Phylogenetic Classification

Claudio Gnoli

University of Pavia. Department of Mathematics.
Library, via Ferrata 1, I-27100 Pavia, Italy, gnoli@aib.it

Claudio Gnoli has been working as a librarian since 1994, and is currently at the Mathematics department of the University of Pavia, Italy. His main interest is in theory of classification and its digital applications. He is chair of ISKO Italy chapter, and member of the Executive Committee of ISKO.

Claudio Gnoli. **Phylogenetic classification.** *Knowledge Organization, 33(3)* 138-152. 83 references.

ABSTRACT: One general principle in the construction of classification schemes is that of grouping phenomena to be classified according to their shared origin in evolution or history (*phylogenesis*). In general schemes, this idea has been applied by several classificationists in identifying a series of integrative levels, each originated from the previous ones, and using them as the main classes. In special schemes, common origin is a key principle in many domains: examples are given from the classification of climates, of organisms, and of musical instruments. Experience from these domains, however, suggests that using common origin alone, as done in cladistic taxonomy, can produce weird results, like having birds as a subclass of reptiles; while the most satisfying classifications use a well balanced mix of common origin and similarity. It is discussed how this could be applied to the development of a general classification of phenomena in an emergentist perspective, and how the resulting classification tree could be structured. Charles Bennett's notion of logical depth appears to be a promising conceptual tool for this purpose.

1. Introduction

Classification schemes are based most often on both pragmatic requirements, and some descriptive features of the classified phenomena. These criteria can satisfy the needs at hand; however, as the resulting schemes have idiosyncratic structures, they are hardly suitable to generalizations and to interoperability with other schemes. To make this possible, instead, schemes should try to follow some general principle, not only internal to their special focus, but also consistent with a general model of knowledge.

Bibliographic classifications have usually assumed as general model a segmentation of knowledge into disciplines, each with its own internal logic and subdivision. This approach tries to reflect the way that each research community classifies its own subjects of study, but has the disadvantage of treating disciplines as separated waterproof chambers, leaving little or no space for interdisciplinary relations among subjects. Facet analysis has been introduced as a tool able to give structure to schemes, by application of general categories such as personality, matter, energy, space and time; however, in most realizations of faceted classification, the categorial analysis is applied only within a discipline, and disciplines remain separated sub-universes (Gnoli 2006). The possibility of a general classification scheme independent of disciplines was explored by the Classification Research Group (CRG) (Foskett 1970).

In order to find alternative approaches to disciplinary classification, and to work on experimental non-disciplinary schemes, some general principle has to be identified according to which these can be structured. As canonical disciplines are not a guide anymore, classification criteria must come from the classified objects themselves. Two such principles have been used to various extent in many different classification schemes: that is, the overall similarity of the classified phenomena between them, and the sequence of their appearance within an evolutionary order. This paper aims to discuss the latter, as well as its possible conflicts with other principles used in classification, especially with similarity.

To this purpose, the presence of evolutionary principles in general and in special classifications will be reviewed in the next two sections. Then, a possible generalization of their use in the methodology of classification will be explored. As a result, a model will be outlined taking into account both evolutionary order and similarity as its main criteria. We will call this model *phylogenetic classification*.

2. Evolutionary principles in general classifications

The general classification schemes used in the course of history have been based on a variety of principles. However, some common patterns can be identified. In Western culture, Aristotle's tripartition into practical, poetical, and theoretical knowledge has been of great importance for a longtime. This was influential even to Francis Bacon, whose system of the sciences was divided according to the human faculties producing them, those of memory, imagination, and reason. Such a tripartition was handed down, through the French encyclopaedists and W.T. Harris, to Melvil Dewey, whose *Decimal Classification* follows that order in the sequence of its main classes (Dahlberg 1978, 29). We can label all these as epistemological systems of knowledge, as they start from the means by which humans know the world and interact with it.

An alternative approach is ontological (Poli 1996). To classify knowledge, this approach takes as its primary principle the nature of the known phenomena, independently from the means by which we know them. In this case, too, it is possible to identify a line of thinkers who have tried to arrange knowledge in a general order, usually starting from the study of the most simple objects (physics, chemistry) towards that of more and more complicated and evolved objects (biology, psychology, sociology). Grolier (1974, 30-32, 37-40) offers a wide review of these authors, including in chronological order Denis Diderot, Paul-Henry Dietrich d'Holbach, Claude-Henry de Saint-Simon, Jean-Baptiste de Lamarck, Isidore Geoffroy Saint-Hilaire, Auguste Comte, Friedrich Engels, Henri Bergson, Karl Jaspers, F. Samuel Alexander, Conwy Lloyd Morgan, Pierre Teilhard de Chardin, Joseph Needham, Nicolai Hartmann, James K. Feibleman, Ludwig von Bertalanffy, and François Jacob.

This latter approach is called "serial classification" by Bhattacharyya & Ranganathan (1974); they say that it is based on "Comte's claim that each subject is virtually an application of the preceding one," and mention André-Marie Ampère and Herbert Spencer as its followers. Ampère's sequence is slightly different, in that it places the applied disciplines just after the corresponding pure one: Physics, Engineering; Geology, Mining; Botany, Agriculture; Zoology, Animal Husbandry, Medicine; this sequence appears to have provided the core for the order of main classes in Ranganathan's *Colon Classification*.

Other researchers in library classification have also pursued the "serial" approach. James Duff Brown's *Subject Classification* is based on a sequence of matter, life, mind, and record:

> Matter, force, motion and their applications are assumed to precede life and mind, and for that reason the material side of science, with its applications, has been selected as a foundation main class on which to construct the system. Life and its forms, arising out of matters, occupy the second place among the main classes ... Human life, its varieties, physical history, disorders and recreations, follows naturally as a higher development;

and so on (Brown 1906, 12).

Ernest Cushing Richardson "may be connected as a precursor" (Grolier 1974) with the school of Alexander and Needham: indeed, he stated that "the order of the sciences is the order of things," and "the order of things is lifeless, living, human, superhuman." Classes should be arranged according to the ontological principle of increasing complexity, rather than according to some epistemological principle connected with the human mind (Richardson 1930; Foskett 1958).

Henry Evelyn Bliss devoted considerable effort to finding the order of main classes which best fits academic consensus and educational purpose. He came to arrange disciplines in a serial order of "gradation by speciality:" first come the disciplines dealing with all phenomena, like physics, followed by those dealing with more and more special phenomena, like biology, psychology, sociology, etc. Like Richardson, he claimed that the order of disciplines is related to the order of nature, which is dynamic and developmental: "this development has evidently arisen from the inorganic and has extended upward thru the biologic into the mental and the social" (Bliss 1929, 179); hence the series of disciplines also reflects the evolutionary series of natural phenomena. "The order of the sciences that is most consistent with the order of nature and the principle of gradation by speciality is at once most logical and most practical, and will prove most permanent" (Bliss 1929, 219).

Bliss's order was reused by Jack Mills, Vanda Broughton and the other editors of the second, faceted edition of the *Bliss Bibliographic Classification* (*BC2*). Interestingly, the *BC2* introduction says (Mills & Broughton 1977, 51):

> Gradation is a theoretical order of the subdisciplines of science. It correlates quite strongly with another theoretical order, that of integrative levels, which has proved of considerable value in classification theory in the last decade or so and may be said to give additional point to the theory of gradation. Integrative level theory refers to phenomena.

Actually, the helpful order of disciplines comes from an evolutionary order of phenomena: "the design of general classification schemes has tended towards following the system of integrative levels, even if this has not been acknowledged as the basis" (Foskett 1961, 145).

As Gatto (2006) observed, this is just the same situation which occurred in biology with the classification of organisms; the basic arrangement was originally established by Linnaeus, who believed that the natural groups were fixed as they were created by God; but, after Darwin's theory of evolution, the Linnean order was found to be largely corresponding to the evolutionary relationships among the groups. This means that a large part of the diversity between phenomena can be explained in terms of their origin.

While most library classifications are based on disciplines, as remarked above, research has also been done towards general schemes based on phenomena, especially by members of the CRG like Barbara Kyle, Douglas Foskett, and Derek Austin. In order to arrange phenomena in a general scheme, a classification principle is needed which be independent of the canonical sequence of disciplines. "I believe that such a principle is available; it is the theory of "integrative levels", which has been discussed in both natural and social sciences" (Foskett 1963, 132). The CRG took the theory of integrative levels in the version of Feibleman (1954), also with some influence from Needham (Gnoli & Poli 2004). Foskett (1970) and Austin (1969) worked on its possible application to a new general scheme for library classification, but produced only some drafts of it. Research in this direction has started again within the Italian chapter of ISKO (ISKO Italia 2004; Gnoli 2006; Gnoli & Hong 2006).

The idea of integrative levels is related with that of emergence (Scott 2005); indeed, each level of phenomena, though being made with parts from the lower levels, forms into a new whole, having emergent properties not present in the lower levels. Thus many of the philosophers mentioned above are also emergentists, though they differ in various details. While some of them, like Diderot, Holbach, Engels, and Needham, are materialists, others, like Jaspers, Alexander, Bergson, and Hartmann, pay most attention to the emergent properties making each level irreducible to the others, and include spirit and sometimes also divinity among the levels. Both Hartmann and Feibleman were also sources for the main classes of Dahlberg's Information Coding Classification, of which the second half are human sciences, sociology, economics, information, and culture.

In order to work as a general principle of classification, however, the ontological approach based on evolutionary relationships among phenomena should apply not only to main classes, but also to detailed subdivisions. Is this possible? How can it be done? To explore the question, in the next section we will consider some cases where an evolutionary perspective is used to classify special domains. We will include examples from the levels of inorganic phenomena (physical geography), organic phenomena (biology), and human culture (musicology).

3. Evolutionary principles in special classifications

3.1. Climates

Climatologists classify the climates of the Earth in order to describe and study them and their relations with other geographic and anthropic phenomena. As climates are a complex phenomenon, their classification is not obvious, and several systems have been proposed over time, grounded in different parameters (Strahler & Strahler 2002).

Most classical systems have been based on the temperature of the lower layers of the atmosphere. Precipitation is another parameter for which data are easily available through a wide range of geographical regions, so classificationists have pinpointed strips of similar amounts of yearly fall across the Earth. Largely dependent on climate are plant associations, so they have been suggested too as a mean to describe climates. A different system was later suggested by Terjung, based on the amount of insolation of each region, which affects the amount of energy available for plants; this can be combined with the amount of available soil water, to give a useful classification of climates. Soil water budget itself has been used by Thornthwaite as the basis for other systems, as it affects both plant growth and running water. Another system, also developed by Thornthwaite, uses precipitation and evapotranspiration, marking each climate class by an elaborated code (Thornthwaite 1948).

The most used classification system of climates is that devised by Köppen and his students (Köppen & Geiger 1936; 1954). It takes into account several of the mentioned parameters (temperature, precipitation, vegetation, etc.), combining them in an empirical way to obtain five main groups (tropical, dry, warm-temperate, nival, glacial), marked by capital letters. To them, another two letters may be added to specify other facets, like the temperatures of some seasons. As an example, most Mediterranean countries fall into class Csa, meaning warm-temperate (C), with dry summer (s), and the hottest month above 22 Celsius on average (a).

All these systems are basically descriptive and empirical. On the other hand, climatology has progressively discovered causal relations between meteorological situations and factors such as wind circulation, the origin and movements of air masses, cyclonic perturbations, and so on. This makes possible a genetic approach to the classification of climates, combining descriptive parameters with causal explanations of climatic phenomena (Oliver 1970). According to Strahler (1975), the most satisfying classification systems in natural sciences are the genetic ones, which are based primarily on the origin of phenomena. As they provide an explanation for the observed characters of classified objects, they can be considered as "explanatory;" in case the explanation is put forward largely by verbal expositions (as distinct from numerical or mathematical formulations), it can be defined as "descriptive." This approach can thus be defined as "explanatory-descriptive," and contrasted with the "empirical-quantitative" one adopted in Köppen's system.

Another application of genetical explanation in physical geography can be found in the classification of landforms, such as valleys, lakes, dunes, etc. It was Davis (1915) who introduced the systematic study of landforms according to their origin and evolution, instead of the simple descriptions of sizes, shapes, slope angles, etc. His approach has been largely followed in the subsequent works in geomorphology. Also, genesis is a key factor in the classification of soils (Buol et al. 1980).

In principle, genetic classification is the most informative. However, it requires that the causal relationships are already known at the moment when the system is designed. Ritter (2006) observes that "though atmospheric science is progressing everyday, we still have a long way to go before we have a *complete* understanding of the workings of our climate. [Genetic ones] are inherently the most difficult classifications to create and use because of the multitude of variables needed".

3.2. Organisms

Biology is one of the domains where classification has developed for the longest time. From the 16th to 18th century, naturalists like Konrad Gesner, John Ray, Joseph Tournefort, and Karl Linné (Linnaeus) were among the first actual classificationists (interestingly, Gesner was both a biologist and a bibliographer; he organized his monumental "Historia animalium" in alphabetical order). Biology deals largely with mesoscopic entities, such as mushrooms, plants and animals, which are familiar to everyone. Thus, biological examples are often used to discuss classification in general.

Modern evolutionary biology has shown how new kinds of organisms appear by modification of preexisting ones. This produces a basically tree-like model of the history of organisms, although in some cases new forms, like the eukaryotic cell, seem to have originated also by fusion and integration of simpler forms (Margulis & Sagan 2002).

It was mentioned already how Darwin's work led to the reconsideration of classification trees as evolutionary trees. Indeed, Darwin was successful in explaining the origin of species by looking at presently observable phenomena in an historical perspective. This method he also applied in other less celebrated studies, like that on coral reefs (Darwin 1842) which is still accepted today as the best explanation for their origin: the three classes of fringings, barriers, and atolls can be seen as different stages of the same historical process, produced by gradual subsidence of an oceanic island under the sea level. According to Gould (1986), three principles are available to infer evolutionary history from its results: extrapolate from directly observed processes at small time scale (like the effects of artificial selection by breeders on pigeon morphology) to processes of large time scale; arrange different observed kinds of a phenomenon (like coral reefs or stars) to form an evolutionary sequence; and use morphological imperfections and oddities (like the panda's thumb) as hints of the past history.

Hull (1998) believes that transforming Linnean structural trees into historical trees poses important problems in the representation of splitting and merging. Actually, as species evolve from older to newer forms, they cannot anymore be defined just as static entities. Some authors think that this even implies the need to abandon the Linnean nomenclature, with its imposition of a fixed number of degrees of specificity (called phylum, class, order, family, genus, and species) to the structure of hierarchies (Queiroz & Gauthier 1992; Ereshefsky 2001). Anyway, the taxonomic essence of a species can now be defined in terms of its history (Griffiths 1999).

Systematic biology has developed many technical terms to deal with the classification of organisms. Classification trees are said to be *phylogenetic*, in that they represent the historical relations in a given group (Greek *phylon)* of related organisms. The similarity among organisms in the same lineage, explained by their common origin, is called *homology*, while that among organisms in separated lineages, explained by accidental convergence, is called *analogy*. The wings of a bird and the fore legs of a mammal are homologous, as both originated from the fore limbs of their common ancestor, while the eyes of octopuses and those of birds are analogous, as they were formed through completely separate evolutionary processes.

Common origin is a preferable criterion for classification, because it is the main factor to explain the present characters. Indeed, objects sharing their origin have more fundamental and numerous characters in common, than can be guessed through a superficial inspection; and identifying them allows for more important generalizations and predictions (Mayr 1982, sections 2.0; 3.6). This is relevant both on the theoretical plane, and for practical applications (Mayr 1981):

> Biological classifications have two major objectives: to serve as a basis of biological generalizations in all sort of comparative studies and to serve as a key to an information storage system. [...] Is the classification that is soundest as a basis of generalizations also most convenient for information retrieval? This, indeed, seems to have been true in most cases I have encountered.

19th century biologist Ernst Haeckel claimed that the use of a tree metaphor (not spread in biology until then) had been suggested to him by linguists, like his colleague Schleicher. Darwin (1871) already pointed out that the evolution of languages has many analogies with that of organisms; see also Hull (1995; 2002). Actually, even in the case of languages, similarity has been progressively explained in terms of historical relationship. According to some authors, all language families are in turn related and come from a single common origin in the history of human populations (Ruhlen 1994). Some similarities, however, are due to convergence (accidental appearance of similar characteristics in two separated evolutionary lines) or to borrowing (cross-influence from a language in another non-related language). These must be recognized and discarded, in order to reconstruct the historical evolutionary tree on the basis of the remaining shared characters. According to Ruhlen (1994), convergence, borrowing, and common origin are the only three possible explanations of similarity in any kind of phenomena, such as proteins, animals, or religions.

The phylogenetic relationships among species can be reconstructed using a variety of datasets, like those coming from comparative anatomy, fossil documentation, similarity of molecular sequences in proteins and genes, etc. However, as in the case of climates, data from these sources are not available for all the groups analyzed: thus, in practice, classification trees are drawn by searching for a reasonable compromise between the observable similarities and the available knowledge of evolutionary history.

To classify large numbers of similar organisms, statistical techniques are used. In particular, numerical or phenetic taxonomy (also applied outside biology) allows one to account for a large number of characters while comparing a group of species, and to represent their similarity in form of a tree called a dendrogram (Sokal & Sneath 1963; Sneath & Sokal 1973). As they are based on statistical evaluation of observed similarities, dendrograms do not guarantee that there is always a real evolutionary relationship among organisms standing on close branches. Indeed, not all characters have the same importance as signs of the evolutionary history: we know from well-known groups of organisms that, for example, the presence of a vertebral column is much more meaningful than the biped posture.

The German entomologist Willi Hennig, then, introduced a different approach, based on a rigorous analysis of the characters actually shared by organisms with their common ancestors. This technique was originally christened "phylogenetic analysis," but has become known as cladistic taxonomy, from the term *clade* by which it refers to any monophyletic branch, i.e. any group coming from a single ancestor (Hennig 1966; Kitching et al. 1998). In this way, only organisms supposed to descend from a single ancestor are grouped in the same class in a cladogram, a tree having historical meanings not present in a dendrogram.

Cladograms may seem to be the ultimate solution in terms of evolutionary biology. However, they produce some oddities. A sensational example is that, according to cladistics, because all birds are originated from a sub-group of reptiles, birds should not form anymore a sister class of reptiles, as in traditional and common sense systematics; rather, birds are now a subclass of reptiles!

A middle way between numerical taxonomy and cladistic taxonomy is represented by so-called evolutionary taxonomy, advocated by ornithologist Ernst Mayr (1981). Evolutionary taxonomy takes into account both the evolutionary relationships among species, and their diversity. Though it is true that birds come from a sub-group of reptiles, they will form a separate class anyway, by virtue of their remarkable differentiation from their ancestors. In other terms, what is relevant is that in the same evolution period one group (classical reptiles) stayed mostly unchanged, while another (birds) changed profoundly and gave origin to a bunch of diverse forms. Radically new forms originated by such profound changes have been called *grades* by some biologists. Thus, a satisfying classification can be obtained by evaluating the ratio between the characters shared with related groups and the novelties which have occurred in one of them (Mayr 1982, section 3.4; 1990; 1995).

The longtime experience of biologists with taxonomy can be relevant to usefully suggest more general principles of classification. This applies especially to the relations and interactions between morphology and history (Gould 1986):

> Darwin was, above all, a historical methodologist. His theory taught us the importance of history, expressed in doing as the triumph of homology over other causes of order. [...] If the primacy of history is evolution's lesson for other sciences, then we should explore the consequences of valuing history as a source of law and similarity.

3.3. Musical instruments

Musical instruments are another interesting case for classification. In Western culture, they have long been classified into traditional large groups, like "string instruments" and "wind instruments," and "percussion instruments." However, it was only in the 19th century that more complete and detailed schemes were attempted, as a reaction to the need for organizing many exotic instruments that were conveyed to European museums. The first of the schemes, indeed, was published in the catalogue of the Brussels conservatory museum (Mahillon 1880).

Some years later, German-speaking ethnomusicologists Erich von Hornbostel and Curt Sachs published a great scheme which has been the main reference until the present day (Hornbostel & Sachs 1914). It retains Mahillon's main partition in four basic classes, according to the nature of the vibrant body: those of idiophones, membranophones, chordophones, and aerophones. Each main class is further subdivided into more and more specific subclasses according to different criteria: mainly morphology for chordophones and aerophones, and playing technique (percussion, scraping, plucking, friction, etc.) for idiophones and membranophones. Interestingly, Hornbostel and Sachs marked each class by a notation, taking its decimal structure from that of the *Dewey Decimal Classification* (actually they were referring to the European version of *DDC* as modified by Otlet and Lafontaine, which originated the *Universal Decimal Classification*). Indeed, they were aware of the techniques of bibliographic classifications, and thought that some of them could be applied to the classification of objects (Ghirardini & Gnoli 2005*b*).

Several classification systems of instruments have been proposed after Hornbostel-Sachs (see Ghirardini & Gnoli 2005*a*), though this has remained the most widely known and used–a situation analogous to that of the *DDC* in libraries. Among the most interesting schemes are those by Schaeffner (1932), Norlind (1939), Dräger (1947), Reinhard (1960), Hood (1971), Montagu & Burton (1971), Heyde (1975), Sakurai (1980; 1981), and Dournon (1992). Besides morphology, they try to account for several characters such as technique of execution, scale, geographical distribution, ethnic groups playing them, and cultural history. Dräger's and Hood's systems include combination techniques having a strong resemblance with the facets introduced in bibliographic classification by Ranganathan and the CRG (Kartomi 1990). In contrast, the Hornbostel-Sachs system is quite enumerative, though providing for combination of characters peculiar of some groups, like the kind and internal shape of pipes in bagpipes, and also for modification of their order by interesting notational devices (Ghirardini & Gnoli 2005*b*). Recent systems, like those proposed by Ramey (1971), Malm (1974), and Lysloff & Matson (1985), apply numerical and computer techniques, in the attempt of accounting for a wider set of characters and balancing them more objectively.

Curt Sachs, one of the authors of the Hornbostel-Sachs system, was also concerned with the history of instruments, as it developed in the various cultures all over the world and can be reconstructed through the examination of specimens collected in the present time. He developed an evolutionary theory, postulating a series of historical layers through which instruments would have passed until their present diffusion and forms (Sachs 1940). The Hornbostel-Sachs classification, as well as other systems, is based both on the present characters of instruments and on their reconstructed history. The basic definitions of the groups are morphological; e.g., chordophones are divided into *zithers* (defined as consisting of a string support which can be separated from the resonator without destroying the apparatus), *lutes* (string support connected with the resonator, and strings parallel to the sound-table), and *harps* (string support connected with the resonator, and strings perpendicular to the sound-table).

As in the case of organisms, the morphological classes of instruments can be conceived as corresponding to evolutionary relationships. Sometimes, however, this correspondence is problematic. *Lyres* are defined as yoke lutes where "the strings are attached to a yoke which lies in the same plane as the sound-table and consist of two arms and a cross-bar". Crowth, a Medieval instrument documented in iconographical sources, in its initial form fell under the definition of lyres; but later it got a neck, so that it is no more a lyre, though being the development of a lyre. A more familiar example is piano, which is classified among table zithers, as in first pianos strings were just tightened on the sound-table; however, later pianos contain a cast iron frame, on which strings are now tightened, so that strictly it should be considered as a frame zither instead. In the latter case, the genetic criterion prevails in the classification, while in the former what prevails is morphology.

A most problematic case was found with an exotic musical bow from Kenya collected at the Dinz Rialto museum in Rimini, Italy (Guizzi & Sistri 1985). Generally, musical bows are classified among zithers, the most simple kind of chordophones. However, in this case the bow acting as the string support is inserted into a cylinder which is the resonator and at the same time the bridge of the string (unless a mobile bridge also existed, which could have been lost before arriving in the museum). Hence, removing the resonator would compromise the whole apparatus, and on this basis the instrument should be classed as a lute! Some researchers claim this, so giving priority to the morphological criterion, while others claim that this bow is still a zither, so giving priority to the genetic criterion, in the same way as cladists say that birds are reptiles.

Evolutionary biologist Niles Eldredge compared the evolution of cornets in his personal collection with that of organisms, applying to cornets a software designed to classify organisms (Eldredge 2000; 2002). The resulting tree was similar to the biological ones, but also showed some differences in shape. Indeed, new characters in artifacts often come through lateral exchange, which does not occur in organisms; furthermore, they can come by sudden innovation, replacing a completely different structure, while in organisms pre-existing structures are a constraint from which change has to start.

Considerations and examples analogous to those provided for climates, organisms, and musical instruments could be developed for other special domains. However we will not continue on it anymore, as the overall picture should now be clear enough for the present purpose.

4. General phylogeny in an emergentist perspective

To summarize what we have seen in various domains, classification can be based on two major principles: similarity, and common origin. Hull (1998) calls them "structural" and "historical", and observes that the latter usually produces trees which are deeper, i.e. more structured in subclasses. According to him, "no one classification can represent both", although this is just what Mayr's evolutionary classification does.

Phylogenetic classification, in its general meaning discussed in this paper, assumes that both these principles should be used, although origin is more relevant, as it allows for more generalizations than naive classes based on similarity. Although dolphins and sharks look similar at first glance, a deeper knowledge of them reveal that dolphins actually share more basic characters with dogs than with sharks, and this is because their origin is closer to that of dogs. Such improved classification allows us to know more about them: if we consider that dolphins, like dogs, are mammals, then we will know that they have lungs, they need air to breath, they breast-feed their young, even if we can't observe these characteristics directly. Although any classification is legitimate, classifications based on both common origin and similarity are more deep and informative, and in this sense more objective.

We have seen that common origin and similarity can conflict. The cladistic criterion of using only common origin yields the weird result that birds fall among reptiles. This is unsatisfying, because we feel birds to be substantially different from reptiles, so we want our classification to reflect this, though heeding common origin. Phylogenetic classification will follow this line, by using both: birds will be listed near reptiles, as they have a common origin, but as a separate class, as they are very different. Another famous example is that of chimpanzees and men, which are very similar in molecular terms, although men have clearly diverged a lot in some key characteristics, such as the ability of using language and developing technological cultures. "A problematic entity should be included in the group with which it shares the greatest number of characteristics. If it is considered to be too dissimilar to be included in an existing group, a separate group should be established for it" (Mayr 1995).

As phenomena get more and more complex, it gets harder and harder to find two instances being identical: while two atoms of helium can be thought as identical, no pair of organisms of a given species can be said to be identical, so that the notion of individuals becomes more relevant. But this does not prevent us from grouping them in classes, and from arranging classes according to their shared characters.

It must be noted, as Bunge did (1979, ch. 1), that the generation of an object from another one does not always mean that a new class of objects originates. Usually the opposite is true: dolphins generate dolphins, and humans generate humans, so the new individuals belong to the same class of their parents. Only when the generated objects differ significantly from their "parents" is a new class originated. Then we have to find the right place to put it in our scheme, and this will depend on both the place of its parent class (origin) and how much is it different from that (similarity). This implies that we need an objective way to measure the degree of similarity among objects of different classes. On which basis can we say that birds are more different than crocodiles from the rest of reptiles, so that birds deserve a separate class while crocodiles do not? We would need an explicit method of stating it, in order to manage more specific and controversial cases.

Modern science tells us that all objects are related by causal chains, and complex objects are derived from simpler ones: birds are made of cells, cells of molecules, molecules of atoms, atoms of subatomic particles. The various degrees of complexity and derivedness, as we saw, have been called the integrative levels or grades. Some new phenomena, emerging from existing ones, can realize completely new kinds of objects, with properties not existing before: cells have a metabolism and can reproduce themselves, while molecules do not.

This makes it even more difficult to evaluate objectively the similarity of a newly emerged class with preexisting phenomena. There is not an empty place in the universe waiting in advance for phenomena of that given level to be originated and to fill it: rather, the level itself arises at the very moment when the new properties emerge in the course of the cosmical evolution. Therefore we do not have any a-priori maximum of complexity, to be labelled as "100%" in order to rank the other phenomena by comparison with it.

Let's imagine a very simple world only consisting of two dimensions (properties), like a sheet. An object in it can be described by two parameters, width and height. Then a new level emerges, having a third dimension, so that a lot of forms can develop in the new three-dimensional space: the new objects are described by three parameters, width, height, and depth. But objects in the two-dimensional world did not have depth; rather, measuring their depth had no sense at all. This is the same situation as if we asked whether molecules are alive or dead: because they belong to a level lower than that of life, concepts such as "alive" and "dead" are nonsensical. Hartmann (1942) and Lorenz (1973) would say that speaking about life properties of molecules is an ontological violation of the stratified structure of the world.

As we cannot compare a phenomenon with an absolute value of complexity, our only possibility is to compare it with already existing phenomena. As suggested by some biologists, the grade of a phenomenon can be evaluated by the ratio between its newly emerged characters (its autapomorphies, in the jargon of cladistic taxonomy) and the characters shared with its parent class (its synapomorphies). Cells have such important new characters, as compared with molecules, that they deserve a new grade: they are not just very complex aggregates of molecules, but are at a new integrative level.

It is also possible to distinguish between cases of "modest emergence", where the whole is different from its parts but still made of them; cases of more "robust emergence," like life emerging from molecules; and even cases of "radical emergence," like consciousness, where the whole is not materially made of the parts at the lower level (Scott 2005). This recalls of Hartmann's distinction between the "layers," like the physical and the chemical, which are connected through a relation of "over-forming," and the more general "strata" in which they can be grouped, namely inorganic (including physical and chemical), organic, psychic, and spiritual, which are connected through a different relation of "building-above" (Poli 2001). What seems to make strata more relevant than layers is that, at least in some cases, the connection between them is a kind of representation, that is a formal but not material correspondence: organisms are materially made of cells, hence they are just a layer within the organic stratum, while mind is not materially made of organic elements but consists in a representation of the material and organic world at an higher level, hence it is a further stratum. Indeed, Jacob (1970) observes that the major breaking points in evolution, being life and thought, both correspond to the appearance of some mechanism of memory (genetic and mental respectively).

5. Representing phylogeny in the classification structure

Thus, integrative levels are not just a simple linear series, but have various degrees of diversity, which can be represented as a hierarchical tree. This tree can be thought of as the structure of a general classification, in the same way as in special domains the evolutionary trees of phenomena can form special classifications.

So, can we use this arboreous-levelled model to outline a general classification? In principle this looks possible, and a good opportunity to arrange knowledge on the basis of unifying principles. Ideally, a general theory of classification (a taxology) should include procedural principles to establish classes and subclasses at the right degree of specificity, their sequence, and a suitable notation to express them. We can just explore the general form that such a classification system should have, leaving an obvious need for further work in order to understand its details and improve its techniques.

The kind of classification system drafted by the CRG (Austin 1969) expresses integrative levels with main classes, marked by letters of increasing value. Their subclasses can be marked with further letters (though they were digits in the original CRG draft):

```
E  atoms
F  molecules
G  bulk matter
L  cells
M  organisms
   Mp  plants
   Mq  animals
       Mqv   vertebrates
             ...
             Mqvl  reptiles
                   Mqvlo  crocodiles
             Mqvo  birds
   ...
N  populations
...
```

In the perspective of the previous section, we can ask ourselves questions like: are mammals as different from other reptiles as crocodiles are? In this case, we could place them at the same degree of specificity, namely as a subclass of reptiles, and mark them by the notation Mqvlm. Or maybe they are as different from reptiles as birds are: then, we will place them directly as a subclass of vertebrates, and will mark them Mqvm. Or they can be even more radically different, so to deserve the status of a class sister of vertebrates, Mqm. The amount of difference, together with origin, should determine the right placement of the new class in the scheme. But how can we evaluate objectively this amount of difference?

Numerical taxonomy would recommend that we describe each object by a list of characters, in order to compare them. But clearly it is not trivial to formulate a balanced list: should the character "having a vertebral column" be of the same importance as the character "be brown?" Probably not, as our general knowledge of animals suggests that external pigmentation is quite secondary with respect to the overall structure of the organisms (brown animals can be found in many classes, while having a vertebral column is a character peculiar to vertebrates).

Complexity theory has searched over the last decades for a good way of defining and measuring complexity, and this seems not an easy task. A relevant idea is to consider complexity as incompressibility of the information used to describe an object. Suppose we are describing zebra crossings: we can say that they consist of a blank strip 0.4 meters wide, then a space 0.3 meters wide, then another strip 0.4 meters wide, and so on. But we can obviously compress this information into a more synthetic form, by saying that they are a sequence of a strip and a space, repeated for the whole width of the street; so, after all, our information was not very complex. On the other hand, we cannot do the same compression in describing bar codes, as bar codes are more complex than zebra crossings (despite their dimensions).

Another side of the problem is that complexity is not always a useful criterion for placing phenomena in the phylogenetic tree of levels. Viruses are simpler than any form of life, but they cannot be ancestor of all other forms, as they need the existence of complex forms in order to act as their genetic parasites; the same is true of most parasite organisms, which are generally simpler than their ancestors. So what we want to map in our scheme is not really complexity, but something usually related to complexity.

A candidate concept is that of order. Objects in high integrative levels, like organisms or societies, are highly ordered, and thus can perform complex activities. However, crystals of quartz are even more ordered, but are very simple and prior to organisms in the emergence of integrative levels. A high integrative level is both complex and ordered (Davies 1988). The union of complexity and order can be called organization (while complexity without order is chaos). The concept of organization seems to be closer to what we are searching for. Indeed, integrative levels are sometimes called "levels of organization."

Searching for useful definitions of organization and measures of the intrinsic complexity of objects, Bennett (1987; 1988) examined as candidates several parameters from contemporary physics and information theory, including thermodynamic potentials, information content, mutual information and long-range order, and self-similarity. Eventually he introduced a new, more suitable concept, that of logical depth. Basically, the logical depth of an object is the (minimal) number of steps which have been necessary to originate it. Elephants cannot be obtained directly and instantaneously by assembling quarks: rather, evolution must follow all the steps of material and living objects, before arriving at such complex and sophisticated objects as elephants. This implies a certain time of computation, in the terms of algorithmic information theory. Highly organized things are felt by us as more valuable than simple things, just because they are possible only after a long evolutionary path has been passed through (Davies 1988).

We can imagine an original simple object: + , which can be represented by only one bit, and has logical depth $D = 1$. By adding random bits one at a time, we can get more complex objects, such as ++-+- consisting of five random bits and hence of $D = 5$. Now we can achieve a more ordered structure by moving one + towards left, so to put it together with the other +s: +++-- . This was one more step, so the new object has $D = 6$ (though still consisting of five bits, but at a more organized stage). And so on.

Now, let's imagine that we know what is the most organized deep-phenomenon in the universe, and that it has $D = x^{25}$. As we want to use the 26 letters of the Roman alphabet to classify phenomena, we assign to level Z the phenonema of $D = x^{25}$, and

level Y: $x^{25} < D \leq x^{24}$
level X: $x^{24} < D \leq x^{23}$
...
level A: $x^{0} < D \leq x^{1}$

As $x^0 = 1$, level A includes the most simple phenomenon, that consisting of only one bit.

Of course this is an ideal situation, as we do not know what exactly is the most deep phenomenon in the world, and anyway are unable to calculate its actual value of D. However, this idealization gives an idea of a possible model to follow, especially as to decide if assigning a phenomenon to the same level of its ancestor, or to a new one: it is a matter of order of magnitude. A level N will be further divided in this way:

```
N: xⁿ < D ≤ xⁿ⁺¹
    Na
    Nb
        Nba
        Nbb
            Nbba
            Nbbb
            ...
            Nbbz
        Nbc
        ...
        Nbz
    Nc
    ...
    Nz
```

We can suppose that the difference between x^n and x^{n+1} be proportionally distributed throughout Na ... Nz; that the subdifference Nc - Nb in turn be distributed throughout Nba ... Nbz; and so on. If a new organism is identified which is derived from Nbba, we should measure the increase in logical depth, in order to decide if the new organism deserves to be assigned to Nbbaa, to Nbbb, to Nbc, to Nc, or even to O. In practice we cannot measure *D* in such objective ways, so that the assignation of phenomena to the various classes and subclasses of a schema is a matter of qualitative judgment, and classification is more an art than a science, or at least has elements of both.

6. Discussion

The illustrated model was functional to explain the basic idea of logical depth and its application to classification. Of course we must be aware that most natural phenomena are very complex and consequently hard to model. However, the appearance of new properties in higher levels, like self-reproduction in living beings, can always be thought as a discrete character, being absent at the lower level and present at the higher one. The ratio between new and already existent characters could give a measure of the relevance of novelties to be reflected in classification. Thus, the model can serve at least as an auxiliary tool for a progressive formalization of the principles to determine the position of a phenomenon in the classificatory tree. Coding theory provides methods to represent a branching graph by a linear code, including techniques (Huffman's algorithm) to solve cases where a node is connected with more than one parent (Gallager 1968). For the purposes of biology, Hull (1966) proposed to represent both the systematic position and the phylogenetic history of organisms by numerical notations.

According to Mayr (1982), inanimate objects should be classified by principles different from those used in biology, because they lack any evolutionary history. This is true in the strict sense that evolution by natural selection appears only from the level of self-reproducing living beings onwards, while different integrative levels have different dynamics. However, we have seen how phylogenesis is a most informative source to classify phenomena at many levels. We just have to define phylogenesis in a broader sense: not just the inheritance of DNA variants by descent and modifications from biological ancestors, but any derivation of a phenomenon from pre-existing phenomena through a path of increasing logical depth. Indeed, Eldredge was able to apply to cornets a method designed to classify organisms.

Making phylogenetic classification does not mean assuming that the methods of biology be able to classify any phenomenon (otherwise it would be a form of biologism, just as mechanistic philosophies are a form of physicalism). It means taking the experience gained by biologists, as well as experts of other domains, in classifying phenomena according to their origin and similarity as a useful model for a more general theory of classification.

Actually, storytelling is a universal way of transmitting knowledge among humans. We like to tell stories, both fictional and real, and are able to remember them better than simple bunches of facts without any historical connection. When we meet someone that we did not know before, we like to ask her where does she come from, who are her parents, or which is her personal history, and once we have learned such things we think we "know" that person. In the same way, we like to describe non-living phenomena in terms of their origin and historical connections with other phenomena. Thus, phylogenetical knowledge is also a basic epistemological modality. It may have been selected positively, in the biological evolution of our mind, just because it is an effective way of knowing the world, as historical origin actually is an important factor explaining how things are.

As it was observed above, we often lack a complete knowledge of the historical origin of a phenomenon, and this is a pragmatic problem while constructing a phylogenetic classification. We will often have to be content with a provisional scheme of things, based more on a superficial description of them. On the other hand, once more phylogenetic knowledge will have been gained, it will be a powerful source for improvement in our classification, and we can expect that it will produce results more satisfying and stable in time. Any classification depends on the state of knowledge at the time when it is built, and phylogenetic classification makes no exception. Linking classification to the most updated evolutionary reconstruction can conflict with pragmatic needs (Gnoli 2004): in case the classification is based on a reconstruction which later is changed, it will need in turn to be modified. But at least the degree to which historical knowledge is incorporated in a scheme can work as a hint of how deep and complete it is until now. Domains of knowledge which are classified only by descriptive principles are likely to be at a preliminary stage of development (unless the use of those principles is intentional for specific purposes), while those classified by more phylogenetic trees are likely to be more mature.

Coming back to bibliographic classifications, we must remember that they classify documents, not directly phenomena. However, a previous investigation (Ghirardini & Gnoli 2005*b*) showed that manifest connections exist between the two kinds of classifications, also in their use: e.g., some music libraries succesfully use the Hornbostel-Sachs classification to arrange documents on non-Western music (Smiraglia pers. comm.). In turn, bibliographic classifications take in account the special characters of music in order to classify documents dealing with it (McKnight 2002). Szostak (2004) also argues that a classification of phenomena can be used to organize documents. As was mentioned at the beginning, most bibliographic classifications are based on disciplines, rather than directly on phenomena. The phylogenetic approach is thought as applied primarily to phenomena. According to Bliss (1929, 229), anyway, the ideal sequence of disciplines is parallel to the "natural" and to the "logical" ones, and all are basically phylogenetic. Of course this does not mean necessarily that disciplines develop historically in the same sequence: astronomy is much older than thermodynamics, though dealing with objects at an higher integrative level. A schedule of disciplines sorted in the sequence of their historical development could be useful only within the class of human knowledge, to represent their systematic position as cultural phenomena.

The joined principles of common origin and similarity, as they have been put into focus here, can be used to develop a classification scheme of phenomena in a more consistent way. The CRG modeled one such scheme according to the generic idea of integrative levels, although their draft was not developed in detail. Its resumption within the ILC project (ISKO Italia 2004) can now aim to apply the principles of phylogenetic classification to the production of more detailed schedules. At the same time, confronting these principles with actual cases of phenomena to be classified can help to individuate problems and inconsistencies, and to work on broader formulations which might be more generally appliable. For example, classes of geographic regions are usually listed and organized according to their spatial proximity, not their similarity or history; and their subclasses tend to be parts of them, rather than kinds of them. Indeed, Ranganathan (1967, ch. F) identified several different principles to be used in establishing helpful sequences of classes, with the following priority order: (1) later-in-time, (2) later-in-evolution, (3) spatial contiguity, (4) quantitative measure, (5) increasing complexity, (6) canonical sequence, (7) literary warrant, (8) alphabetical sequence. While 1, 2 and 5 are related to the phylogenetic principles discussed in this paper, and 6 to 8 concern pragmatic issues, the relation of 3 and 4 with phylogenetic classification remains to be investigated.

At the present state of classification theory, the exact class to which a phenomenon is assigned remains often a matter of common sense, and can be influenced by various practical needs. However, the phylogenetic method seems to have some potential to give a significant contribution to the development of more satisfying and generally valid classification schemes.

Acknowledgments

Work on the classifications of musical instruments, their similarities with bibliographic classifications, and examples of instruments problematic to classify was carried out together with dr. Cristina Ghirardini, PhD student in ethnomusicology at the joined Universities of Turin and Milan. Discussion on how to represent the tree of integrative levels, and the search for some mathematical formula helping to do it, developed partially with dr. Viviana Doldi, PhD in coding theory at the University of Pavia Department of Mathematics. A large part of my knowledge of evolutionary biology is owed to prof. Marco Ferraguti's class at the University of Milan. Marco Ferraguti, Roberto Poli and Rick Szostak provided valuable comments to a previous version.

References

Austin, D. 1969. Prospects for a new general classification. *Journal of librarianship* 1: 149-69.

Bennett, C.H. 1987. Dissipation, information, computational complexity and the definition of organization. In Pines, D. ed. *Emerging syntheses in science*. Boston: Addison-Wesley, pp. 215-31. Also <http://researchweb.watson.ibm.com/people/b/bennetc/bennettc19876b2b7460.pdf>.

Bennett, C.H. 1988. Logical depth and physical complexity. In Herken, R. ed. *The universal Turing machine: a half-century survey*. Oxford: Oxford University Press, pp. 227-57. Also <http://researchweb.watson.ibm.com/people/b/bennetc/UTMX.pdf>.

Bhattacharyya, G. & Ranganathan, S.R. 1974. From knowledge classification to library classification. In Wojciechowski, J.A. ed. *Conceptual basis of the classification of knowledge: proceedings of the Ot-*

tawa conference. *Pullach bei München: Saur*, pp. 119-43.

Bliss, H.E. 1929. *The organization of knowledge and the system of the sciences*. New York: Holt.

Bunge, M. 1979. *Ontology 2: a world of systems.* Dordrecht, Boston, London: Reidel.

Buol, S.W., Hole, F.D. & McCracken, R.J. 1980. *Soil genesis and soil classification*. 2nd ed. Ames: Iowa State University Press.

Dahlberg, I. 1978. *Ontical structures and universal classification.* Bangalore: SRELS.

Darwin, C. 1842. *The structure and distribution of coral reefs*. London: Smith Elder.

Darwin, C. 1871. *The descent of Man*. London: Murray.

Davies, P. 1988. *The cosmic blueprint*. New York: Simon & Schuster.

Davis, W.M. 1915. The principles of geographical description. *Annals of the Association of American Geographers* 5: 61-105.

Dournon, G. 1992. Organology. In Myers, E. ed. *Ethnomusicology: an introduction.* London, Houndmills: Macmillan, pp. 245-300.

Dräger, H.H. 1947. *Prinzip einer Systematik der Musikinstrumente*. Kassel, Basel: Bärenreiter.

Eldredge, N. 2000. Biological and material culture evolution: are there any true parallels? *Perspectives in ethology* 13: 113-53.

Eldredge, N. 2002. An overview of piston-valved cornet history. *Historic Brass Society journal* 14: 337-90.

Ereshefsky, M. 2001. *The poverty of the Linnean hierarchy: a philosophical study of biological taxonomy*. Cambridge: Cambridge University Press.

Feibleman, J.K. 1954. Theory of integrative levels. *British journal for the philosophy of science* 5n17: 59-66. Republished in Chan, L.M., Richmond, P.A. & Svenonius, E. eds. *Theory of subject analysis: a sourcebook.* Littleton: Libraries Unlimited, 1985, pp. 136-42.

Foskett, D.J. 1958. *Library classification and the field of knowledge*. London: Chaucer House.

Foskett, D.J. 1961. Classification and integrative levels. In D.J. Foskett, B.I. Palmer eds. *The Sayers memorial volume*. London: Library Association, pp. 136-50. Republished in Chan, L.M., Richmond, P.A. & Svenonius, E. eds. *Theory of subject analysis: a sourcebook.* Littleton: Libraries Unlimited, 1985, pp. 210-20.

Foskett, D.J. 1963. *Classification and indexing in the social sciences*. London: Butterworths.

Foskett, D.J. 1970. *Classification for a general index language: a review of recent research by the Classification Research Group.* London: Library Association.

Gallager, R.G. 1968. *Information theory and reliable communication.* New York: Wiley.

Gatto, E. 2006. Variazione locale e comunicabilità globale. In *Classificare la documentazione locale: giornata di studio, San Giorgio di Nogaro, 17 dicembre 2005*. ISKO Italia, <http://www.iskoi.org/doc/locale4.htm>.

Ghirardini, C. & Gnoli, C. 2005a. *Bagpipes and books on bagpipes: relations between object classifications and bibliographic classifications.* ISKO Italia <http://www.iskoi.org/bagpipes/>.

Ghirardini, C. & Gnoli, C. 2005b. Zampogne e libri sulle zampogne: classificazioni diverse? *Bibliotime, n.s.* 8n3. <http://www2.spbo.unibo.it/bibliotime/num-viii-3/gnoli.htm>.

Gnoli, C. 2004. Naturalism vs. pragmatism in knowledge organization. In McIlwaine, I.C. ed. *Knowledge organization and the global information society: proceedings Eight international ISKO conference, London, July 13-16, 2004*. Würzburg: Ergon, pp. 263-68.

Gnoli, C. 2006. The meaning of facets in nondisciplinary classifications. In Budin, G., Swertz, C. & Mitgutsch, K. eds. *Knowledge organization for a global learning society: proceedings of the Ninth international ISKO conference, 4-7 July 2006, Vienna*. Würzburg: Ergon, pp. 11-18.

Gnoli, C. & Hong M. 2006. Freely faceted classification for Web-based information retrieval. *New review of hypermedia & multimedia* 12: 63-81.

Gnoli, C. & Poli, R. 2004. Levels of reality and levels of representation. *Knowledge organization,* 31: 151-60.

Gould, S.J. 1986. Evolution and the triumph of homology, or Why history matters. *American scientist* 74: 60-69.

Griffiths, P.E. 1999. Squaring the circle: natural kinds with historical essence. In Wilson, R.A. ed. *Species: new interdisciplinary essays*. Cambridge (Mass.): MIT Press, pp. 209-28.

Grolier, E. de. 1974. From knowledge classification to library classification. In Wojciechowski, J.A. ed. *Conceptual basis of the classification of knowledge: proceedings of the Ottawa conference*. Pullach bei München, pp. 20-118. English translation of excerpts in ISKO Italia. *Integrative level classification*. <http://www.iskoi.org/ilc/grolier.htm>, 2005.

Guizzi, F. & Sistri, A. eds. 1985. *Uomini e suoni: strumenti musicali del Museo arti primitive Dinz Rialto di Rimini*. Firenze: Casa Usher.

Hartmann, N. 1942. *Neue Wege der Ontologie*. Stuttgart: Kohlhammer. (English translation: *New ways of ontology*. Westport: Greenwood Press, 1952.)

Hennig, W. 1966. *Phylogenetic systematics*. English revised ed. Urbana: University of Illinois Press.

Heyde, H. 1975. *Grundlagen des natürlichen Systems der Musikinstrumente: Beiträge zur musikwissenschaftlichen Forschung in der DDR*. Leipzig: VEB.

Hood, M. 1971. *The ethnomusicologist*. New York: McGraw Hill.

Hornbostel, E.M. von & Sachs, C. 1914. Systematik der Musikinstrumente: ein Versuch. *Zeitschrift für Ethnologie* 46: 553-90. (English translation 1969 by A. Baines & K. Wachsmann, *Galpin Society Journal* 14: 3-29.)

Hull, D.L. 1966. Phylogenetic numericlature. *Systematic zoology* 15: 14-17. Also in *JSTOR* <http://www.jstor.org/>.

Hull, D.L. 1995. La filiation en biologie de l'évolution et dans l'histoire des langues. In Gayon, J. & Wunenburger, J.J. eds. *Le paradigme de la filiation*. Harmatten: Paris, pp. 99-119.

Hull, D.L. 1998. Taxonomy. In *Routledge encyclopedia of philosophy, v.* 9. London, New York: Routledge, pp. 272-76.

Hull, D.L. 2002. Species, languages and the comparative method. *Selection* 3n2:. 17-28.

ISKO Italia. 2004. *Integrative level classification: research project.* ISKO Italia <http://www.iskoi.org/ilc/>.

Jacob, F. 1970. *La logique du vivant.* Paris: Gallimard.

Kartomi, M.J. 1990. *On concepts and classifications of musical instruments.* Chicago, London: University of Chicago Press.

Kitching, I., Forey, P., Humphries, P. & Williams, D. 1998. *Cladistics: theory and practice of parsimony analysis*. 2nd ed. Oxford: Oxford University Press.

Köppen, W. & Geiger, R. eds. 1936. Das geographische System der Klimate. In *Handbuch der Klimatologie* 1 C. Berlin: Borntraeger.

Köppen, W. & Geiger, R. 1954. *Klima der Erde*. Darmstadt: Perthe.

Lorenz, K. 1973. *Die Rückseite des Spiegels: Versuch einer Naturgeschichte menschlichen Erkennens.* München: Piper. (English translation 1976: *Behind the mirror: a search for a natural history of human knowledge*. London: Methuen; Harcourt Brace: Jovanovich.)

Lysloff, R.T.A. & Matson, J. 1985. A new approach to the classification of sound-producing instruments. *Ethnomusicology* 29: 213-36.

Mahillon, V.C. 1880-1922. *Catalogue descriptif et analytique du Musée instrumental du Conservatoire royale de Musique de Bruxelles*. Gand.

Malm, V.P. 1974. A computer aid in musical instrument research. In *Studia instrumentorum musicae popularis* 3: 119-22.

Margulis, L. & Sagan, D. *Acquiring genomes: a theory of the origins of species*. Basic Books: New York, 2002.

Mayr, E. 1981. Biological classification: toward a synthesis of opposing methodologies. *Science* 214: 510-16.

Mayr, E. 1982. *The growth of biological thought: diversity, evolution, and inheritance.* Cambridge (Mass.), London: Belknap Press.

Mayr, E. 1990. A natural system of organisms. *Nature* 348: 491.

Mayr, E. 1995. Systems of ordering data. *Biology and philosophy* 10: 419-34.

McKnight, M. 2002. *Music classification systems*. Metuchen: Scarecrow.

Mills, J. & Broughton, V. 1977. Introduction and auxiliary schedules. In *Bliss bibliographic classification.* 2nd ed. London, Boston: Butterworths.

Montagu, J. & Burton, J. 1971. A proposed new classification system for musical instruments. *Ethnomusicology* 15: 49-70.

Norlind, T. 1932. "Musikinstrumentensystematik", *Svensk Tidskrift for Musikforskning* XIV: 95-123. (Translated in *Systematik der Saiteninstrumente*, Stockholm, 1936 e Hannover, 1939.)

Oliver, J.E. 1970. A genetic approach to climatic classification. *Annals of the Association of American Geographers* 60*:* 615-37.

Poli, R. 1996. Ontology for knowledge organization. In Green, R. ed. *Knowledge organization and change: proceedings of the Fourth international ISKO conference, Washington, 15-18 July 1996.* Frankfurt: Indeks, pp. 313-19.

Poli, R. 2001. The basic problem of the theory of levels of reality. *Axiomathes* 12: 261-83. Also <http://www.mitteleuropafoundation.it/Papers/RP/The%20Basic%20Problem.pdf>.

Queiroz, K. de & Gauthier, J. 1992. Phylogenetic taxonomy. *Annual review of ecology and systematics* 23: 449-80.

Ramey, M. 1971. *A classification of musical instruments for comparative study.* PhD dissertation. Los Angeles: University of California.

Ranganathan, S.R. 1967. *Prolegomena to library classification*. Bangalore: SRELS.

Reinhard, K. 1960. Beitrag zu einer neuen Systematik der Musikinstrumente. *Die Musikforschung* 13: 160-64.

Richardson, E.C. 1930. Classification, theoretical and practical. 3rd ed. New York: Wilson.

Ritter, M.E. 2006. Climate classification. In *The physical environment: an introduction to physical geography.* University of Wisconsin <http://www.uwsp.edu/geo/faculty/ritter/geog101/textbook/title_page.html>.

Ruhlen, M. 1994. *The origin of language*. New York: Wiley.

Sachs, C. 1940. *The history of musical instruments*. New York: Norton.

Sakurai, T. 1980. An outline of a new systematic classification of musical instruments. *Journal of the Japanese Musicological Society* 25: 11-21.

Sakurai, T. 1981. The classification of musical instruments reconsidered. *Kokuritsu Minzokuyaku Hakubutsukan* 6: 824-32.

Schaeffner, A. 1932. *D'une nouvelle classification méthodique des instruments de musique*. *Revue musicale* 1932: 215-31. (Republ. in *Origine des instruments de musique*. Paris, 1936.)

Scott, A. 2005. Emergence. In Scott, A. ed. *Encyclopedia of nonlinear science*. New York: Routledge, pp. 261-62.

Sneath, P.H.A. & Sokal, R.R. 1973. *Numerical taxonomy.* San Francisco: Freeman.

Sokal, R.R. & Sneath, P.H.A. 1963. *Principles of numerical taxonomy.* San Francisco: Freeman.

Strahler, A.N. 1975. *Physical geography.* 4th ed. New York: Wiley.

Strahler, A. & Strahler, A.N. 2002. *Physical geography: science and systems of human environment*. 2nd ed. New York: Wiley.

Szostak, R. 2004. *Classifying science: phenomena, data, theory, method, practice*. Berlin: Springer.

Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. *Geographical review* 38: 55-94.