



## Signs and terminology Science caught between language and perception

Lars VOGT

*Institut für Evolutionsbiologie und Ökologie, Universität Bonn, An der Immenburg 1, 53121 Bonn, Germany.*

*<lars.m.vogt@gmail.com>*

### Contents

Abstract .....	2
Key words .....	2
Introduction .....	2
The linguistic problem of biology .....	4
Terminology and semiotics .....	5
Terminology—one word, two meanings .....	5
Terminology <sup>terms</sup> .....	5
Terminology <sup>science</sup> .....	6
Semiotics—the science of signs .....	6
Conventional signs, natural signs and limits of interpretation .....	7
Semantic conceptual content, aesthetic nonconceptual content, and the limits in scope of terminologies <sup>terms</sup> .....	8
Distinguishing three levels of entities and their roles in a general semiotics .....	9
Linguistic conventional signs. Part 1: proper names for representing particulars .....	11
Proper names and the classical representational conception of conventional signs .....	11
Defining proper names and determining their meaning .....	12
Linguistic conventional signs. Part 2: general terms for representing universals, natural kinds and types of entities .....	12
The semiotic triangle .....	13
Sets, classes and natural kinds .....	14
The <i>type</i> approach: dealing with diversity and variability of biological objects .....	17
Terms and concepts in biology .....	19
Scope, problems and limits of terminologies <sup>terms</sup> .....	20
Organizing terminologies <sup>terms</sup> : increasing semantic transparency and communicability .....	21
Distinguishing “raw data”, data, metadata, information and knowledge .....	22
What role do images and other media recordings play in science? .....	22
Metadata, information and knowledge .....	23
Standardization and terminology <sup>terms</sup> .....	24
Taxonomy: an ordering system of terms based on class-subclass relations .....	25
New tools and approaches: knowledge representation and ontologies .....	27
The usefulness of terminological standardization and formalization and its downside .....	31
Conclusions .....	35
Acknowledgements .....	35
References .....	36

## Abstract

After briefly discussing various problems that can result from linguistic ambiguities, I attempt to provide an introduction and overview of various aspects and theories that are relevant to scientific terminology. These include a general semiotics (i.e., theory of signs), the distinction of semantic conceptual content and aesthetic nonconceptual content and their relationship to each other, a discussion of different types of scientific concepts (e.g., essentialistic and cluster classes, natural kinds, type approach), and the importance of specifying epistemological recognition criteria for empirical concepts in addition to their ontological theoretical definitions and the specification of contexts in which the concepts are used and on which theories they depend upon. I then provide a distinction of raw data, data, metadata, information and knowledge, and discuss the relation between images and data and how efforts to standardize data and metadata can affect scientific terminology. I briefly introduce new methods and techniques for increasing semantic transparency and communicability in science, which include the organization and the management of scientific terms within taxonomies, and their formal representation in ontologies. The usefulness of terminological standardization and its possible negative effects on scientific progress is then discussed, and finally the question is addressed of whether one can distinguish types of terminologies that benefit from standardization from those that could suffer from it.

**Key words.** Communicability, conceptual content, nonconceptual content, ontology, semantic transparency, semiotics, standardization, terminology

## Introduction

The quality of scientific research depends on various criteria, and **objectivity** is one of its most important and central ones. If one specific observation relies less on the particular scientist who describes it than another observation, it is considered to be more objective (*aperspectival objectivity* sensu Daston 1992, 1998 and Daston & Galison 1992; *procedural objectivity* sensu Heintz 2000). Objectivity can only be understood in reference to its complement, subjectivity (Daston & Galison 2007). Objectivity requires ruling out all individual and subjective influences of body and mind and forbids judgment and interpretation in documentation and reports of observation; it usually involves the establishment of specific experimental and measurement standards, with the purpose of ruling out the fallibility and deficiency of human cognition (*mechanical objectivity* sensu Daston & Galison 1992; *methodical objectivity* sensu Heintz 2000). Although it represents an absolute ideal that is not fully accomplishable, objectivity is nevertheless an epistemic virtue that every scientist should strive for.

There are several conditions that potentially increase overall objectivity in science. **Communicability** of scientific findings is one of the more important ones, since it represents a necessary precondition for other important aspects of objectivity, such as **reproducibility** and **transparency**. Scientific findings are commonly communicated in the form of verbal statements. Therefore, a clear and highly unambiguous assignment of sign (i.e., word, digit, symbol) should be given to its referent, the underlying real object or condition to which it refers. Typically, this requires the development of a specialized, often highly standardized, and sometimes even formalized **terminology** or language. Therefore, terminology is central to any scientific discipline and every discipline has developed its own set of technical terms.

The purpose of having discipline-specific specialized vocabularies is to provide the highest possible degree of communicability of the discipline's most commonly used types of data and knowledge. But not only data and knowledge have to be represented with a specialized terminology. Terminology is also used for the detailed documentation of the experimental design employed in generating data. Again, objectivity requires this usage, because data should be independent from the investigator generating them, a goal that can be achieved by making all necessary steps for generating the data transparent and reproducible for other scientists, and by making data production as far as possible automated (see *mechanical objectivity* Daston & Galison 1992; *methodical objectivity* Heintz 2000). Therefore, a discipline-specific terminology usually includes not only terms for representing data (and knowledge), but also terms for their annotation (i.e., data