

**What we don't see we see and don't see:
Selection bias in linguistic description**

Rik De Busser
National Chengchi University

Bias has been a major theoretical concern in many fields of science; in many linguistic subdisciplines it has curiously been overlooked. In this talk, I present morphosyntactic evidence from two languages which illustrates that very often the knowledge we acquire through descriptive analysis and other forms of inductive linguistic inquiry is a selective and often inaccurate representation of complex linguistic systems as they are realized in the real world. I argue that these are at least partly the result of selection bias induced by presuppositions and conceptual limitations of the theoretical frameworks used in descriptive analysis.

I use data from two studies to illustrate this problem. An elaborate study of predicate-argument structure in the Takivatan dialect of Bunun (Austronesian, Taiwan) suggests that the language has no unified system for argument alignment, but rather a number of partly conflicting sub-systems. In many theoretical frameworks, this analysis is not acceptable, since argument alignment is supposed to be systematically organized in accordance with a limited number of pre-established patterns.

A study of Dutch causative and causative-like constructions suggests that the traditional view of causativization in Dutch only describes a subset of a larger set of related constructions. Descriptions commonly distinguish only two causative verbs: *moeten* 'must' and *laten* 'let'. These neatly coincide with the classical distinction between direct and indirect causation. However, a corpus search shows that there are at least four verbs which have causative-like meaning and all these have a number of related functions. These two examples both illustrate forms of selectional myopia that have a serious impact on evidence-based linguistic research, in that they unintentionally prime data to accommodate certain linguistic preconceptions.