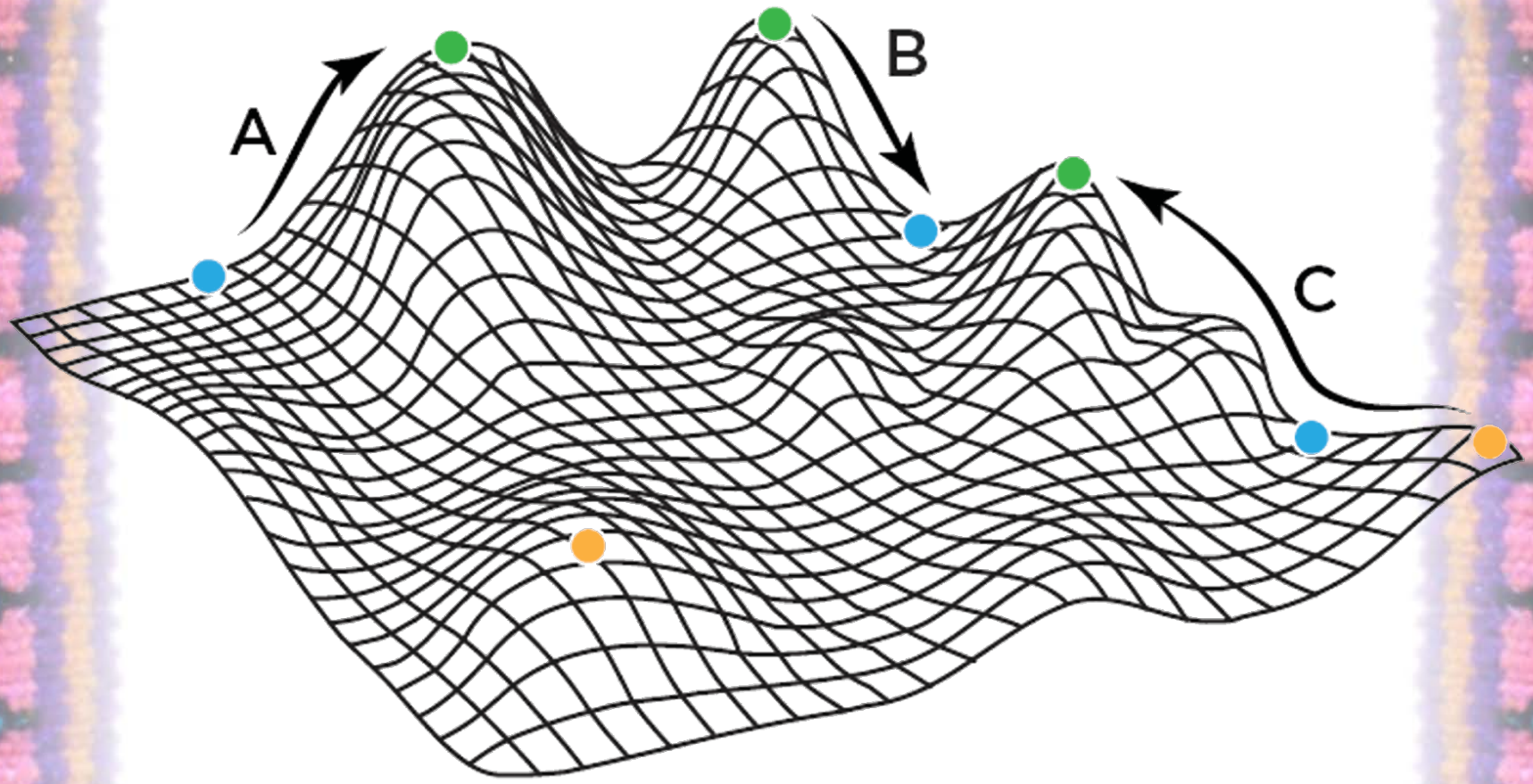






An initial hypothesis

- Grammar is a hierarchical complex adaptive system in an abstract n-dimensional fitness landscape
  - Competition for survival under environmental pressure (represented as elevations of the abstract space)
  - Environment = linguistic + non-linguistic



Afterthought: Language and evolution as a metaphor

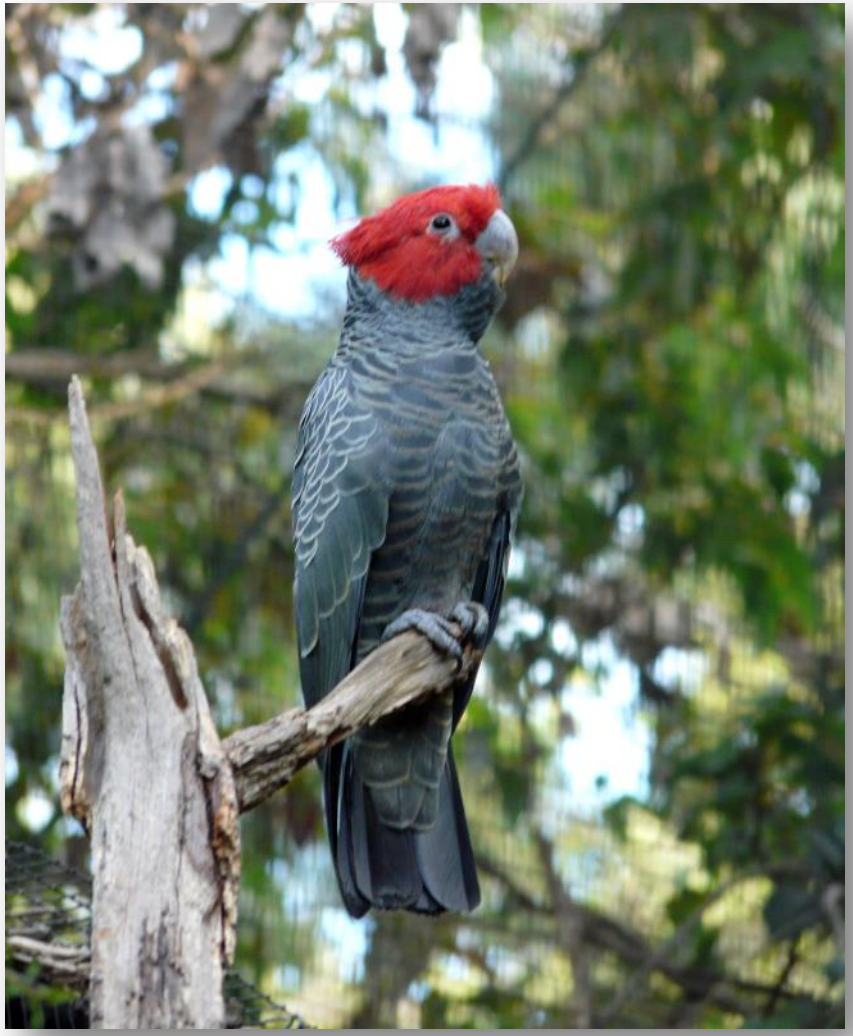
Evolutionary metaphors

- Language evolution
- 19th century metaphor:

> "Languages were born and died, like living organisms. They had their life spans, they grew and changed like men and animals, they had their little ills which could be cured by appropriate remedies prescribed by good grammarians."

Haugen (2001)

Evolutionary metaphors



Evolutionary metaphors

- Language is a superorganism

  (cf. Beckner & al. 2009)

- Language is like a bacterial colony

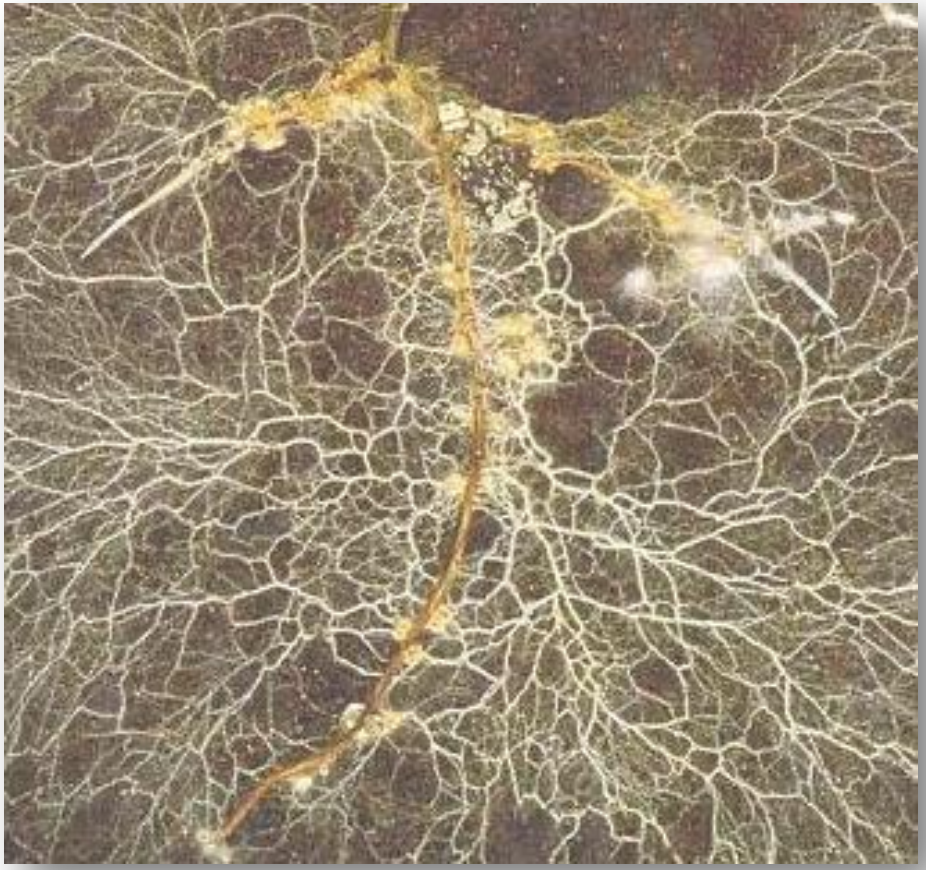
Evolutionary metaphors

- Language is like a mycelial network

"… mycorrhizal mycelia can also act as a conduit for signalling between plants, acting as an early warning system for herbivore attack."

(Babikova & al. 2003)

Natural metaphors



Mycorrhizal Fungi

Conclusion

- Bias should be avoided
- Understand bias
- Understand complexity
- Language has no rules
- Grammar ≈ mushrooms

Bibliography

Adami, Christoph. 2002. What is complexity? *BioEssays* 24(12). 1085–1094.

Babikova, Zdenka, Lucy Gilbert, Toby J. A. Bruce, Michael Birkett, John C. Caulfield, Christine Woodcock, John A. Pickett & David Johnson. 2013. Underground signals carried through common mycelial networks warn neighbouring plants of aphid attack. *Ecology Letters* 16(7). 835–843.

Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman & Tom Schoenemann. 2009. Language is a complex adaptive system: Position paper. In Nick C. Ellis & Diane Larsen-Freeman (eds.), *Language as a Complex Adaptive System*, 1–26. Ann Arbor, MI: Language Learning Research Club, University of Michigan.

Bybee, Joan. 2013. Looking back, looking forward: the emergence of grammatical meaning. Plenary presentation at the *12th International Cognitive Linguistics Conference (ICLC-12)*, University of Alberta, Edmonton, Canada. http://vimeo.com/74965255.

Bibliography

Chan, Serena. 2001. *Complex Adaptive Systems*. Course notes of the ESD.83 Research Seminar in Engineering Systems, Massachusetts Institute of Technology, 31 October 2001.

De Busser, Rik. 2011. Towards an analysis of argument alignment in Takivatan Bunun. *Studies in Language* 35(3): 523–555.

Coppen, Peter-Arno, Walter J. M. Haeseryn & Folkert de Vriend. 2007. *Electronische Algemene Nederlandse Spraakkunst (E-ANS), version 1.2*.

Croft, William. 2001. Radical Construction Grammar. Syntactic Theory in Typological Perspective. Oxford: Oxford University Press

Dixon, Robert M. W. 2010. *Basic Linguistic Theory, Volume 2: Grammatical Topics*. Oxford: Oxford University Press.

Haugen, Einar. 2001. The ecology of language. In Alwin Fill & Peter Mühlhäusler (eds.), *Ecolinguistics Reader: Language, Ecology and Environment*, 57–66. London: Continuum.

Bibliography

Mithun, Marianne. 1994. The implications of ergativity for a Philippine voice system. In Barbara Fox & Paul J. Hopper (eds.), *Voice: Form and Function*, 247–278. Amsterdam: John Benjamins.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee & Nathan P. Podsakoff. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88(5): 879–903.

Ross, Malcolm. 2006. The argument structure of undergoer voice clauses in Formosan and other Philippine-type languages. Talk at the *13th Annual Meeting of the Austronesian Formal Linguistics Association (AFLA 13)*, Hsinchu, Taiwan, 24-26 March 2006.

Sackett, David L. 1979. Bias in analytic research. *Journal of Chronic Diseases* 32 (1–2): 51–63.

Steels, Luc. 2011. Modeling the cultural evolution of language. *Physics of Life Reviews* 8(4). 339–356.

Bibliography

Tversky, Amos & Daniel Kahneman. 1982. Judgment under uncertainty: Heuristics and biases. In Daniel Kahneman, Paul Slovic & Amos Tversky (eds.), *Judgment under Uncertainty: Heuristics and Biases*, p. 3–20. Cambridge: Cambridge University Press.

Verhagen, Arie & Suzanne Kemmer. 1997. Interaction and causation: Causative constructions in modern standard Dutch. *Journal of Pragmatics* 27 (1): 61–82.

Uninaŋ miqumisaŋ!