Zootaxa 1950: 123–152 (2008)
www.mapress.com/zootaxa/

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

Learning from Linnaeus: towards developing the foundation for a general structure concept for morphology*

LARS VOGT

Systematik und Evolution der Tiere, Institut für Biologie, FU Berlin, Königin Luise Str. 1–3, D-14195 Berlin, Germany.
E-mail: lars.vogt@zoosyst-berlin.de

* *In:* Minelli, A., Bonato, L. & Fusco, G. (eds) *Updating the Linnaean Heritage: Names as Tools for Thinking about Animals and Plants. Zootaxa,* 1950, 1–163.

Table of contents

Accepted by A. Minelli: 24 Oct. 2008; published: 5 Dec. 2008

Abstract

Morphology has fundamental problems regarding aperspectival objectivity of its data—morphological terminology is often based on homology assumptions, lacks standardization, and has problems with comparability, reproducibility, and transparency. This is astonishing given that with his sexual system Linnaeus had already established a high degree of aperspectival objectivity in morphology that unfortunately has been lost subsequently. In the first part of the article a brief introduction to the history of classification is given that provides an answer to the question why morphology only initially has been gripped by the general trend towards objectification that started in the seventeenth century. The conceptual shortcomings of Aristotle's concept of essences and its link to the definition of species and taxa in natural philosophy play an important part in this development. The only solution to the problem of essences was to link it to the evolutionary concept of homology, which explains why morphological terminology today often rests on homology assumptions. By taking a closer look at Linnaeus' sexual system, basic principles for developing a general structure concept for morphology are discussed, which would provide the conceptual basis for establishing a high degree of aperspectival objectivity for morphological data. The article concludes with discussing the role of data bases and ontologies for developing a data standard in morphology. A brief introduction to the basic principles of Resource Description Framework (RDF) ontologies is given. A morphological ontology has high potential for establishing a general morphological structure concept if it is developed on grounds of the following principles: morphological terms and concepts must be defined taxon-independently, homology-free, preferably purely anatomically, and if functionally only by clearly indicating the trait's active participation in a specific biological process.

Key words: Aperspectival objectivity, Bio-ontology, Essentialism, Morphological data, Linguistic problem of morphology, RDF, Standardization

Introduction

Morphology represents a set of methods and techniques for producing data about anatomical and organizational facts of organisms. As such, it does not represent a theory or an explanatory hypothesis. When it comes to preparing morphological descriptions, morphology is all about the textual representation, documentation, and comparison of structural diversity and patterns of structural equivalences between organisms and their traits, thereby being only *assisted* by various imaging techniques for the empirical substantiation of these descriptions. Therefore, morphological terminology and language assume a central methodological role in morphology. Only if the language and terminology used in morphological descriptions are capable of reliably transporting the relevant information in an unambiguous way and independent of individual morphologists, and only if they enable the comparison of morphological data across a broad taxonomic range, will morphology meet the high degree of comparability and communicability of data that is being increasingly demanded in the age of a growing importance of data bases in biology.

Unfortunately, morphology lacks standardization and common acceptance of morphological terms and lacks a formalized method of recording and documenting morphological descriptions (Vogt *et al. submitted*). Thus, morphology has fundamental problems with its terminology. As a consequence, morphological terminology and morphological descriptions vary from author to author, the meaning of morphological terms often changes through time, and the applicability of morphological terms is often restricted to a specific taxonomic group and cannot be easily adapted to other groups. In scientific research practice, this non-standardization of morphological terminology and the diversity in quality, organization, and style of morphological descriptions frequently lead to divergent descriptions of equivalent traits or to identically described morphological traits that are in fact not identical (see linguistic problem of morphology, Vogt *et al. submitted*; see also Ramírez *et al.* 2007).

These linguistic ambiguities pose fundamental problems for comparative morphological studies, being the source for repeated misunderstandings among morphologists, undermining the possibility to reliably communicate morphological data. Reliable communication of data, however, represents a necessary prerequisite for the division of labor not only among morphologists conducting comparative studies over a broad taxonomic range, but also for all kinds of co-operations in which morphologists are involved or morphological data are analyzed. Thus, it seems that morphology is hard pushed these days to prove that its standards of objectivity, comparability, and communicability still hold up to non-morphological biological data, as for instance DNA sequence data. Considering these fundamental problems, it is not surprising that some biologists even claim that morphology has already lost its traditionally prominent role in phylogenetics (see e.g., Scotland *et al.* 2003), since comparability of data represents a *sine qua non* of phylogenetic research practice.

The interaction between phylogenetics on the one hand, or more traditionally biological taxonomy and classification, and morphology on the other hand represents a liaison with many different facets and a continuous story of mutual interference. This is not surprising since taxonomy and classification have been one of the initial fields of application of morphology, in which traditionally it always had been very strong. As a consequence, much of morphological terminology and methodology has been strongly influenced by the needs and requirements of generating classifications.

In the following I will provide a brief introduction to the history of classification and its impact on the development of morphological terminology and methodology, including conceptions of naturalness and the epistemic status and role of observation and empirical investigation. This historical excursion is intended to give a historical explanation for the question why objectification did not catch on in morphology, while in many other biological disciplines objectification has advanced to a level that established a high degree of transparency, reproducibility, communicability, and inter-subjective consensus regarding empirical data. This question becomes even more interesting when considering that, initially, morphology has been gripped by the general trend towards objectification that started in the seventeenth century. In the second part of the article I will show that there is a lot to learn from Linnaeus' approach to classification in terms of increasing objectification in morphology. Based on Linnaeus's sexual system, I develop the basics of a general structure concept for morphology and argue that it takes in a key role in the context of objectification of morphology. I conclude the article with a brief introduction to biological data bases and standardized controlled vocabularies (i.e., bio-ontologies) and how they can serve as a basis for establishing a general structure concept in morphology and its broad dissemination.

The historical burden of essentialism

Aristotle and essentialism

In order to be able to identify regularities in biology and to generalize about the biology of organisms, a concept for abstracting individual organisms into types (i.e., classes, kinds, families) is required. From a logical point of view any given set of organisms shares an infinite amount of equivalent properties—in other words, any set of organisms could be conceptualized as some sort of kind. Thus, unfortunately, possessing same traits does not necessarily imply ontological equivalence of the respective organisms. As a consequence, generalizations in biology would become impossible as long as biologists would not manage to differentiate between 'real' kinds and artificial kinds; but how to recognize and define 'real' kinds?

This problem was known to ancient Greek scholars, who recognized the necessity to develop a concept for differentiating between essential and accidental properties. Aristotle defines 'real' kinds according to their *essential* properties and classifies them according to the method of logical division of *per genus et differentiam* (Fristrup 2001). The English word 'essence' comes from Latin *essentia* (from *esse*, 'to be'), which represents a translation of Aristotle's ancient Greek phrase *to ti ēn einai* (i.e., 'what it is for a thing to be'), denoting a thing's essence.

An Aristotelian essence represents the attribute or set of attributes that make an entity what it fundamentally is. Without its essence, the entity would lose its identity. Essential properties are real physical properties of the 'nucleus' (i.e., substance) of a thing. For any specific kind of entity, there is a set of essential properties, all of which any entity of that kind *must* have. As a consequence, if two objects share the same essence, they can be considered to be truly *identical* with respect to this aspect of their 'nucleus' and therefore can be classified as instances of the same 'real' kind. That is the reason why, according to Aristotle, essences are fundamentally linked to definitions of different *kinds* of entities.

Aristotelian definitions are hierarchically organized, resulting in a hierarchy of classes and their subclasses (Fig. 1). The defining attributes of a class are inherited downstream to its subclasses (i.e., downward propagation). Thus, if a given entity is an instance of a specific class, it is necessarily also an instance of all those classes of which this class is a sub-class. This hierarchy represents a taxonomy (i.e., *taxonomy* in a broad sense) of more and more specialized concepts, which implies a hierarchical organization of terms (i.e., taxonomic inclusion, Bittner *et al*. 2004).

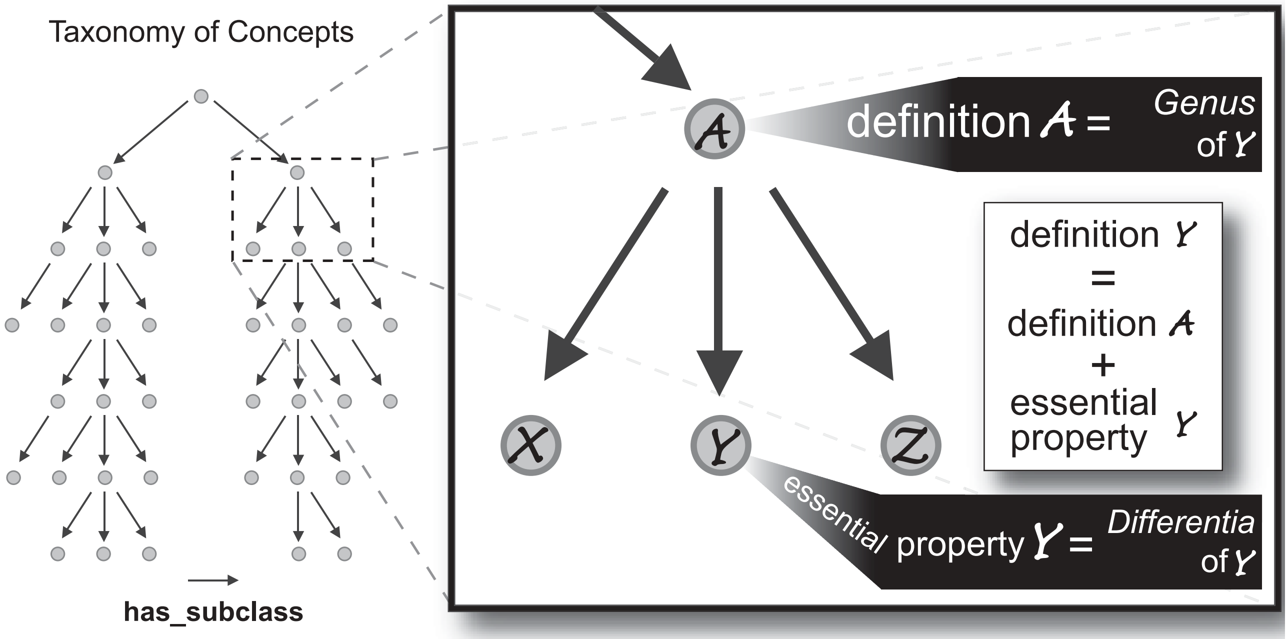

FIGURE 1. Aristotelian definitions: The definition of the 'parent' kind is inherited by all of its 'first child generation' kinds and forms the *Genus* part of their definitions. These represent properties that an instance of a 'child' kind *necessarily* has to possess. On the other hand, the essential property of a 'child' kind represents the distinguishing property that is in combination with the *Genus* part *sufficient* for the recognition of an instance of that kind, since all and only individuals that are instances of this kind do possess this property. This forms the *Differentia* part of the kind's definition. *Genus* and *Differentia* together represent the essence of the kind and at the same time its definition.

When defining a new kind of entity, Aristotle refers to this conceptual relationship of hierarchical specializations. In his definitions, Aristotle distinguishes between what he calls 'genus' and 'differentia'. 'Genus' represents attributes referring to essential properties of all of the respective 'parent' kinds, which an instance *necessarily* has to possess for membership to the 'child' kind. However, these properties are not sufficient for membership recognition. 'Differentia', on the other hand, represent attributes referring to the essential properties of the 'child' kind. 'Differentia' are required to distinguish the kind to be defined from all other kinds of the same hierarchical level. If no other entity, but the instances of the kind to be defined, possesses a certain property and if *all* instances of that kind possess it without any exception, that property qualifies as 'differentia'—the essential property of the kind. Only the combination of 'genus' and 'differentia' is *sufficient* for recognition of membership to the kind. Thus, on *any* hierarchical level of a classification based on Aristotelian definitions the essential properties of all 'parent' kinds provide the 'genus' part of the definition, and the essential properties of the 'child' kind its 'differentia' part; and 'genus' and 'differentia' together represent the essence of the kind (Fig. 1).

Naturalness, essences, and the epistemic role of observation

To Aristotle, observation and accurate description of biological individuals (i.e., single specimens) obtain a central epistemic function in natural philosophy. Following Aristotle, the study of individual natural phenomena is a substantial prerequisite for a philosophical representation of the natural world. Thus, it is not surprising that Aristotle bases his definitions on observation and the study of specimens. His terminology is clearly grounded in observation. According to Aristotle, essences are real physical properties that can be, in principle, discovered through observation.

However, whether a given morphological trait of an organism represents an essential property of its corresponding kind is not directly testable and had to be inferred through comparison and empirical investigations. Aristotle's concept of essences lacked clearly and unambiguously applicable recognition criteria, with the consequence that statements about essences always remain hypothetical and cannot be sufficiently validated in principle.

Aristotle's essentialism had a major influence on all subsequent classificatory attempts in biology and represents a paradigm concept for most biologists to follow. After Aristotle, every biologist who wanted to define a taxonomic group or a morphological trait had to deal in one way or another with Aristotle's concept of essence. Thereby it underwent several major alterations, all of which also significantly influenced morphological methodology and terminology. It is no overstatement to say that what follows in the history of biological classification can be characterized as an enduring reaching out for Aristotelian essences, which at its turn strongly influenced the epistemic status of morphology and the conceptualization of morphological data.

The influence of medieval hermetism

While during the time of Aristotle empirical investigations take on an important role within sciences, this fundamentally changes at the latest in the Roman Empire of the second century. Culturally it is a melting pot of most diverse peoples and languages, a mixture of different ideas and ideologies, in which, officially, all different kinds of religions and deities are more or less tolerated (Eco 1988), but in which the young Christian community was significantly growing and constantly gaining influence. In this time the concept of truth, as it was delivered from the Greek rationalistic tradition, experiences a large crisis. On the search for a *single* truth within the multiplicity of most different religions and cultures, hermetism gains importance and influence.

Many ideas, most of which were spiritually related, influenced the development of hermetic thinking. For instance the Jewish-Christian idea of the existence of a universal language that all peoples spoke until the building of the tower of Babel, when God decided to give the workers different languages to prevent them from finishing their work. This universal language was believed to be closest to the language of Paradise, in which, since it was the language of God, only the truth and nothing but the truth could be spoken. The hermetic idea was that every language carries pieces of this old universal language of Babel, with Hebrew being closest to it, followed by ancient Greek. Therefore, hermetists believed that manuscripts in different languages, preferably Hebrew and ancient Greek, all carry traces of the language of Babel and, thus, hidden pieces of truth that only had to be revealed and discovered by studying and comparing the texts. Thus, in hope that each book holds a spark of truth and that all books confirm each other someway somehow, hermetists focused in their search for truth exclusively on the content of books. Thereby, the principle of *tertium non* *datur* (i.e., excluded middle principle: a statement can only be true or false—either A=B or A≠B, a third possibility being considered to be impossible) is invalidated. Consequently, various things are regarded to be simultaneously true, even if they contradict each other, and empirical research loses its importance (Eco 1988).

If books, however, tell the truth although they contradict each other, each of their words is to be interpreted as an allusion, as an allegory. The universe is interpreted as a network of relations, in which each thing reflects and means all other things—a universe of universal sympathy, which a human being will only comprehend through a web of allegories.

Obviously hermetists were seeking the truth that is lying *behind* the objects and that *transcends* observation and description. In doing so, they did not use language as Aristotle did—words and things were not conceived as necessarily separate concepts and categories, but were believed to belong essentially together, connecting humanity, nature, and the divine with one another, with the ultimate goal to restore the paradisiacal union between them. This hermetic philosophy renders the inference of empirical evidence, with which one could eliminate or refute hypotheses, impossible.

As a consequence and in reference to the Christian influence on hermetism, the medieval notion of truth was more complex and multilayered than ours today: to a literal truth, a complex structured, transferred moral-spiritual truth was associated. On the one hand, this drastically constrained the possibilities of interpreting empirical phenomena and the number of potential explanations, since they were not allowed to contradict the religiously shaped world view of people of this time. On the other hand, they were tremendously extended due to the idea of a potentially infinite repertoire of possible mechanisms of godly intervention.

From the second century until the early Middle Ages, it is characteristic for biological classification that one does not differentiate between observation, document, and fable, which represents an obvious influence of medieval hermetism. Signs are understood as intrinsic properties of their denotations—they are thought to be essentially connected. Thus, following the ideas of hermetism, one assumes things possess at the most visible point of their surface *signatures* or markers, which are to indicate what is substantial (i.e., essential) with them. The signature assigns the meaning to the organism. Heart-shaped leaves of a plant, for instance, must possess a certain effect on the human heart due to their similarity, or they are at least related to it through some cosmic connection. Another example is the belief in mercury salve helping to treat syphilis because mercury is signed by the planet Mercury, which at its turn signs the market place where syphilis has been usually contracted.

Paradoxographies and early encyclopaedists

The classical natural miracles and marvelous creatures of the paradoxographies, to which the Blemmyes (headless creatures with their mouth on their belly), the Sciapodes (who like to lie in the shade of their single large foot) and the Cyclopes (one-eyed creatures) belong, represent an ancient literary genre about which even Aristotle wrote and whose meaning and role is still uncertain (the Greek texts may have served as collections of samples for rhetoricians; Daston & Park 1998). Since paradoxographies were considered to represent essential components of any encyclopedia they have been copied into each medieval encyclopedia and were presented therein as knowledge to be true without any critique—obviously, people of the Middle Ages actually believed in the existence of the fabulous creatures of the paradoxographies.

Starting from third and up to the fourteenth century, the typical compilers of historical and topographic encyclopedias rarely left their own hometown, but, instead, supported their work solely with the content of older encyclopedias. In doing so they knew to tell tales from real and legendary countries inhabited by all kinds of fabulous animals and strange creatures, including those from the paradoxographies (Daston & Park 1998). Thereby, the influence of the Christian world view also had a strong impact. The lion, for instance, was believed to possess the habit to obliterate traces with his tail so that no hunter can take up his track. This proposition was believed to be true since the lion was considered to be a symbol of Christ who, for his part, erased the sins of mankind. In this way Christian hermetism provided a rational basis that gave a certain internal coherence to this proposition. The same applies to Phoenix, a bird-like creature that burns every 500 years on an altar and three days later resurges again from his own ashes, because in the Middle Ages the Phoenix represents the symbol of the redeemer (Eco 1988).

At the end of the twelfth century some scholars and naturalists start to be critically opposed towards the ancient knowledge. Gerald of Wales (lat. Giraldus Cambrensis) for instance stresses the importance of personal experience of flora and fauna. He is one of the first to criticize the mistakes of antiquity and to doubt the existence of the fabulous creatures described in the paradoxographies. For the first time a medieval topography does not exist as pure literary compilation—Cambrensis' work on the history and topography of Ireland, written around 1185, is enriched with personal observations and criticism of Bede, Solinus, and others.

This development has also been influenced by the fact that with the eleventh century the time of the extensive oriental expeditions and voyages begins. Europe opens its cultural borders and ends its isolation as a consequence of, for instance, the opening of the commercial routes and the Mongolian peace. Missionaries, ambassadors, mercantilists, and researchers start to travel distant countries.

Nevertheless, to many authors it still applies: what is to be found in exotic countries is written in the ancient books and the Bible. Therefore, it does not surprise that at the beginning of the fourteenth century illustrations were still added to the report of Marco Polo's journey to China. Those illustrations were added by third persons and showed creatures like the Blemmyes, the Sciapodes, and the Cyclopes, which, following the ancient writings, were to be expected in India, although they had not been described by Marco Polo (Eco 1988). This impressively points out the enormous authority that the ancient writings had in this period of time and the negligible value that was given to observation and personal experience.

The influence of medieval hermetism extended even into the seventeenth century. In books such as his *History of Serpents and Dragons* (Aldrovandi 1648), the influential Ulisse Aldrovandi, professor in Bologna and owner of the largest and most famous collection of naturalia in Europe during the seventeenth century, still mentions for every creature and organism listed the recommended method to catch it, its allegorical use, on which coat of arms it is to be found, known legends and narrations, as well as the best way to serve it with sauce. All this information is listed in addition to and mixed within pieces of knowledge about its biology, following no specific categorical order. Thus, even well into the late Renaissance, organisms are still understood hermetically within the *entire* semantic net that connects them with the world, thereby *not* distinguishing words and things as necessarily separate concepts and categories.

Naturalness, essences, and the epistemic role of observation

With the rise of hermetism, the epistemic role of observation fundamentally changes. This has a major impact on biological classification. While Aristotle gives observation a central epistemological function, his early medieval commentators do not follow him therein—to them all observation that is based on single occurrences is accidental. Observation and empirical comparison no longer serve as a basis for classifications. Definitions of kinds of biological entities are no longer given in terms of physical properties only. Hermetic essences are physical properties that have to bear spiritual meaning. Morphological terminology is not anymore grounded only in observation. What is considered to be real is not necessarily observable but must be referable to the Bible or ancient texts.

Therefore, the medieval philosophers do not regard it as their task to observe natural phenomena and to discover explanations for them. They rather dedicate their time to the study of the universal truths, which they believe to exist as a network of relations of analogy, all dependent on an underlying, singular big metaphysical cause. They were convinced that the universal truths can only be discovered by studying the books or by receiving them from their teachers, which they then only had to improve and refine. This method is also known as *doctrina* because it depends on the passing on of knowledge by instructors (*doctores*).

Emancipation from the ancient heritage

With the invasion of the Iberian Peninsula by Islamic Moors in 711 and the following establishment of the emirate of the Ommiades of Córdoba in 756, the cultural exchange between Arab and European cultures dramatically increased. In the Islamic cultural centers (Baghdad, Damascus, Cairo, Mecca, Samarqand) Oriental and Hellenic knowledge was merged by the Arabian scholars and provided the foundation of an Islamic science, which reached its climax in the ninth and tenth century. Among Arabian Moors who occupied Spain there were scholars who earned their living by translating into Latin ancient texts that had previously been translated into Arabic. This contributed considerably to the intellectual Renaissance in Europe.

A major consequence of this influence is to be seen in the introduction of new opinions about the order of nature, in which nature was seen no longer as the direct expression of divine arrangements, but was thought to be subordinated via internal orders formed by causal chains. With Adelardus and other scholars of that time, the idea of an autonomous natural order was brought into the scientific discourse. This marks an important break with medieval hermetism, as it implies the assumption that the structure of nature and the universe must exist *independent* of humankind and human culture. This represents a precondition for the establishment of *ontological objectivity* (see e.g., Daston & Galison 1992).

The discovery of the 'New World' and the collections of pharmacists and physicians

A significant influence, not only culturally and economically, but also regarding scientific methodology was provided by the discovery of the '*New World*' of America by Christopher Columbus in the year 1492. An immense quantity of unknown species and exotic naturalia are made accessible for study. Girolamo Cardano, lawyer, mathematician, professor of medicine in Pavia and Bologna, and owner of a collection of naturalia, refers to naturalia from America to impressively point out the gaps of ancient knowledge (Cardano 1557). Insights of this kind lead to the consequence that from this point the authority of the Greek and Roman authors experienced a collapse.

Giovanni Battista Olivi is one of the first who studies the naturalia of the New World for their own sake (Olivi 1584). Others will follow him. Gradually, something like the discipline of *natural history* develops, emancipating itself initially only slowly from medicine. In natural history, natural objects are investigated independently from their possible therapeutic applications. With the emergence of the discipline of natural history the great time of biological classifications begins.

Natural history is understood to represent the task of moulding language as to represent things – natural history cares about an unambiguous *designation of the visible*. Regarding content and meaning, natural history is basically a classificatory discipline that deals with the description and classification of plants, animals, and minerals (Kanz 2002). Later, by Francis Bacon, this new discipline will be given an important role within his reformed system of natural philosophy, the 'new philosophy'.

The encyclopedic work of Jan Jonston (1657) represents an impressive example for the corresponding change in thinking, in which a majority of *formerly* essential entries are now missing, which would have been listed by, for instance, Ulisse Aldrovandi. Jonston limits his description of organisms to known biological data as well as the possibilities of the utilization of the organism for human purposes — well-known legends and the like are not thought to belong to the organism in principle and are therefore ignored. Henceforth, as a consequence of the advent of the discipline of natural history, observation and its documentation are separated from fables, but are from then on understood as representational types in their own right. This represents a very important step towards the establishment of *ontological objectivity* (Daston & Galison 1992) in biology.

Bacon's empirical facts

By the growing popularity of scientific collections and the increasingly effective publishing of scientific essays due to the new book printing techniques, contacts among scientists increase within Europe, crossing cultural and national borders. As a result, an active scientific exchange develops. This is the ground on which, during the seventeenth century, a new epistemology of empirical *facts* develops together with a community of empirical researchers.

Francis Bacon was one of the leading figures of this ‘new philosophy’. He had the idea that nature has just to be interpreted correctly, that nature tells its own tale and scientists only have to learn how to listen to her. Nature would reveal herself as soon as scientists manage to rule out their intervention (Galison 1998). This is based on the idea that the ultimate structure of reality is independent of humankind and human culture (*ontological objectivity sensu* Daston & Galison 1992).

In trying to do so, Bacon conceives natural history on the ground of his own epistemological system, in which natural history takes in an epistemological key function (Spedding *et al.* 1857–1874). According to Bacon’s *Novum organum* (Bacon 1620), natural history must serve as source for facts and as the empirical foundation for natural philosophy. Facts represent a new category of scientific experience, detached from explanations, illustrations, or conclusions. Following Bacon, scientists should put aside their terms and should start to deal with facts.

Bacon’s empirical facts are grounded in observation. However, observation is always spatio-temporally fixed and therefore represents a localized event. Observation not only depends on the conditions present at a specific location but also on a given individual with all its idiosyncrasies. Thus, the important question that Bacon had to answer was, how necessarily subjective experience can be transformed into objective empirical facts. Which conditions have to be met to turn a ‘view from somewhere’ (Porter 1992) into a ‘view from nowhere’ (Nagel 1986)?

According to Bacon, the process from the study of particulars to the identification of facts is considered to be burdensome, since the senses have to cross the corset of a strict method of order, processing, and evaluation to be immune against deceptions. Bacon develops boards, so-called *praerogative instances*—boards of agreement, differences, gradations, and repulsion of properties in nature. They are to serve as methodological aids to the discovery of the laws of natural properties and their fundamental forms. The praerogative instances provide the methodological core to Bacon’s conception of natural history, which in its turn was conceived to be free of any theory, and to represent one of the first attempts to establish *aperspectival objectivity* in science.

Aperspectival objectivity (see Daston 1992, 1998; Daston & Galison 1992; *procedural objectivity sensu* Heintz 2000; for a critique of the claim for perspective-independent objectivity in science see, e.g., Kukla 2006) is about communicability of scientific results and claims that something is more objective than something else if it relies less on the specific individual who generated the results, their social position and character.

As a consequence of Bacon’s conception of natural history, observation and its documentation in form of descriptions are *methodologically sharply separated* from the conclusions drawn from them, including all explanatory hypotheses and theories. This represents a major methodological improvement. The focus on facts and their consistent distinction from conclusions should also prove to be favorable for the just emerging scientific societies and their meetings, since one can usually talk about facts more objectively and without personal arguments than it is possible with theories. Bacon’s ‘natural histories’ subsequently developed into the ‘*facts*’ of natural philosophy of the late seventeenth century, which in their turn represent the attempt to generate pure descriptions that should be free of any theory or conclusion and which represent the precursor of our modern notion of empirical data.

During the seventeenth century more and more scientists and philosophers became aware of the subjectivity of perception. Locke (see also Galilei, reprinted in Drake 1957), for instance, distinguishes between primary qualities such as shape, size, distance, solidity, and volume that, according to him, exist in the external world in the same way humans perceive them, and secondary qualities such as color, taste, texture, smell, and sound that, following him, do not exist in things themselves but depend on the perceiver’s senses (Locke 1689/1979). Following this distinction, Bacon’s natural history had to focus on primary and not on secondary qualities.

Naturalness, essences, and the epistemic role of observation

During the seventeenth century more and more scientists and philosophers became aware of a) the subjectivity of perception and b) the requirement of natural philosophy to be grounded on perception. As a consequence, observation and empirical comparison regain their epistemic role. But scientists are at the same time suspicious of singular observations of individual specimens and apply rigorous methods and techniques. They believe that to discipline observation requires the distinction of different observational categories in order to minimize its subjectivity (e.g., by distinguishing primary and secondary qualities). Consequently, essences are considered to exist *independent* of human beings (ontological objectivity) but require specific methodological procedures for their identification and documentation (aperspectival objectivity).

The age of classification

Bacon also had a major influence on biological classification. Jungius, a student of Cesalpino, started to develop criteria for generating something like Baconian empirical facts for morphology. He applied a philosophically derived rationale to the observation of plants. Jungius believed that well defined terms, just like numbers in mathematics, represent stable and objective values. Based on this assumption, he defined clear and unambiguously applicable botanical terms that he stripped off of all untested physiological interpretations, thereby clearly separating observation from conclusion (Jungius 1678; von Sachs 1875; Beck 1969). Jungius' goal was to develop a *primary language* (i.e., a purely descriptive one), with which observation and its documentation through descriptions should become independent of the respective individual observer and universally communicable. By standardizing morphological terminology and descriptions of morphological traits, Jungius attempted to establish *aperspectival objectivity* (Daston & Galison 1992) within botanical morphology and classification.

The standardization of experimental and linguistic practices becomes more and more popular during the nineteenth century, thereby stressing the importance of *communicability* of scientific results among scientists. Objective knowledge comes to be defined as communicable knowledge and requires scientists to standardize their methods of measurement and communication. Jungius thereby follows the most important strategy for linguistic standardization: quantification and formalization—to use formulae, numbers, and graphs whenever possible (Heintz 2000).

One of the first to adopt Jungius' terminology for the purpose of biological classification was Ray. Thereby Ray followed Cesalpino (1583) and Tournefort (1694) in focusing on the properties of flowers and used only distinct and exactly definable properties of morphological traits for classification (Ray 1703).

Cesalpino, Jungius, Ray, and Tournefort prepared the ground for Linnaeus and his new and very successful approach to biological classification. Linnaeus reduced and limited observation to a few categories only, so that one is left not only with what is analyzable in the somewhat confusing opulence of representations but also with what anybody can recognize and identify and, thus, with what can receive a name that everyone understands (Linnaeus 1735, 1751). With his sexual system (i.e., *Clavis systematis sexualis*; for more detail see below), Linnaeus proposed a very pragmatic way to make biological classification a less subjective procedure, at least for botany. Linnaeus' sexual system provides a theory for *taxonomic character* that takes in the function of a *secondary language* (i.e., analytical/explanatory), with which classification should become independent of the individual taxonomist.

Biological classification before Linnaeus is characterized by a plurality of contradictory approaches (Ereshefsky 1997). Actually, Heywood (1985) concludes that before Linnaeus biological taxonomy must have been a rather chaotic discipline, which stands out by its miscommunications and misunderstandings. The tremendous success and the broad acceptance of the Linnaean system are probably due to his rather pragmatic choice of criteria for the conceptualization of classification (for his contribution to the theory of biology see Müller-Wille 1999). Linnaeus' approach provided comparatively clear and simple rules for the construction of classifications, which also included rules for the denomination of species and taxa. This significantly increased the possibility of communication among taxonomists (Ereshefsky 1997; Stevens 1997).

Naturalness, essences, and the epistemic role of observation

The ‘new philosophy’ of Bacon, Descartes, and Locke, had a major influence on taxonomists. In order to satisfy the claim of *transparency* and *reproducibility* that accompanied this new style of doing empirical research, taxonomists strived for making classification a less subjective procedure by relying on mathematics and logic. They limited themselves to the study of distinct, clearly and unambiguously definable morphological traits, with which they attempted to establish a terminological standard in morphology-based classification. This also improved possibilities to *communicate* one’s findings and discuss and agree upon possible classificatory relationships (i.e., aperspectival objectivity). As a consequence, essences are understood to represent real physical properties of traits that are observable, unambiguously describable, and that can be identified independently from a particular taxonomist.

Causal reasoning and the role of function in classification

Development of the laboratory method

Medieval natural philosophers strive for *scientia* (e.g., theology and theoretical medicine), which, according to the definition of Aristotle, represents safe knowledge based on classical syllogisms and, thus, ultimately on postulation and deduction. In contrast to *scientia* stands *artificium*, which is associated with handicraft and which represents another form of researching that is based on experience and the development of laboratory techniques and instruments. Instead of the safe knowledge of *scientia*, *artificium* was considered to only be able to produce reliable opinions, as can be found in practical medicine and agriculture.

Due to the increasing urbanization and the increasing trade at the end of the fourteenth century, the urban bourgeoisie flourishes and a solvent commercial, crafts, as well as intellectual urban elite emerges. These elites let the market for professional medical supply boom. As a consequence of this development and due to its increased social importance, the field of medicine that is concerned with the diagnosis, the description, and the treatment of individual diseases, receives an enhanced position. Noble patrons promote for instance the research on spas, i.e. medical springs. Giovanni Dondi, physician and professor of medicine, is engaged in the observation and description of thermal springs in the proximity of its hometown Padua (Dondi 1372–1374). Giovanni is not alone as Ugolino da Montecatini (1471), Michele Savonarola (1448–1449), and many others of that time were busy with the study of thermal springs as well. During their research they quickly become aware of the fact that the classical methods of *scientia*, which reduce their reasoning to deduction from first principles and the study of literature, are not suitable for the study of thermal springs. Thus, only with the aid of empirical experience and the application of the methods of *artificium* could successful research be accomplished in this field, leading to the development of methods and techniques for experimental exploration and measurement of more complex observable relations. Thus, it is in the tradition of *artificium* that methods of experimental exploration and measurement developed, which represent the core of the emerging new *laboratory method*.

Influenced by the principles that Francis Bacon introduced in his *Novum Organum* (Bacon 1620), the ‘Invisible College’ was founded in the seventeenth century—a precursor of the Royal Society of London. Its members devoted themselves to Bacon’s ‘new philosophy’ and were dedicated to acquire knowledge through experimental investigation (Gingrich 2004). Among its members was Robert Boyle. Assisted by Robert Hook, Boyle went through numerous modifications and changes in design and construction of his air pump, finally leading in 1659 to his ‘Pneumatic Engine’, with which he began to run a series of experiments. Through this experimentation Boyle discovered that the volume of gas varies inversely to its pressure. This, by now famous law resulted from his extensive experimental work with the air pump (Boyle 1660). Boyle’s experimental approach marks a change in paradigm in science, representing a consequence of what is often referred to as the scientific revolution. It is paradigmatic for the enthronement of empiricism and experimentation as primary instruments for gaining knowledge (Shapin & Schaffer 1985; Shapin 1996).

In the beginning of experimentation, experimental results had to be validated and authenticated by trustworthy witnesses. At that time, trustworthiness of a witness was primarily defined in terms of their social position—only the unbiased judgment of a *gentleman* can witness experimental results. Scientific societies take in a central position in the establishment of this *social objectivity* (Heintz 2000). With an increase of standardization of experimental methods and techniques, with the development of instruments for data production, which replace the scientist as an observer, and with the (international) standardization of measurement procedures as well as measuring units, *gentlemen*-science is suppressed and social objectivity is replaced by mechanical (and aperspectival) objectivity (see Heintz 2000). *Mechanical objectivity* (Daston & Galison 1992; *methodical objectivity sensu* Heintz 2000) requires the ruling out of all individual and subjective influences of body and mind and forbids judgment and interpretation in documentation of observation and reports on it (Daston & Galison 1992).

Causal reasoning

Isaac Newton formulates the first of the ‘rules for the operation of comprehension’, a call for economy of thinking in natural philosophy, using only as many causes for the explanation of natural phenomena as necessary (Newton 1687). This methodological principle, also called parsimony, which goes back as far as to Aristotle, but is commonly attributed to William Ockham (‘Ockham’s razor’), enjoys at the end of the seventeenth century increasing popularity. Parsimony provides the methodological restriction for the choice of the best explanation which is necessarily required for empirical research. With it, functionality moves into the focus of naturalists—the simplicity of nature and the economy of their instruments are connected with the sobriety of purposes (Daston & Park 1998). The conception of the regularity of causes is associated with that of the regularity of effects. Also among the anatomists of this time, the coupling of morphology/anatomy and functionality plays a prominent role. To them, function begins to represent the most important issue. Thus, already in 1718, the Parisian anatomist Jean Mery thinks of the ‘machine of the human body’ (Mery 1718).

The role of functional morphology in classification

In the distinct and usually very reliably and easily preservable forms of plants, natural history finds an ideal object for its research (additionally, by mailing seeds, specimens could by traded between botanical gardens). This contributes substantially to the boom of botany, resulting in the formation of many botanical chairs at the universities during this time. The actual practice of classification, however, proves that not every trait provides differentiating properties and can serve as a taxonomic character—not every property represents an essential property. The scientific task of taxonomists therefore consists in finding suitable traits in order to receive accurate names for the objects to be classified. This brings up some difficulties. One has to ask oneself what is ‘suitable’—a problem that results from the concept of essence lacking clear recognition criteria. As a consequence, a multiplicity of different classifications are conceivable, each of which is based on different sets of traits. Michel Adanson (1763), for instance, came up with 65 different classification systems this way.

Classification and taxonomy prior to causal reasoning arranged the knowledge about organisms according to the possibility of representing them within a system of names. As a consequence, many different classifications are possible and their scientific value is evaluated on grounds of rather pragmatic criteria (e.g., Stevens 1997). Classification experiences a major change in paradigm with the development of the idea of a *hierarchical* natural order that can be discovered by taxonomists through causal reasoning, in which function takes on a central role.

At the beginning of the nineteenth century the term and concept of *organization* develops, which refers to the internal physique and physiology of an organism. The term and concept of *function* receives specific attention, since it relates individual morphological traits to the entire organism and, thus, to other morphological traits. The morphological traits together with their functional relations establish the organization of the organism. On the basis of the concept of organization, a trait can be evaluated in reference to its functions and their importance to the organism: important traits provide functions that are essential for the survival of the organism, whereas less important traits do not. Thus, by defining an internal law (i.e., the function), organization mediates between *morphological trait* (i.e., primary language, purely descriptive) and *taxonomic character* (i.e., secondary language, analytical/explanatory). Function is understood to permit a certain trait to adopt the value of a taxonomic character. As a consequence, classification receives a completely new conception. With the criterion of function classification becomes a natural system. The hierarchical order of classification is neither merely given by God nor only dependent on the cognitive constraints and requirements of humans anymore. Function determines the hierarchy within the system.

Cuvier, the developer of the modern discipline of comparative anatomy (Coleman 1964), assigns sets of organs to a specific *function* and tries to reveal similarities by comparing them (Cuvier 1800, 1817, 1825)—although Linnaeus justified the epistemic role of his sexual system as providing the key traits for classification on the function of the sexual organs for reproduction, thereby following the tradition of Cesalpino, this reference to function is owed to an Aristotelian concept of 'being' (see, e.g., Larson 1967) rather than a modern causal account of function. According to Cuvier, function is not assigned to anatomy anymore, but instead receives primacy—which was the matter of debate of the famous argument between Cuvier and Geoffroy Saint-Hilaire in 1830 (see e.g., Appel 1987), in which Geoffroy Saint-Hilaire argued that the body plan of an organism constrains how organ functions are manifested (i.e., form determines function), whereas Cuvier argued that function determines how organs are designed (i.e., function determines form). According to Cuvier, any similarities between organisms are due to common functions. During his comparative studies of the internal organization of organisms Cuvier noticed that individual organs can no longer be conceived without the other organs and that all other organs would have to change as soon as one of them changes (Cuvier 1800; principle of functional correlation, Russell 1916). Thus, Cuvier introduced the idea of organic integration into biological thinking, which is fundamental for our modern conception and understanding of the organism (Fristrup 2001).

Furthermore, Cuvier is convinced of being able to recognize an internal hierarchy of morphological traits, with some traits possessing a greater functional importance for the organism than others (principle of subordination of characters; e.g., Coleman 1964; Farber 1976; Eigen 1997). On the basis of the most important function he tries to find the most important type of taxonomic character. This type of character, in its turn, has to serve as foundation for the higher ranked taxa in a classification. In doing so he initially dedicates himself to the investigation of blood circuits, then to digestion, and later to nervous systems. On the basis of having identified four types of nervous systems, Cuvier classifies animals into four distinct basic classes, the '*embranchements*' (i.e., morphological types—for a discussion of Cuvier's type concept see Eigen 1997): Articulata (i.e., arthropods and segmented worms), Mollusca (i.e., all other soft-bodied bilaterally symmetrical invertebrates), Radiata (i.e., cnidarians and echinoderms), and Vertebrata.

Naturalness, essences, and the epistemic role of observation

Many different and mutually contradicting classifications were proposed for animals on the basis of Linnaeus' method of classification. But also regarding the classification of higher ranked taxa in botany Linnaeus' sexual system failed to provide a consistent solution, as he himself had to admit (Linnaeus 1751; see also Larson 1967). This was unsatisfactory and required the introduction of additional criteria to classification. With the advent of causal reasoning and physiology in biology, the investigation of functional relationships within an organism became more and more important. This influenced not only classification but also morphology. As a consequence, essences were understood to represent real physical properties of traits that serve *functions* which are very important for the *survival* of the organism.

The problem with essences and its unsatisfactory solution

This brief introduction to the history of biological classification illustrates the important role that the concept of essence had in biological classification and morphology. With his concept of essence, Aristotle provided a means to rationally distinguish between artificial and ‘real’ kinds, thereby allowing for abstraction and generalizations over the overwhelming diversity of biological beings. Thus, it is not surprising that Aristotle’s essentialism played a very important role in biological classification ever since and that it significantly influenced morphology too. Unfortunately, already when introduced by Aristotle himself, the concept of essence was somehow ill-conceived and could not be satisfactorily clarified. This might be due to the fact that the concept of essence, which refers to specific properties of a given kind of things, has always been linked to the concepts of species and taxa, to which essences are supposed to provide diagnostic criteria. Considering the problems that biologists had and still have with agreeing upon a sound concept of species and taxa (for an overview see, e.g., Claridge *et al.* 1997, especially the contribution of Mayden therein; Wilson 1999, especially the contribution of Boyd therein; Pleijel & Rouse 2000; Wheeler & Meier 2000; Pigliucci 2003), the enduring ambiguity that accompanies the concept of essence is not very surprising.

As for the concept of species and taxa, the concept of essence brings about two problems. First, the *ontological* problem of what exactly is an essence (i.e., its theoretical definition), and second, the *epistemological* problem of how to recognize and identify essential properties (i.e., its recognition criteria). Aristotle’s conception was unclear regarding what exactly essences are ontologically, except that they were considered to be real physical properties. In order, in theory at least, to be able to distinguish essential from accidental properties, Aristotle had to assume the existence of an *invisible* ‘nucleus’, the ‘substance’, that every thing possesses and that bears essential properties. Other than that, essential properties could only be discovered through observation and comparison, guided by the classificatory method of logical division *per genus et differentiam*. Unfortunately, this procedure is ambiguous and can result in many different, mutually contradictory classifications. This does not necessarily pose any problems, as long as one does not assume the existence of a natural order, which would allow only one classification to be true and all others to be false.

The emergence of Christian hermetism did not really help to clarify the concept of essence, either. Quite contrary, on the basis of understanding fables as another type of property possessed by an organism, sharing the same epistemic status as biological properties, medieval hermetists modified the definition of essence to bear spiritual meaning. The invisible ‘nucleus’ was not merely the bearer of real physical properties anymore, but also of relationships of similarity to spiritually meaningful things and characters from the Bible or from ancient Greek texts. Since its cultural connotations are understood as essentially belonging to an organism in the same way as its biological properties, reality *cannot* be understood as existing independent of humankind and human culture (i.e., ontological *subjectivity*). Since the Bible and ancient Greek texts are considered to provide the main source for gaining knowledge, observation and empirical investigation were not sufficient anymore to identify essences, and the applicability of the concept of essences was further hampered.

The increasing exchange between Christian European and Islamic Arab culture marks the turning point in European history of science. Arab scholars introduced the idea of a natural order that exists independently from humankind and cultural knowledge. As a consequence, names and fables have to be separated from biological knowledge, because they are now considered to belong to different ontological categories. This marks the beginning of the establishment of ontological objectivity in Western science and philosophy and a break with one of the central paradigms of medieval hermetism.

The scientific revolution, with its claim of transparency and reproducibility, and Linnaeus’ new approach to taxonomy, with its focus on communicability, marks a significant step forward towards clarification of the concept of essences. By identifying the problem of subjectivity of individual observations and the necessity to clearly separate observation from explanation, methods for data production and documentation became more important. From then on, essences were considered to be real physical properties that are observable, unambiguously describable and that can be identified independent from individual morphologists. As a consequence, morphological terminology became more standardized and morphologists became aware of the critical role that language plays in objective data representation. The age of classification, with such protagonists as Jungius and Linnaeus, marks the high time of aperspectival objectivity in morphological terminology. When comparing their standards with current standards of morphological data conceptualization and documentation, the question immediately arises why the standards that they developed have not been reached ever since?

Unfortunately, in the long term, the trend of objectification of morphological terminology did not yield the expected success in biological classification, at least not in zoology. Cuvier, who regarded Linnaeus as the greatest genius in biological classification (Eigen 1997), realized that Linnaeus' sexual system fails to provide the foundation for a classification that unambiguously represents the existing natural order. As a consequence, and along with the hype that accompanied the emergence of causal reasoning, taxonomists were eager to further modify the concept of essence in classification, understanding essences to represent real physical properties of traits that serve functions that are very important for the survival of an organism. This can be interpreted as an improvement of the theoretical definition of the concept of essence, since it replaces the spiritual context of medieval hermetism and adds with function a component that is experimentally accessible to the idea of a 'nucleus'. However, comparative anatomical studies reveal that similar traits can have very different functions and equivalent functions can be fulfilled by morphologically diverse traits, suggesting that the relation between form and function is very flexible. Moreover, it is not clear how one can determine the importance of a function of a trait. In other words, recognition criteria for essential properties were still very ambiguous and, therefore, the application of the concept of essence still very problematic.

However, the idea of a hierarchical natural order of organisms allowed taxonomists to differentiate between accidental and essential properties on a *heuristic* basis *without* requiring reference to the functions of the respective traits: With the increase of comparative anatomical studies the idea develops that sameness relations between morphological traits can be differentiated into two different categories. On the one hand, there are those properties that appear to be accidentally equivalent. These properties, which are called *analogies*, occur isolated. Their sameness can be traced back to an equivalence of form and function (Rieppel 1993). On the other hand there are those properties that appear to be 'truly' identical. These properties, which are called *affinities* (Strickland 1840a, 1840b) or *homologies* (Owen 1843), can be distinguished from analogies by their occurrence in reciprocally corroborating aggregates. Starting point for this distinction was the idea that a hierarchical natural order of organisms would have to stand out because of the natural affinity of the corresponding organs (Whewell 1840). In other words, characteristic for affinities/homologies is that classifications based on different affinities/homologies tend to confirm each other by congruence.

Thus, by assuming a hierarchical natural order and, thus, the existence of real correlates for species and taxa, which in its turn provides a means to reasonably organize organisms into classes, an independent criterion for testing whether a trait represents an Aristotelian essence became available. A putative essential property can be tested against sets of other putative essential properties in terms of congruence. Obviously, for the first time in the history of the concept of essences, something like a (heuristic) recognition criterion is available. However, a satisfactory explanation for the existence of affinities/homologies was still lacking and, thus, the ontological status of essences (i.e., their theoretical definition) was still unclear. Nevertheless, the concept of affinity/homology was commonly accepted among nineteenth century comparative anatomists (Panchen 1999).

In the light of the theory of evolution, the concepts of affinity/homology and analogy experienced considerable modifications (for overviews see Hall 1994; Bock & Cardew 1999; Rieppel 1993). As a consequence, nowadays we understand (morphological) homologies as traits that share equivalent properties with one another due to common ancestry, whereas homoplasies represent traits that share equivalent properties due to other reasons, but not common ancestry (Lankester 1870). For the first time ever, with the theory of evolution the basis for an unambiguous, historically grounded theoretical definition of essence becomes available. From then on, homologous traits take in the role of essences in classification: whenever similarity between organisms is interpreted to be based on homologous traits, these traits are considered to be truly identical, providing the grounds for concluding identity of the trait bearing organisms, which at its turn establishes the identity of the corresponding species or taxon. Thus, the concept of homology provides a solution to the problem of essences, with which biologists have struggled for such a long time.

Morphological terminology and homology—the downside of the solution

Morphological methodology has been strongly influenced by the concept of essences. Thus, it is not surprising that we nowadays have to deal with a multiplicity of morphological terms that imply homology. The influence of essentialism on morphological thinking is so strong that even today most morphologists cannot imagine a morphological terminology free of homology assumptions. However, homology transcends the perceptually given by providing an explanation for the sameness of traits. If descriptions of morphological traits are based on homology assumptions, they depend on particular phylogeny hypotheses, which in their turn provide historical explanations for the perceived distribution pattern of sameness and differences of traits. As a consequence, much of morphological terminology is phylogeny-sensitive, requiring a change in terminology whenever a currently preferred phylogenetic hypothesis is replaced with another one due to new data, leading to a continuous change of meaning in many morphological terms. If morphological terminology is phylogeny-sensitive, morphological data are conceptually not clearly separated from conclusions.

Unfortunately, terminological standardization has been further hindered by lack of communication between morphological specialists of different taxa. As a consequence, morphological terminology has developed and grown independently within different taxonomic communities in the past and still does so today, with the effect that morphologists assign different terms to equivalent morphological traits or the same terms to different traits. This leads to a major problem regarding communicability of morphological data across large taxonomic ranges. Especially with respect to the comparative method, it represents a fundamental problem that morphology has to face. Unfortunately, Jungius and Linnaeus were the last biologists to successfully attempt to develop a general standard for morphological terminology. It seems as if morphology does not strive for the major achievements of the scientific revolution anymore: establishing a high degree of *aperspectival objectivity.*

What can we learn from Linnaeus

First attempts to establish a basic degree of mechanical and aperspectival objectivity in morphology can be seen in Locke's (1689/1979) distinction of primary (subject independent) and secondary (subject dependent) qualities in classification. In order to exclude some of the subjectivity that is necessarily connected to individual observations, taxonomists like Jungius, Tournefort, and Ray already excluded most secondary qualities from classificatory considerations. Furthermore, by focusing primarily on fructification traits, they also assumed that specific traits are more suitable for generating a consistent classification than others. As a consequence, the relevant area of matter for classification has been confined and restricted: to Jungius, Tournefort, and Ray not all morphological empirical phenomena were relevant to classification anymore. Distinguishing relevant from irrelevant phenomena, however, requires an epistemological criterion that goes beyond Locke's differentiation. This is where the *concept of structure* comes into play.

Linnaeus' morphological structure concept

Linnaeus' tremendous success can be traced back to four aspects of classification, to which Linnaeus made significant contributions—at least, when considering them in combination:

1) *Defining taxa* on the basis of Aristotelian definitions. Linnaeus defined taxa on the basis of five predi-

cates, which are derived from Aristotle's definition by *genus* and *differentia* (see e.g., Ereshefsky 1997; for a criticism that Linnaeus followed Aristotelian essentialism see Winsor 2003, 2006a, 2006b; for a reply see Stamos 2005):

Definition	A *statement* about necessary traits (i.e., essence)
Genus	Genus part of an Aristotelian definition, inherited from its parent taxon
Differentia	Distinguishing part of the definition
Property	The necessary traits as such (i.e., the taxon's essence)
Accidents	Typical traits that are not essential

2) *Classification* by Aristotelian logical division *per genus et differentiam*. Linnaeus offered clear and simple rules for constructing classifications. For pragmatic and other reasons, he introduced new ranks and rules for naming genera and species (Stevens 1997; Larson 1967). Linnaeus was also the first to propose a classification with a *strictly* encaptic hierarchy of non-overlapping classes, a Linnaean hierarchy, thereby significantly contributing to the theoretical advancement of biology (see Müller-Wille 1999).

3) Linnaeus' confinement to the sexual system—his *taxonomic characters*. Linnaeus owed a lot to the work of Cesalpino, Jungius, Ray, and Tournefort, who formalized morphological descriptions in botany and who already used fructification characters for botanical classification. Linnaeus was convinced that, due to pragmatic reasons (but see also Larson 1967 for his Aristotelian reasons), the various traits of a plant's sexual organs are best suited for botanical classification. Linnaeus considered them to be easy to work with, being most complex organs that incorporate many characters (31: calyx with 7 parts, corolla with 2, stamen with 3, pistil with 3, pericarp with 8, seed with 4, and receptacle with 4 parts; Atran 1990), which can be described precisely (Ereshefsky 1997). Thus, he used them as '*Property*' for defining botanical taxa.

4) Linnaeus' *taxonomic facts*—a botanical ***structure concept*.** Linnaeus described each of the 31 fructification traits according to four categories (Linnaeus's defining attributes; Linnaeus 1735, 1751; see also Larson 1967), each of which is based on a single perceptual judgment: 1) the quantity of observed elements (i.e., *numerus*), 2) their basic geometrical form (i.e., *figura*), 3) their relative size (i.e., *proportio*), and 4) their spatiotemporal distribution (i.e., *situs*). Applied to all fructification traits of a plant, one receives 31 descriptions, each of which consists of four 'values'. As a consequence, the description of fructification traits became parametrized. This not only established terminological standardization, but also a standardization of description which established a degree of aperspectival objectivity that was formerly not known to morphology and that has not been reached ever since.

The combination of these four aspects of Linnaeus' method of classification allowed plant taxonomists of his time to arrive at similar conclusions and to unambiguously communicate their morphological findings (his method failed, however, on the level of orders and classes, as Linnaeus himself had to admit; see Larson 1967).

Linnaeus *taxonomic facts* are obtained as a result of the concision and reduction of perception to four *categories* that exclusively refer to Locke's primary qualities. On the basis of observation, comparison, and *perceptual judgment*, morphologists decide which 'value' a given trait adopts. One could also say that each category poses a question that can be answered in reference to morphological investigations. A category is only applicable, and thus a trait only describable, if morphologists can unambiguously assign a specific 'value' to it. In other words, organizing morphological descriptions on the basis of these categories forces morphologists to make clear *perceptual judgments* in reference to criteria that demand mathematization or formalization, and thus always a standardization of statements about traits and their properties. The 'values' that a trait obtains in a description should be independent of the individual morphologist—the ideal would be that different morphologists assign the same 'values' to a given trait. The respective 'values' of a trait thus represent 'facts' about the trait. In combination, the four values or variables—one for each category—describe what represents the *morphological structure* of a trait.

The idea of restricting morphological descriptions to a predefined set of categories and their corresponding value-spaces represents an ingenious way to deal with the overwhelming diversity of morphological traits. The restriction and abstraction of the phenomenal field to only those phenomena that can be grasped by the structure concept not only *translates* morphological diversity into standardized and analytically accessible bits of information, but also establishes a high degree of communicability and comparability of morphological data and, therefore, a high degree of *aperspectival objectivity.*

Developing a general structure concept for morphology

What is 'structure'?

Structure, in general, can be understood as a fundamental notion covering the observation, recognition, dependencies, and stability of patterns and relationships of objects and processes. The concept of structure is as old as Western philosophy and science and provides an indispensable foundation of nearly every mode of inquiry and discovery in science, philosophy, and art (Pullan 2000). The term 'structure' evokes connotations of organization, connection, orientation, framework, and others—but it is *order,* which is most central to the concept of structure (Pullan 2000).

The set of relations between different parts and aspects of a given complex whole determines the latter's structure. Structure represents a way to conceive properties and relations of a complex whole, and without some notion of structure it would be very difficult for anybody to develop a conceptualization of something.

While a general notion of structure provides a general concept for structuring the overwhelming diversity of the phenomenal world, thereby mediating between phenomena (as representations of sense impressions) and their corresponding concepts (as representations of real objects and processes), when dealing with the world we live in, a potentially infinite plurality of specialized structure concepts have necessarily to be developed (Pullan 2000).

Basic principles for developing a structure concept

A structure concept is developed with a specific practical purpose in mind: it should facilitate in generating data of a specific type and quality that are relevant for a specific scientific discipline, research program, or investigation. In order to successfully develop such a proper structure concept, it is inevitable to understand the characteristics of high quality data for the given scientific question and theoretical framework.

Probably the most basic characteristic of data, commonly shared by most if not by all fields of empirical research, is that it documents some sort of observational experience, conducted by either a human being or by instruments and machines. However, scientifically relevant observational experiences usually involve information about properties and relations of real objects and their behaviour. Thus, on a very basic level, data represent descriptions of properties, relations, and behaviour of specific types of objects and processes. As such, they represent descriptions which are existence statements that do not only go beyond the necessarily private phenomenal world of an observer's experience, but also beyond descriptions of particular phenomena. Instead, these descriptions represent hypotheses about the existence of entities and their properties, which are based on *observational judgments*. In other words, these descriptions provide answers to questions regarding the entity's properties and relations, such as for instance what shape does the entity have; what is adjacent to it; whether it is continuous with some other entity; what is its temperature; how does it react to exposure to light.

A structure concept should provide a method for standardizing and formalizing such descriptions. It should consist of perceptual categories, which pose questions that can be answered in reference to empirical investigation, observation, and measurement. Ideally, the structure concept is formalized to a degree that it restricts the observer in what is allowed as an answer for each question posed by the structure concept. In other words, the structure concept should provide a set of empirical questions (i.e., categories) and with each question a set of ‘values’ (i.e., a ‘value-space’) that are allowed as an answer. Each ‘value-space’ is determined by a range of allowed numerical values (e.g., natural numbers), Boolean values (i.e., ‘*YES*’ or ‘*NO*’), or by a limited set of defined terms (i.e., a controlled vocabulary). Linnaeus’ sexual system, whose application was restricted to sexual organs of plants, came close to being such a formalized structure concept.

A clear and unambiguous structure concept should furthermore provide criteria for distinguishing relevant from irrelevant information. This includes discounting all information that does not meet previously specified and commonly accepted criteria for objectivity (i.e., ontological, aperspectival, and mechanist objectivity) as well as differentiating between phenomena that refer to real entities that are relevant to the ongoing investigation from those that are irrelevant. As a consequence, a specific structure concept necessarily always depends on the theoretical and methodological framework of a given investigation and is therefore always context-dependent. In other words, for different scientific purposes and different domains of matter, different structure concepts have to be developed and applied.

Foundations of a morphological structure concept

In order to develop the foundations for a general morphological structure concept, some questions have to be addressed first. The first question to be answered is what morphological data represents. As I have argued above, descriptions in form of existential statements grounded and substantiated in observation and experimentation represent empirical data. In the context of phylogenetics, many morphologists consider phylogenetic characters and character matrices to represent morphological data. Taking Bacon’s separation of empirical facts from scientific conclusions (i.e., explanatory hypotheses and theories) as a paradigm of scientific objectivity, however, phylogenetic characters cannot represent morphological data in the strict sense, since they incorporate homology hypotheses (i.e., putative character and character state homologies; Brower & Schawaroch 1996; see also Freudenstein 2005) and are therefore explanatory and not purely descriptive. Unfortunately, the documentation of morphological facts as discrete characters and character states becomes more and more popular among biologists, especially in the context of morphological data bases, and seems to become a standard for summarizing comparative morphological data (e.g., Ramírez *et al*. 2007).

Images of morphological traits, just like morphological character matrices, do not represent morphological data in the strict sense, either. An image cannot represent data since, as long as no description accompanies the image, its perception remains stuck in the necessarily subjective private phenomenal realm, which is to a large degree open to personal interpretation. Thus, only morphological descriptions qualify as morphological data in the strict sense.

The second question to be answered is what properties morphological descriptions should have in order to meet standard criteria for mechanical and aperspectival objectivity in morphology. First and foremost, morphological descriptions require a highly formalized and standardized morphological terminology. However, many morphological terms presuppose homology of traits. If the correct application of morphological terminology requires individual morphological traits to be homologous, one would have to know the phylogeny of the trait-bearing organisms before one could give traits a common name and describe them, since homology relations between traits can only be decided upon reference to a phylogeny. The phylogeny, in its turn, can only be reconstructed on the basis of data about distribution patterns of similar morphological traits, which, however, can only be documented and analyzed using morphological concepts and terminology in the first place. In other words, the problem is that if morphological terminology rests on homology assumptions, we cannot produce morphological data without knowing the homology relations beforehand, which, in its turn requires knowledge about the underlying phylogeny that we can only obtain on the basis of morphological data. Obviously, resting morphological terminology on homology assumptions inevitably leads to circular reasoning in phylogenetics. Therefore, in order to avoid circularity, it is essential that all morphological concepts that are used for morphological descriptions be defined without reference to homology relations.

Unfortunately, the notion of basing morphological terminology on homology assumptions represents the currently prevalent practice in morphology. Although it obviously violates Bacon's claim of separation of empirical facts from scientific conclusions, which was previously recognized as an epistemological hallmark regarding transparency and reproducibility of modern sciences, it nevertheless became commonly accepted. As I have argued above, it is most likely that this practice resulted from the impact of both the theory of evolution on biological thinking in general and the constraints of essentialist thinking of morphologists in particular.

In order to re-establish the high degree of aperspectival objectivity in morphology that Linnaeus reached for fructification traits, and in order to expand it to the entire structural diversity of morphological traits, it is inevitable that morphological terminology must be freed of all homology assumptions. Thus, in order to establish a high degree of comparability of morphological data, morphological terms should only represent structural kind terms, which are purely descriptive and free of evolutionary or other explanatory connotations. Furthermore, for allowing comparisons over broad taxonomic ranges, the applicability of morphological terminology should be taxon-independent in principle.

What is the structure of a morphological trait?

Considering the aforementioned criteria, the structure of a morphological trait consists of a set of properties corresponding to the trait and their particular values. Thereby, ideally, the list of possible properties and their definitions are provided by a general morphological structure concept. In order to describe the morphological trait, the morphologist only has to ask the corresponding question about each possible property in the list and study the particular morphological trait for an answer. Ideally, an answer, which takes in the form of a 'value', is chosen from a defined and controlled vocabulary or from a defined interval of numbers that refers to the specific property. As a consequence, morphological data would consist of pairs of property-value descriptions, referenced to a particular morphological trait and based on a general morphological structure concept, which in its turn provides the definitions and meanings to the terms (i.e., possible properties and their possible 'values') used in the descriptions. In other words, the structure of a morphological trait is a standardized list of all of its intrinsic and properties that are describable and relevant to a given scientific research program.

Data bases, ontologies, and data standards

The role of data bases in biology

In life sciences, the rate at which new data, especially molecular data, are generated increases exponentially, and this continuous increase requires the development of tools for easy sifting through and analyzing of large amounts of data (Brazma 2001). This is one of the reasons why data bases become more and more popular in life sciences. Some data bases, such as Pubmed, Ensembl or the UCSC Genome Browser, have already become essential resources, which are being used by many scientists on a daily basis (Stein 2003).

Besides many general data bases for molecular data, a lot of specialized data bases have been developed that are restricted to data from a specific model organism (e.g., FlyBase for *Drosophila*, flybase.bio.indiana.edu; Arabidopsis Information Resource for *Arabidopsis thaliana*, www.arabidopsis.org). Other data bases are devoted to a specific taxonomic group (e.g., Antbase, antbase.org; Fishbase, www.fishbase.org; AmphibiaWeb, amphibiaweb.org).

With their technical possibilities, including the convenient management of all kinds of different information, such as images and other media files, the mapping of for instance collection sites of specimens on global maps and satellite images, the possibility to link all sorts of entries with one another, such as information of a specimen in a morphological data base to its corresponding information in a data base of the museum where it is permanently deposited, biological data bases have the potential to significantly contribute to an increase of transparency and reproducibility of biological data and thus to an increase in objectivity of biological data in general (Vogt *in press*).

Data bases can thus provide a valuable resource for enabling detailed documentation of all relevant information regarding the generation of all kinds of particular empirical data. Thereby, every data base has to define *what* information can be uploaded by *whom* in which *format*. As a consequence, each data base develops its own standardized way of storing and presenting data, which requires the development or the adoption of a corresponding structure concept. Thus, it is not surprising that already today some data bases take on an important role in biological research practice, with the effect of significantly increasing the degree of mechanical and aperspectival objectivity within biology. Terminological problems, such as the lack of standards of gene names and spellings (Brazma 2001; Stein 2003), caused, for instance, fundamental problems with comparability of molecular data, turning the initial purpose of the development of molecular data bases upside down. This forced molecular data base developers to put a lot of effort into the development of defined and controlled vocabularies, in order to deal with these problems. As a consequence, the comparability of molecular data within data bases has significantly increased, with new and better standards of data documentation and representation becoming commonly accepted.

Morphological data bases

Within the last decade, some interesting morphological data bases became available. MorphBank (morphbank.net) is an open web repository of images for the documentation of specimens and vouchers for sharing research results in taxonomy, morphometrics, morphology, and phylogenetics. Another project, MorphoBank (http://morphobank.geongrid.org), is a GenBank-like repository for storing digital images (Pennisi 2003). It catalogues images and allows the labeling of structures on the images and the display of editable phylogenetic matrices, which are linked to images within the data base. A different project, Digital Morphology (DigiMorph, http://www.digimorph.org), is an archive of digital morphological images and 3D models.

Unfortunately, none of the aforementioned morphological data bases stores and documents morphological descriptions and, thus, morphological data in the strict sense. Instead, they focus on providing convenient tools for management of images, specimen information, and homology hypotheses in the form of character matrices. Thus, it is not surprising that none of these data bases provide a defined and formalized, taxon-independent, and homology-free morphological terminology for preparing morphological descriptions.

The Morphological Descriptions Data Base (MorphDBase, http://www.morphdbase.de) *attempts* to provide a platform for uploading different types of phenotypic information including all kinds of media files *and* morphological descriptions. These descriptions will be based on a morphological ontology (i.e., *MorphOntology*, http://www.morphdbase.de; for more information on ontologies see following paragraph), which is currently being developed and will be available in the near future.

Ontologies for standardizing structure concepts

Some biological data bases use *ontologies* (not to be mistaken with *Ontology* in philosophy, which is the study of ‘being’ or ‘existence’), which provide a defined and controlled vocabulary. An ontology consists of a vocabulary of terms with their corresponding concepts and some specifications of their meaning that are used to describe a certain reality. The concepts of an ontology are described both by their meaning and their relationship to each other (see also Bard 2003; Bard & Rhee 2004). An ontology is a formal way of representing knowledge of a particular scientific field through concepts and represents, as such, a data standard (Wang *et al.* 2005). It is based on a set of formal rules and assertions that describe the relationships between the concepts in a computer parsable form.

The Gene Ontology (GO; Gene Ontology Consortium 2006) represents a well-established ontology and probably the most commonly known within biology. GO provides a standardized, controlled vocabulary for genome annotation systems, cataloguing information about the structural and cellular location of gene products, about the processes to which these products contribute, and the functions that they fulfill (Stevens *et al.*

2000; Bard 2003). Hitherto, many data bases that manage molecular data have incorporated the GO annotation sets, such as for instance the Saccharomyces Genome Database (SGD, http://www.yeastgenome.org), FlyBase (http://flybase.bio.indiana.edu), Mouse Genome Informatics (MGI, http://www.informatics.jax.org), Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org), and other genome centers, such as for instance the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov) (Blake 2004). Unfortunately, regarding their applicability, most bio-ontologies available today are restricted to one specific model organism, with GO representing a rare exception.

An introduction to resource description framework (RDF) ontologies

An ontology has to be highly standardized and formalized in order to be applicable with description logics and utilizable for many different software applications. The Resource Description Framework (RDF, http://www.w3.org/RDF) has become the most accepted general method for modeling knowledge. RDF is a (meta-) data model and not a specific description language for metadata—it is data describing *all* kinds of web resources. In order to serialize (i.e., make it computer-parsable) RDF it requires syntax. Typically, RDF uses a defined XML syntax (Beckett 2004) or N3 (Berners-Lee 2005) and the semantics via reference to RDF Schema Language (RDFS) (Brickley 2004) or Ontology Web Language (OWL) (McGuinness & van Harmelen 2004). RDFS and OWL represent languages that are based upon RDF and offer support for machine processing and inferences (Wang *et al.* 2005).

In RDF, relationships between resources are described by connecting one resource to another through a relation, resulting in a RDF triple: '*Resource_X* **relation** *Resource_Y*'. A resource is anything that is identifiable by a uniform resource identifier (URI; e.g., a web address) reference (Manola & Miller 2004). By convention, the resource to the left of the relation is called '*Subject*', while the resource to the right is called '*Object*', and the relation '**property**' (in the remainder of this article, every '*Subject*' and '*Object*' will be written in italics while every '**property**' will be in bold font), resulting in the typical RDF triple formalism of '*Subject* **property** *Object*'. The '*Subject*' represents the object that is being described, the '**property**' specifies the relationship or property type between '*Subject*' and '*Object*', and the '*Object*' specifies the value of the property and is either another resource (i.e., a URI) or a literal string (i.e., a sequence of letters or numbers that is only stored by the computer without applying semantics to it, as for instance comments and numbers).

Each RDF triple can be modeled as a graph comprising two nodes connected by a directed arc (Fig. 2). A collection of such RDF graphs can jointly form a directed labeled graph (DLG) (Fig. 3). Such a DLG in its turn can, in theory, model most domain knowledge (Wang *et al.* 2005) and is a useful tool for analysis using graphs logics. A collection of RDF triples or graphs can be used to represent an ontology.

FIGURE 2. A RDF triple modeled as a directed labeled graph (DLG). *Subject* and *Object* represent the nodes and the '**property'** the edge that connects the nodes.

Defining concepts in RDF

Within an ontology, concepts are defined by a set of RDF triples. Ideally, all concepts are defined on the basis of Aristotelian definitions—*per genus et differentiam* (it is, however, possible to define a concept only on the basis of the 'genus' part and a specification of the concept of which it represents a specialized sub-concept, without explicitly specifying the 'differentia' part of its definition). As a consequence, specialized concepts inherit all defining triples of their more general 'parent' concepts.

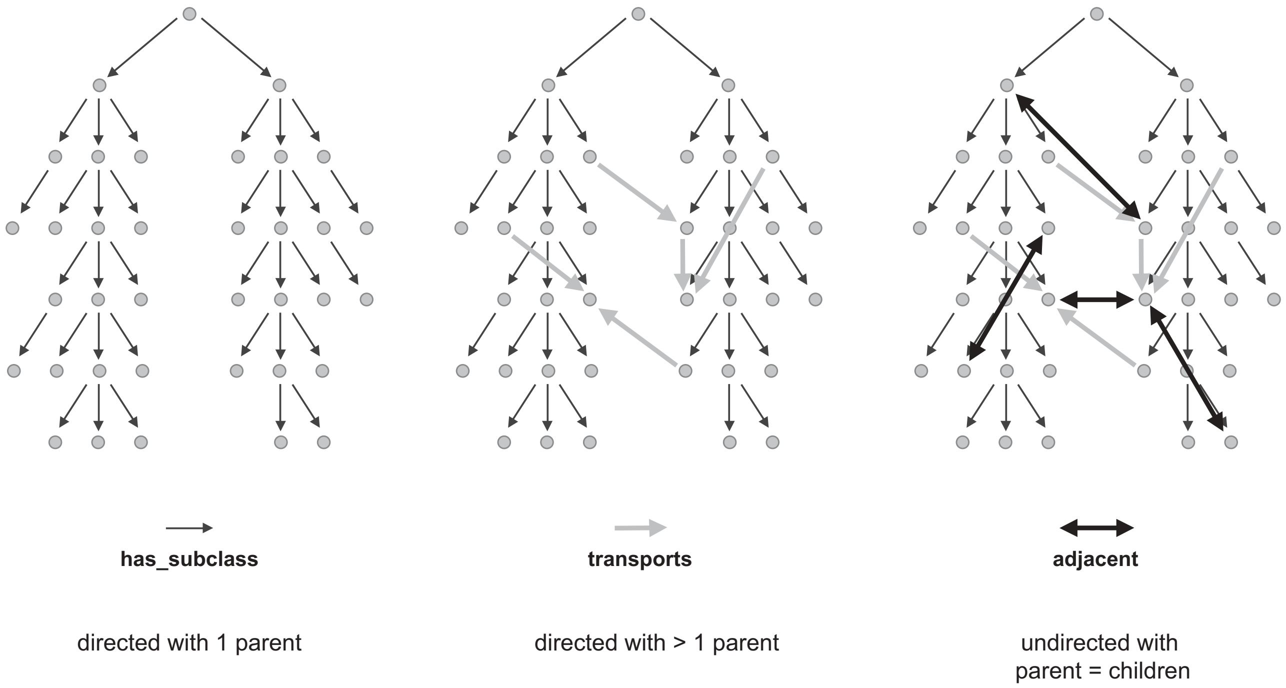

FIGURE 3. Different types of graphs. **a)** A unidirectional rule that allows only a single parent (e.g., '**has_subclass**', which is the inverse property to '**is_a'**). It can be modeled as a simple directed graph representing a tree. **b)** A unidirectional rule that allows for more than one parent (e.g., '**transports'**) can be modeled as a directed acyclic graph, in which the graph itself can be traversed in several ways, with more than one path linking two nodes. **c)** A bidirectional rule that imposes no directional constraints (e.g., '**adjacent**'), resulting in an undirected graph.

For example, a polarized junctioned cell could be defined as a junctioned cell (i.e., *genus*) that has an apico-basal orientation (i.e., *differentia*). Since it represents a specialized junctioned cell, it *necessarily* also has to possess the defining properties of junctioned cells (i.e., a junctioned cell is a cell that has an intermolecular bond with at least one cell-junction of another cell); and since junctioned cells represent a special kind of cell, a polarized junctioned cell would *necessarily* also have to possess all defining properties of cells (e.g., having as its parts a cell membrane and at least one organelle). This definition can be visualized as a graph (Fig. 4a). By organizing different property types into general categories such as topological properties versus functional properties, and by color coding them, one can also easily visually differentiate between different aspects of a definition, as well as differentially navigate through the network of relationships that exist between different concepts of an ontology by only focusing on the properties of interest and blinding out those that are not of interest.

Since most terms and concepts in an ontology should be defined through Aristotelian definitions, terms and concepts are related to one another in a network of different '**property**' relations, with a hierarchical taxonomy of class-subclass relations (i.e., '**is_a**') as a backbone, which at its turn results in a *taxonomy* of more and more specialized concepts, implying a hierarchical organization of terms (i.e. taxonomic inclusion, Bittner *et al.* 2004).

The concepts of an ontology represent classes of defined terms and their inter-relationships and should *not* contain empirical data (i.e., instances) in principle. However, statements about individual objects or individual processes can be linked as instances of concepts to the ontology. This can be done through the '**instance_of**' property. If within a data base empirical data are linked to an ontology in such a way, one receives what is called a *knowledge base* (Stevens *et al.* 2000).

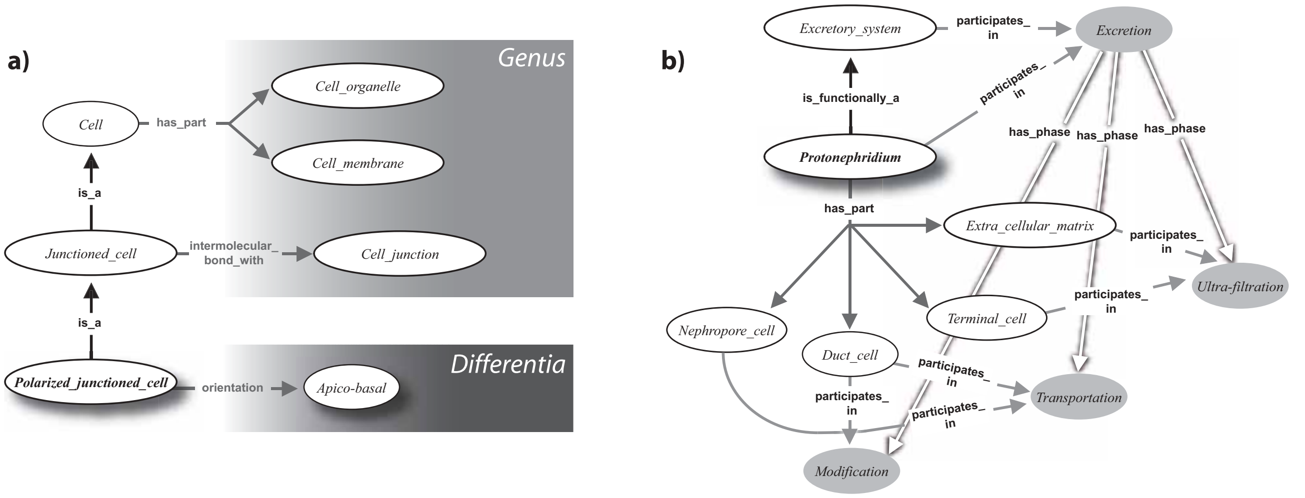

FIGURE 4. Example of two definitions of concepts expressed in RDF: **a)** the definition of 'polarized junctioned cell' (for more information see text); **b)** the definition of 'protonephridium'—a protonoephridium always consists of a nephropore cell, a duct cell, a terminal cell, and extracellular matrix. However, some structures cannot be defined satisfactorily without reference to dispositions of being able to actively participate in specific biological processes. This is also the case for protonephridium, which participates in the process of excretion. This process can be partitioned into different phases, which at their turn have different parts of the protonephridium as their participants.

Defining relations in RDF

The relations of an ontology play a very important role since they carry all the semantic content. Thus, all types of relations specified in an ontology must be carefully defined. In addition to providing free text definitions for each property of an ontology, one can define them according to their logical properties.

The '**is_a**' property, for instance, which stands for the class–subclass relationship between a specialized concept and its more general 'parent' concept, is *transitive* (if 'A_1 **is_a** A_2' AND 'A_2 **is_a** A_3', then 'A_1 **is_a** A_3'), *reflexive* ('A_1 **is_a** A_1'), and *antisymmetric* (if 'A_1 **is_a** A_2' AND 'A_2 **is_a** A_1', then A_1 and A_2 are identical), but it is not symmetric; whereas '**adjacent**', when applied as a relation between instances, is only symmetrical (if 'A_1 **adjacent** A_2', then 'A_2 **adjacent** A_1'). The '**instance_of**' property, on the other hand, is neither transitive, nor reflexive, nor symmetric, nor antisymmetric.

Further logical properties can be specified for each property of an ontology. One can define the concepts or types of literal strings (e.g., numerical values, specific intervals, Boolean values, or free text) that are allowed to be used as a possible '*Subject*' (called the *domain* of the property) and those that are allowed as a possible '*Object*' (called the *range* of the property) in a RDF triple together with this specific property. The property '**actively_participates_in**', for instance, specifies a material object that participates in a process. Thus, the domain of '**actively_participates_in**' has to be restricted to material objects only and its range to processes only. The specification of the domain and the range of each property that is defined in an ontology thereby not only constraints its applicability, but can also be utilized for enforcing, at least to some degree, logical coherence within sets of RDF triples, as for instance triples about particular relations of individual objects representing instances of concepts of the ontology.

Using an ontology for inferences

The possible applications of traditional dictionaries and glossaries, which only represent indexed sets of terms and definitions, are by far outclassed by those of an ontology. By applying descriptions logics, one can utilize the logical properties of relations defined in an ontology in order to make inferences. For instance, given the information that my left arm *A* is part of my body *X*, my left hand *B* is part of my arm *A*, and my left index finger C is part of my left hand B, appropriate software tools can infer that, due to the *transitivity* of the parthood relation, not only my left arm A is part of my body X, but also my left hand B and my left index finger C (if ‘*Arm*$_A$ **part_of** *Specimen*$_X$’ AND ‘*Hand*$_B$ **part_of** *Arm*$_A$’ AND ‘*Finger*$_C$ **part_of** *Hand*$_B$’, then ‘*Hand*$_B$ **part_of** *Specimen*$_X$’ AND ‘*Finger*$_C$ **part_of** *Specimen*$_X$’).

While this application seems to be trivial at first sight, it turns out to be invaluable when it comes to searching for relevant information within very large data bases. So for instance, when annotating the content of images using an ontology, one could annotate the information that the image depicting a complete organism also depicts its parts (e.g., its head and thorax) and all further subparts, simply by annotating that the image depicts an instance of a specific body organization. This would be enough in case this body organization is defined within the ontology as necessarily possessing a head and thorax as its parts. As a consequence, when searching for thorax within the data base, all images depicting this body organization could be retrieved as well, thereby guaranteeing that all images showing heads will be found.

Ontology as a structure concept

A structure concept requires the standardization and formalization of a specialized terminology that is required for making scientific descriptions (i.e., empirical data). Fortunately, to provide such a specialized terminology is exactly one of the key purposes of scientific ontologies. Each ‘**property’** of an ontology that refers to properties and relations of the things and processes to be described can be understood as a particular question that the structure concept poses to the scientist in reference to this given thing or process, just like Linnaeus’ categories of his sexual system (e.g., what **shape** does the entity to be described have; what is **adjacent** to it; whether it is ‘**continuous_with’** some other entity; what is its **temperature**; how does it **react** upon exposure to light). The thing or process to be described is represented by the ‘*Subject*’ in a RDF triple. The ‘*Object*’ of an RDF triple, on the other hand represents the answer to this question and specifies a specific value for the trait to be described (see Fig. 5). Actually, Linnaeus’s definitions of plants can be easily translated into RDF statements (see Fig. 6).

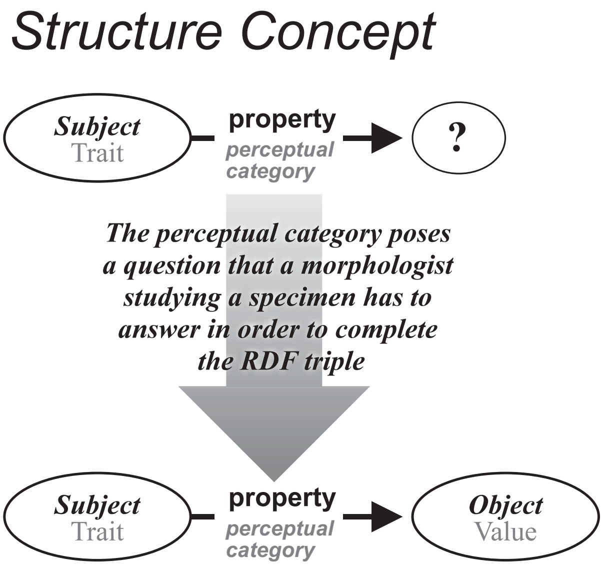

FIGURE 5. Implementation of the structure concept in RDF ontology: The trait to be described is represented by the ‘*Subject*’ of a RDF triple. The ‘**property**’ represents one perceptual category of the structure concept and functions as an empirical question that can only be answered by studying the trait. The answer to the question is represented by the ‘*Object*’ of the RDF triple and corresponds with one of the values that are allowed for this category according to the structure concept. One such describing RDF triple represents a morphological datum – the smallest piece of morphological information possible.

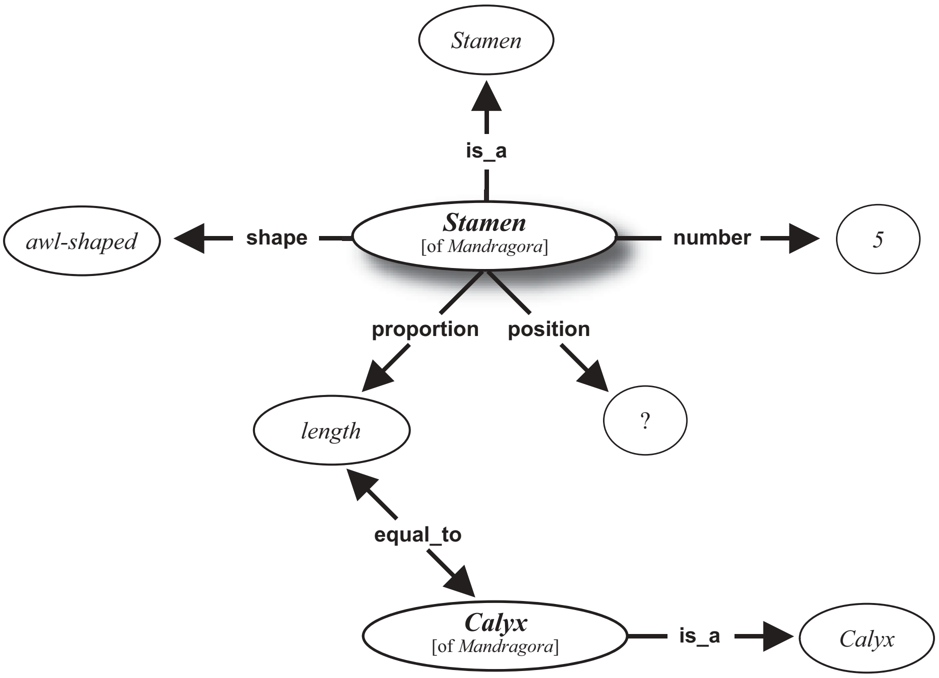

FIGURE 6. The part of the definition of the genus *Mandragora* that refers to its stamen, taken from Linnaeus's *Genera plantarum* (1737) and transformed into a RDF graph.

Which '**property**' is relevant for the description of a given entity is thereby controlled by the ontology via the specification of the *domain* and the *range* of each '**property**'. These questions can only be answered on the basis of observational judgments substantiated by experimental investigations and observations. When implementing an ontology in a data base, the advantage of both technologies is combined and the descriptions, which are based on terms and relations provided by the ontology, can be empirically substantiated by respective images from the data base.

A morphological ontology can provide a general morphological structure concept

Characteristics of biological objects

In biology, structures exist at all levels of organization, ranging hierarchically from the atomic and molecular to the cellular, tissue, organ, multicellular organism, population, and ecosystem level (see 'scalar hierarchy', Salthe 1985, 1993; 'levels of organization', Wimsatt 1976, 1994; 'cumulative constitutive hierarchy', Valentine & May 1996; 'Theorie des Schichtenbaus der Welt', Riedl 2000). Usually, a higher-level structure is composed of multiple copies of a lower-level structure. Thus, a morphological ontology has to cover all these different levels of organization, providing for each level the adequate terminology, without allowing for redundancies and inconsistencies.

Moreover, since morphological traits actively participate in specific types of processes, which is commonly understood as a property of the trait and referred to as its function, and since morphological traits also represent the result of morphogenetic processes, a morphological ontology has to cover relevant biological processes as well. The challenge here is to develop the ontology in such a way that it enables coherent representation of all the relevant inter-relationships between morphological traits and biological processes, thereby integrating structural, functional, and developmental aspects of morphological traits. This enterprise is far from being trivial.

Principles for developing a general morphological ontology

Beyond these conceptual challenges that have to be dealt with, a general morphological ontology should meet the following criteria:

- All morphological concepts should be, in principle, taxon-independent regarding their applicability. This is essential for establishing a high degree of comparability of morphological descriptions.
- A definition of a morphological kind should focus on its structural properties. This allows for unambiguous recognition of instances of the kind exclusively on the basis of morphological studies and does not require experimentation, as it would be the case with functional or developmental definitions.
- In some cases functional definitions will be inevitable, but must be clearly indicated as referring to active participation in a specific type of biological process (see example of '*Protonephridium*', Fig. 4b).
- All morphological concepts should be defined without reference to homology relations. This is essential in order to circumvent circular reasoning and is also required with respect to transparency and reproducibility of data generation.

A morphological ontology as a general morphological structure concept

The combination of a morphological ontology that meets the aforementioned criteria, imposed on a data base for morphological descriptions, would provide an integrative platform—although restricted to those particular data bases that use the ontology—within which comparative morphological studies through a broad taxonomic range would be possible in principle, since the ontology would guarantee a high degree of comparability of morphological data. In all fields in which morphological data are used, such morphological knowledge bases could take in a central methodological function comparable to GenBank for molecular data.

A premise for the success of such an approach for solving terminological problems in morphology is the development of a general structure concept for morphology. RDF ontologies, with their properties and with all their possible applications, represent the most promising tool for attempting to develop such a general morphological structure concept. Ontologies provide promising tools for the development of an easily and intuitively accessible terminology for morphology and provide a high degree of transparency of their basic underlying rules and axioms. Moreover, ontologies have the potential to provide a basis for establishing a general data standard not only for morphological data but for the entire field of biology (see Vogt *in press*), which would substantially facilitate all kinds of co-operations among the different fields in biology. Morphology as a whole would significantly benefit from this development, if it manages to participate in this already ongoing process. All it takes for its success are experienced morphologists who are willing to share their knowledge and who are willing to invest some of their time in helping to develop and improve a general morphological ontology.

Acknowledgments

I am very thankful to Prof. Alessandro Minelli for inviting me to give a talk at the meeting "*Updating the Linnaean Heritage: Names as tools for thinking about animals and plants*" of the Linnaean Society of London, held in Padova, Italy, in May 2008.

This study was supported by the German Research Foundation DFG (VO 1244/3–2). I am also grateful to the taxpayers of Germany.

References

Adanson, M. (1763) *Familles des plantes.* Vincent, Paris, 664 pp.
Aldrovandi, U. (1648) *Monstrorum historia.* Tebaldini, Bononiae, 401 pp.

Atran, S. (1990) *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge University Press, Cambridge, Mass, 360 pp.

Appel, T. A. (1987) *The Cuvier–Geoffroy Debate: French biology in the decades before Darwin*. Oxford University Press US, New York, 305 pp.

Bacon, F. (1620) *Novum organum scientiarum*. J. Billium, Londinii.

Bard, J. (2003) Ontologies: formalising biological knowledge for bioinformatics. *BioEssays*, 25, 501–506.

Bard, J. & Rhee, S.Y. (2004) Ontologies in biology: design, applications and future challenges. *Nature Reviews Genetics*, 5, 213–222.

Beck, L.W. (1969) *Early German philosophy: Kant and his predecessors*. Harvard University Press, Cambridge, Mass, 567 pp.

Beckett, D. (2004) *RDF/XML syntax specification (revised)*. W3C recommendation published online 10 February 2004. Available from: http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/ (August 2008).

Berners-Lee, T. (2005) *Primer: Getting into RDF & Semantic Web using N3*. Published online 29 June 2005. Available from: http://www.w3.org/2000/10/swap/Primer.html (August 2008).

Bittner, T., Donnelly, M. & Smith, B. (2004) Individuals, universals, collections: on the foundational relations of ontology. *In*: Varzi, A. & Vieu, L. (Eds), *Proceedings of the International Conference on Formal Ontology in Information Systems*. IOS Press, Amsterdam, pp. 37–48.

Blake, J. (2004) Bio-ontologies — fast and furious. *Nature Biotechnology*, 22, 773–774.

Bock, G.R. & Cardew, G. (eds.) (1999) *Novartis Foundation Symposium 222: Homology*. John Wiley & Sons, Chichester.

Boyd, R. (1999) Homeostasis, species, and higher taxa. *In*: Wilson, R. A. (Ed), *Species — New interdisciplinary Essays*. MIT Press, Cambridge, MA, pp. 141–185.

Boyle, R. (1660) *New experiments physico-mechanical: Touching the spring of the air and their effects*. H. Hall, Oxford.

Brazma, A. (2001) On the importance of standardisation in life sciences. *Bioinformatics*, 17, 113–114.

Brickley, D. (2004) *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C recommendation published online 10 February 2004. Available from: http://www.w3.org/TR/rdf-schema/ (August 2008).

Brower, A.V.Z. & Schawaroch, V. (1996) Three steps of homology assessment. *Cladistics*, 12, 265–272.

Cambrensis, G. (1185) *History and topography of Ireland*.

Cardano, G. (1557) *De vita propria*.[First published 1643, Villery, Paris.]

Cesalpino, A. (1583) *De plantis libri XVI*. Georgius Marescottus, Florentiae.

Claridge, M.F., Dawah, H.A. & Wilson, M.R. (1997) *Species — The units of biodiversity*. The Systematics Association Special Volume Series 54. Chapman & Hall, London, 460 pp.

Coleman, W. (1964) *Georges Cuvier, zoologist: A study in the history of evolution theory*. Harvard University Press, Cambridge, Mass, 212 pp.

Cuvier, G. (1800-5) *Leçons d'anatomie comparée*. Baudoin, Paris.

Cuvier, G. (1817) *Le règne animal; distribué d'après son organisation; pour servir de base à l'histoire naturelle des animaux et d'introduction à l'anatomie comparée. Volume I*. Deterville, Paris.

Cuvier, G. (1825) *Discours sur les révolutions de la surface du globe et sur les changements qu'elles ont produit dans le règne animal*. G. Dufours et Ed. D'Ocagne, Paris et Amsterdam.

Daston, L. (1992) Objectivity and the escape from perspective. *Social Studies of Science*, 22, 597–618.

Daston, L. (1998) Fear and loathing of the imagination in science. *Daedalus*, 127, 16–30.

Daston, L. & Galison, P. (1992) The image of objectivity. *Representations*, 40, 81–128.

Daston, L. & Park, K. (1998) *Wonders and the order of nature, 1150–1750*. Zone Books, New York, 512 pp.

Dondi, G. (1372–1374) *De fontibus calidis agri Patavini consideratio*. [Published 1553, Iunta, Venetiis.]

Drake, S. (1957) *Discoveries and opinions of Galileo*. Doubleday, New York, 320 pp.

Eco, U. (1988) *Über Spiegel und andere Phänomene*. Carl Hanser Verlag, München/Wien, 261 pp.

Eigen, E. A. (1997) Overcoming first impressions: Georges Cuvier's types. *Journal of the History of Biology*, 30, 179–209.

Ereshefsky, M. (1997) The evolution of the Linnaean hierarchy. *Biology and Philosophy*, 12, 493–519.

Farber, P.L. (1976) The type-concept in zoology during the first half of the nineteenth century. *Journal of the History of Biology*, 9, 93–119.

Freudenstein, J. V. (2005) Characters, states, and homology. *Systematic Biology*, 54(6), 965–973.

Fristrup, K.M. (2001) A history of character concepts in evolutionary biology. *In:* Wagner, G. P. (Ed), *The character concept in evolutionary biology*. Academic Press, San Diego, pp. 13–35.

Galison, P. (1998) Judgement against objectivity. *In*: Jones, C. A. & Galison, P. (Eds), *Picturing science, producing art*. Routledge, New York, London, pp. 327–359.

Gene Ontology Consortium (2006) The gene ontology (GO) project in 2006. *Nucleic Acid Research*, 34, D322–D326.

Gingrich, O. (2004) *The book nobody read: Chasing the revolutions of Nicolaus Copernicus*. Walker Publishing Company, Inc., New York, 360 pp.

Hall, B.K. (1994) *Homology — The hierarchical basis of comparative biology*. Academic Press, New York, London, 504 pp.

Heintz, B. (2000) *Die Innenwelt der Mathematik. Zur Kultur und Praxis einer beweisenden Disziplin.* Springer, Wien, 318 pp.

Heywood, V. (1985) Linnaeus — the conflict between science and scholasticism. *In*: Weinstock, J. (Ed.), *Contemporary perspectives on Linnaeus*. University Press of America, Lanham, Maryland, pp. 1–16.

Jonston, J. (1657) *Historiae naturalis de quadripedibus libri.* J.J. Schipper, Amstelodami.

Jungius, J. (1678) *Isagoge phytoscopica*. Vagetius, Hamburg.

Kanz, K.T. (2002) Von der BIOLOGIA zur Biologie – Zur Begriffsentwicklung und Disziplingenese vom 17. bis zum 20. Jahrhundert. *In:* Hoßfeld, U. & Junker, T. (Eds), *Die Entstehung biologischer Disziplinen II – Beiträge zur 10. Jahrestagung der DGGTB in Berlin 2001*, Verhandlungen zur Geschichte und Theorie der Biologie, Volume 9. VWB Verlag, Berlin, pp. 9–30.

Kukla, R. (2006) Objectivity and perspective in empirical knowledge. *Episteme: A Journal of Social Epistemology,* 3, 80–95.

Lankester, E.R. (1870) On the use of the term homology in modern zoology, and the distinction between homogenetic and homoplastic agreements. *Annals and Magazine of Natural History,* (series 4) 6, 34–43.

Larson, J.T. (1967) Linnaeus and the natural method. *Isis*, 58, 304–320.

Linnaeus, C. (1735) *Systema naturae sive Regna tria naturae. Systematice proposita per Classes, Ordines, Genera, & Species*. Theodor Haak, Lugduni Batavorum.

Linnaeus, C. (1737) *Genera plantarum.* Salvius, Holmiae.

Linnaeus, C. (1751) *Philosophia botanica*. Kiesewetter, Stockholm.

Locke, J. (1689/1979) *An essay concerning human understanding*. Oxford University Press, Oxford.

Manola, F. & Miller, E. (2004) *RDF Primer,* W3C Recommendations published online 10 February 2004. Available from: http://www.w3.org/TR/rdf-primer/ (August 2008).

Mayden, R.L. (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. *In*: Claridge, M. F., Dawah, H. A. & Wilson, M. R. (Eds), *Species: The units of biodiversity.* Chapman & Hall, London, pp. 381–424.

McGuiness, D.L. & van Harmelen, F. (2004) *OWL Web Ontology Language Overview.* W3C recommendation published online 10 February 2004. Available from: http://www.w3.org/TR/owl-features/ (August 2008).

Mery, J. (1718) *Sur un exomphale monstrueuse. HARS, Année 1716.* Paris.

Montecatini, U. da (1471) *De balneis mineralibus et artificialibus.* [Publ. 1950, Olschki, Firenze.]

Müller-Wille, S. (1999) *Botanik und weltweiter Handel. Zur Begründung eines Natürlichen Systems der Pflanzen durch Carl von Linné (1707–78).* Studien zur Theorie der Biologie, Band 3. Verlag für Wissenschaft und Bildung, Berlin.

Nagel, T. (1986) *The view from nowhere.* Oxford University Press, Oxford, 255 pp.

Newton, I. (1687) *Philosophiae naturalis principia mathematica*. Societas Regia, Londini.

Olivi, G.B. (1584) *De reconditis et praecipuis collectaneis ab onestissimo et solertissimo Francisco Calceolari Veronesi in Musaeo adservatis.* Zanfrettus, Venetiis.

Owen, R. (1843) *Lectures on comparative anatomy and physiology of the invertebrate animals delivered at the Royal College of Surgeons*. Longman, Brown, Green and Longmans, London, 689 pp.

Panchen, A.L. (1999) Homology – history of a concept. *In:* Bock, G. R. & Cardew, G. (Eds), *Homology, Novartis Foundation Symposium 222*. John Wiley & Sons, Chichester, pp. 5–18.

Pennisi, E. (2003) Modernizing the tree of life. *Science*, 300, 1692–1697.

Pigliucci, M. (2003) Species as family resemblance concepts: The (dis-)solution of the species problem? *BioEssays*, 25, 596–602.

Pleijel, F. & Rouse, G.W. (2000) Least-inclusive taxonomic unit: a new taxonomic concept for biology. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 267, 627–630.

Porter, T. (1992) Objectivity as standardization: The rhetoric of impersonality in measurement, statistics, and cost–benefit analysis. *Annals of Scholarship*, 9, 19–59.

Pullan, W. (2000) Structuring structure. *In*: Pullan, W. & Bhadeshia, H. (Eds), *Structure in science and art*. Cambridge University Press, Cambridge, pp. 1–8.

Ramírez, M.J., Coddington, J.A., Maddison, W.P., Midford, P.E., Prendini, L., Miller, J., Griswold, C.E., Hormiga, G., Sierwald, P., Scharff, N., Benjamin, S.P. & Wheeler, W.C. (2007) Linking of digital images to phylogenetic data matrices using a morphological ontology. *Systematic Biology,* 56, 283–294.

Ray, J. (1703) *Methodus plantarum emendata et aucta*. Smith and Walford, London.

Riedl, R. (2000) *Strukturen der Komplexität – Eine Morphologie des Erkennens und Erklärens*. Springer, Berlin, 367 pp.

Rieppel, O. (1993) The conceptual relationship of ontogeny, phylogeny, and classification – the taxic approach. *Evolutionary Biology*, 27, 1–32.

Russell, E.S. (1916) *Form and function: a contribution to the history of animal morphology.* John Murray, London, 394 pp.

Salthe, S.N. (1985) *Evolving hierarchical systems: their structure and representation*. Columbia University, New York, 343 pp.

Salthe, S.N. (1993) *Development and evolution: complexity and change in biology.* MIT Press, Cambridge, Mass, 368 pp.

Savonarola, M. (1448–1449) *De balneis seu thermis naturalibus omnibus Italiae.* [Publ. 1845, Andreas Belfortis, Ferrariae.]

Scotland, R.W., Olmstead, R.G. & Bennet, J.R. (2003) Phylogeny reconstruction: The role of morphology. *Systematic Biology*, 52, 539–548.

Shapin, S. (1996) *The Scientific Revolution*. University of Chicago Press, Chicago, 218 pp.

Shapin, S. & Schaffer, S. (1985) *Leviathan and the air pump: Hobbes, Boyle and the experimental life, including a translation of Hobbes, Dialogus Physicus De Natura Aeris*. Princeton University Press, Princeton NJ, 456 pp.

Spedding, J., Ellis, R.L. & Heath, D.D. (1857–1874) *The works of Francis Bacon.* 14 volumes. Longman, Green, Longman & Roberts, London.

Stamos, D.N. (2005) Pre-Darwinian taxonomy and essentialism – a reply to Mary Winsor. *Biology and Philosophy*, 20, 79–96.

Stein, L.D. (2003) Integrating biological databases. *Nature Reviews Genetics*, 4, 337–345.

Stevens, P.F. (1997) Mind, memory and history: How classifications are shaped by and through time, and some consequences. *Zoologica Scripta*, 26, 293–301.

Stevens, R., Goble, C.A. & Bechhofer, S. (2000) Ontology-based knowledge representation for bioinformatics. *Briefings in Bioinformatics*, 1, 398–414.

Strickland, H.E. (1840a) On the true method of discovering the natural system in zoology and botany. *Annals and Magazine of Natural History*, 6, 184–186.

Strickland, H.E. (1840b) Observations upon the affinities and analogies of organized beings. *Magazine of Natural History*, 4, 219–226.

Tournefort, J.P. de (1694) *Elemens de botanique*. Imprimerie Royale, Paris.

Valentine, J.W. & May, C.L. (1996) Hierarchies in biology and paleontology. *Paleobiology*, 22, 23–33.

Vogt, L. (*in press*) The future role of bio-ontologies for developing a general data standard in biology – Chance and challenge for zoo-morphology. *Zoomorphology.*

Vogt, L., Bartolomaeus, T. & Giribet, G. (*submitted*) The linguistic problem of morphology – Challenges to the future role of morphology in phylogenetics.

von Sachs, J. (1875) *Geschichte der Botanik*. Oldenburg, München.

Wang, X., Gorlitsky, R. & Almeida, J.S. (2005) From XML to RDF: how semantic web technologies will change the design of 'omic' standards. *Nature Biotechnology*, 23, 1099–1103.

Wheeler, Q.D. & Meier, R. (2000) *Species concepts and phylogenetic theory: a debate.* Columbia University Press, New York, 256 pp.

Whewell, W. (1840) *Philosophy of the inductive sciences*. J.W. Parker, London.

Wilson, R.A. (ed.) (1999) *Species — New interdisciplinary essays*. MIT Press, Cambridge, Mass, 349 pp.

Wimsatt, W.C. (1976) Reductionism, levels of organization, and the mind–body problem. *In:* Globus, G., Maxwell, G. & Savodnik, I. (Eds), *Consciousness and the brain: A scientific and philosophical inquiry.* Plenum Press, New York, pp. 202–267.

Wimsatt, W.C. (1994) The ontology of complex systems: Levels, perspectives, and causal thickets. *Canadian Journal of Philosophy*, Supplemental Volume 20, 207–274.

Winsor, M.P. (2003) Non-essentialist methods in pre-Darwinian taxonomy. *Biology and Philosophy*, 18, 387–400.

Winsor, M.P. (2006a) Linnaeus's biology was not essentialist. *Annals of the Missouri Botanical Garden*, 93, 2–7.

Winsor, M.P. (2006b) The creation of the essentialism story: an exercise in metahistory. *History and Philosophy of the Life Sciences*, 28, 149–174.