What we don't see we see and don't see:

Confirmation bias in linguistic description

Rik De Busser
National Chengchi University
Taipei, Taiwan
rdbusser@nccu.edu.tw

What we don't see we see and don't see:

~~Confirmation~~ bias in linguistic *Methodological* description

Rik De Busser
National Chengchi University
Taipei, Taiwan
rdbusser@nccu.edu.tw

Introduction

- 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?

Ex. 1: Dutch causatives

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

Verhagen & Kemmer (1997)

Coppen et al. (2007), *ANS*

Ex. 1: 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)

Ex. 1: 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)

Ex. 1: Dutch causatives

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

(5)

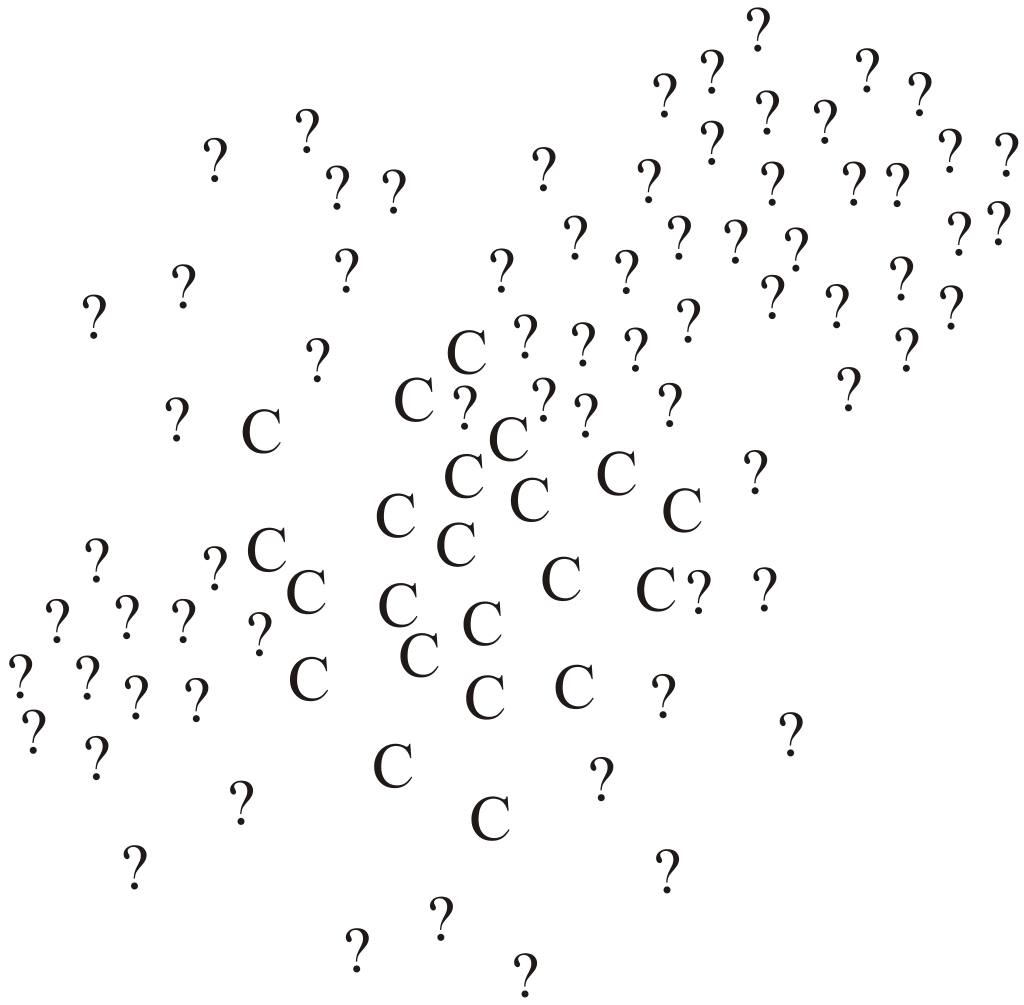
Ø	***geef***	*me*	*gras*	*te*	*eten.*
	give	1S.NOM	grass	PRT	eat.INF

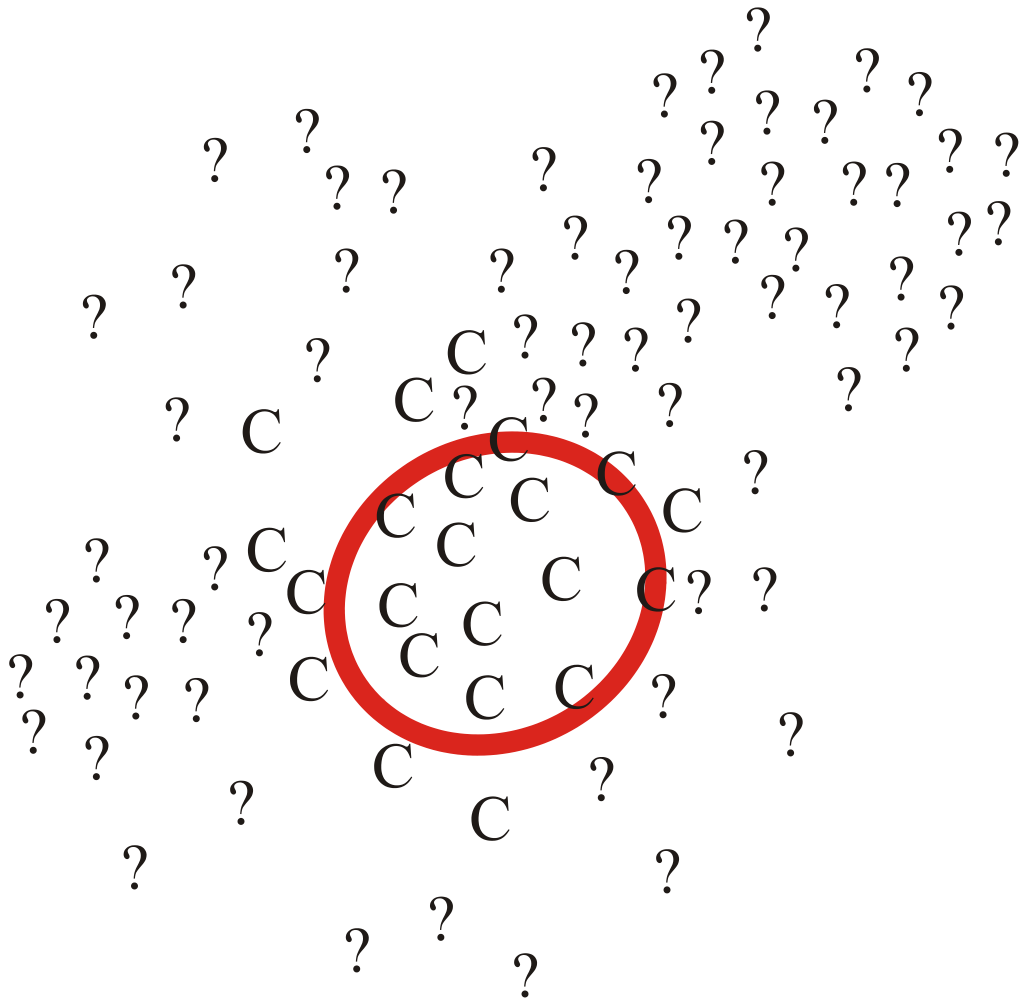
‘… **make** me eat grass.’ (fv800618)

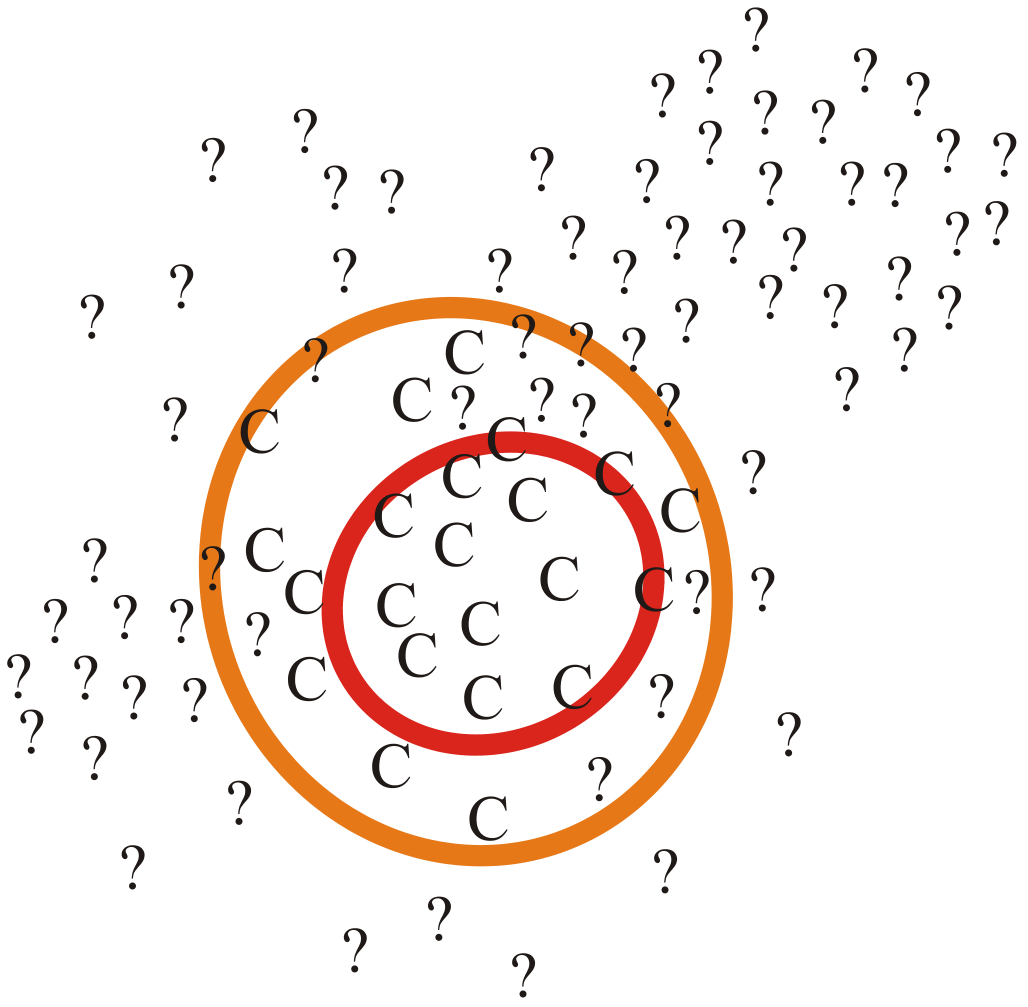
(6)

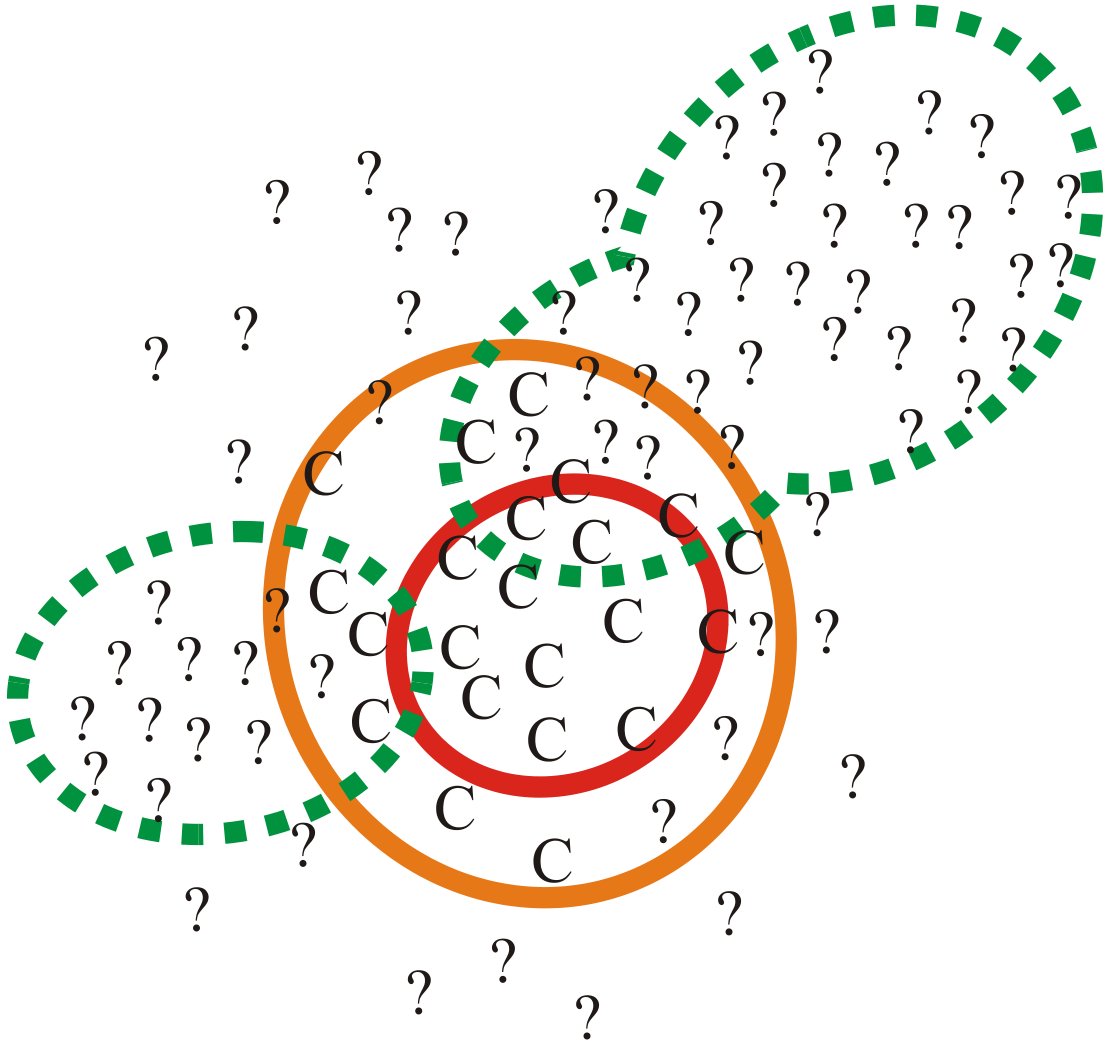
…	***geef***	*ons*	*iets*	*te*	*doen...*
	give	1P.ACC	something	PRT	do.INF

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






C
?



Ex. 1: Dutch causatives

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

Ex. 1: 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

Ex. 2: Bunun argument alignment

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

Ex. 2: Bunun argument alignment

- 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'

Ex. 2: Bunun argument alignment

- 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)

Ex. 2: Bunun argument alignment

- 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.’

Ex. 2: Bunun argument alignment

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

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

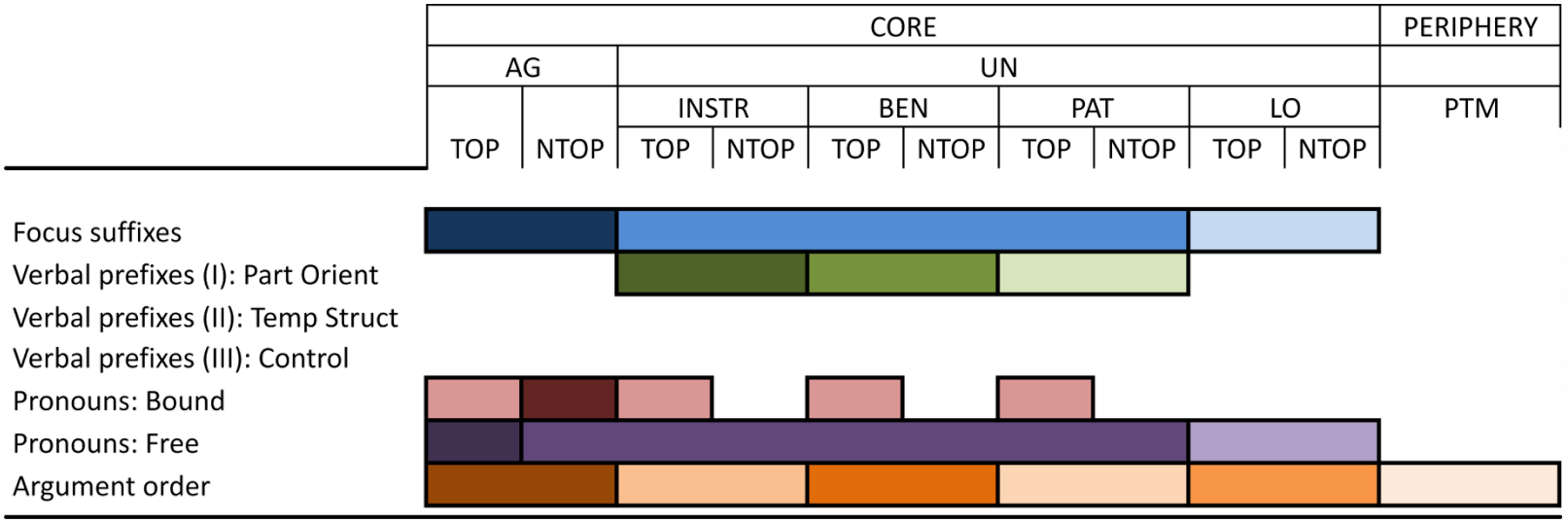
(7) ***ka-****daŋað* *baðbað*
ASSOC.DYN-help have.conversation
‘I’ll help you talk (by speaking in your place).

Ex. 2: Bunun argument alignment

- 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*

Ex. 2: Bunun argument alignment



- Different subsystems, different grammatical distinctions

Ex. 2: Bunun argument alignment

- 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)

Ex. 2: Bunun argument alignment

- 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)

Ex. 2: Bunun argument alignment

- 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:

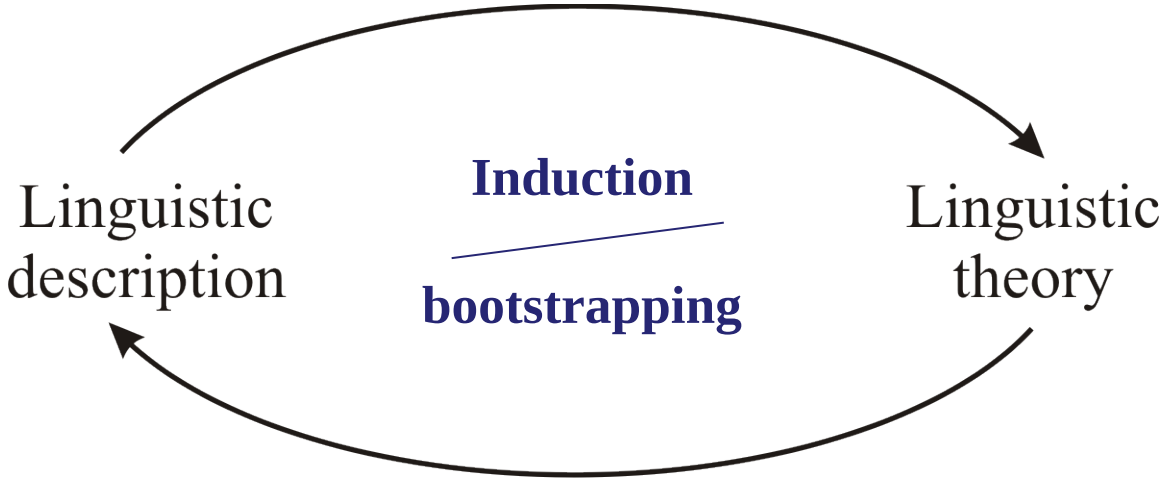
  "Method variance refers to variance that is attributable to the measurement method rather than to the construct of interest."

  (Podsakoff & al 2003 quoting Bagozzi & Yi 1991)

  – How can research based on biased descriptive data be certain that its conclusions are not due to bias?

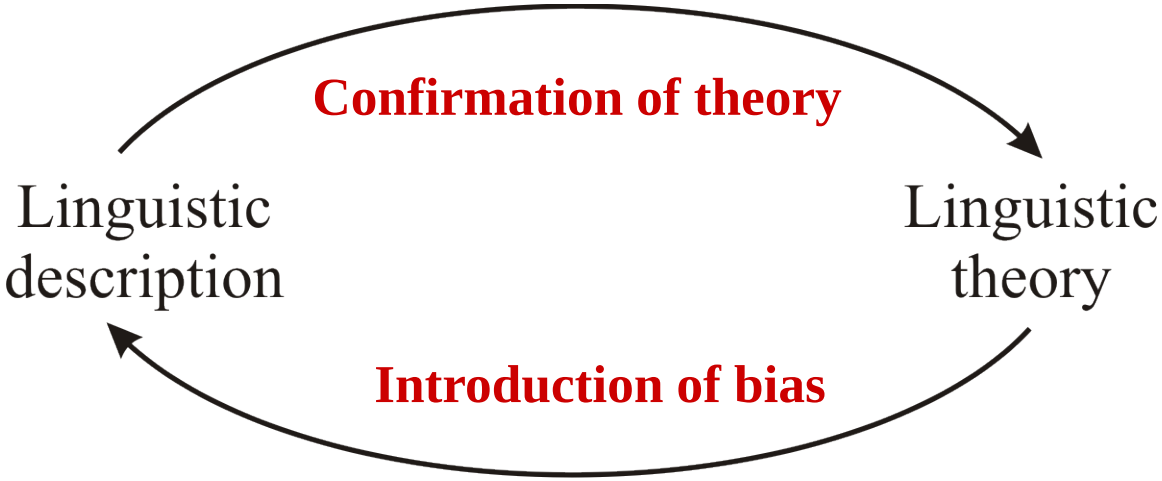
Why should we care?

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



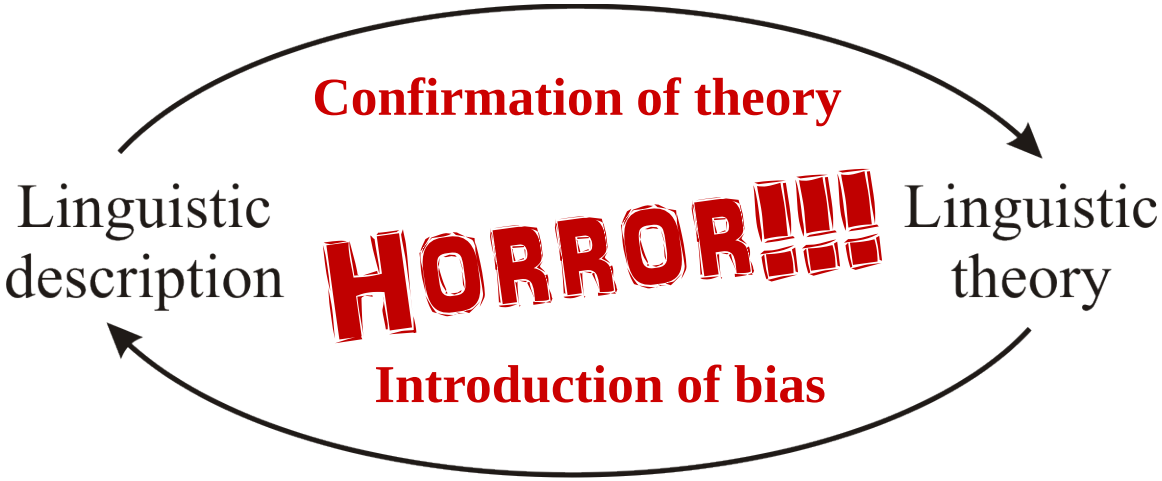
Why should we care?

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



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



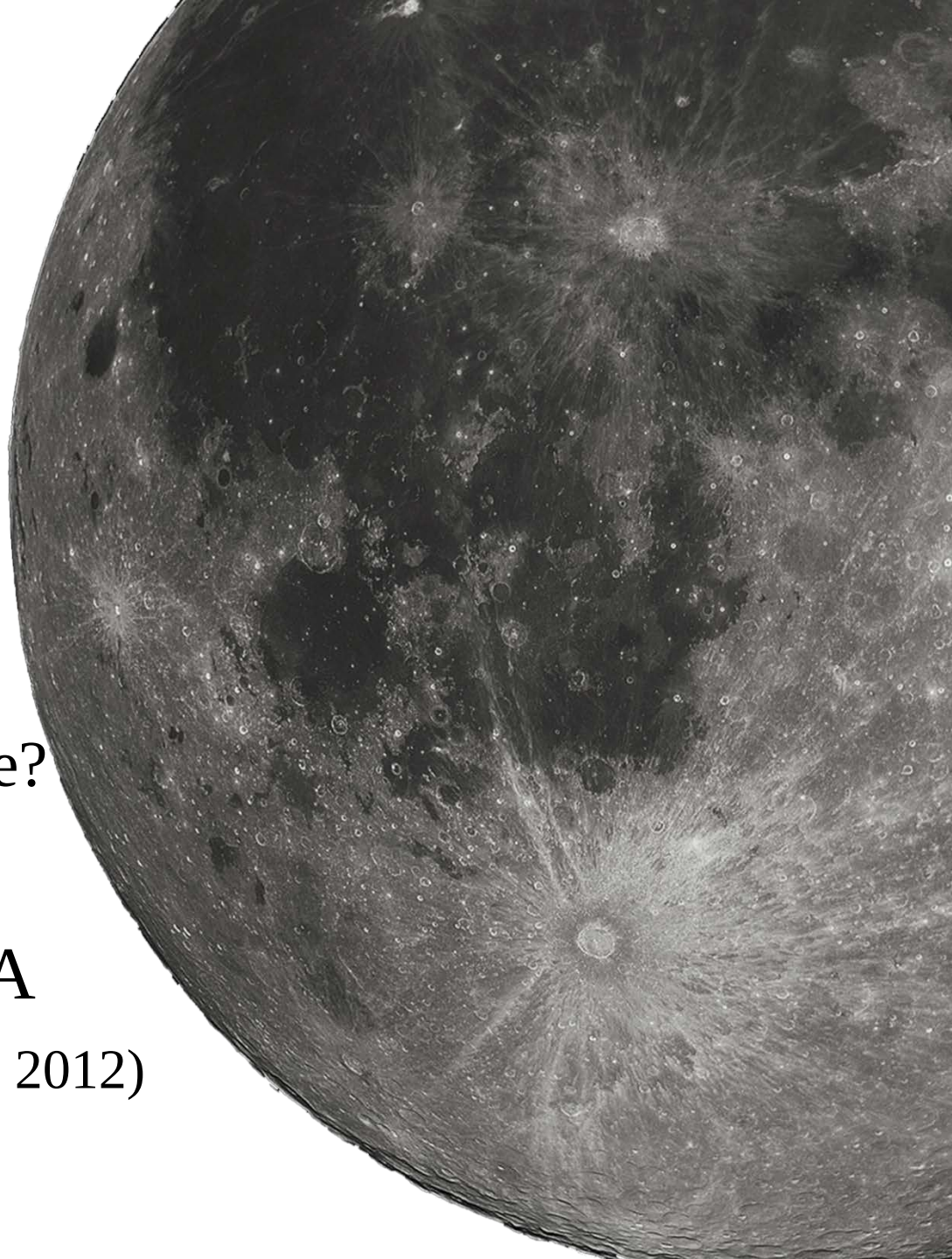
What now?

- 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

- What about the empty spaces between the basins?
  - Are they just insignificant?
  - Chaotic?
  - Something else?

- Cf. “junk” DNA

(Pennisi 2012)



Bibliography

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*.

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

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.

Bibliography

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.

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.