



DESCRIBING NEW SPECIES

Sectional Editor: Colin Groves

Submitted: 29 March 2010, Accepted: 18 December 2010

Alain Dubois

Muséum national d'Histoire naturelle, CP 30, 25 rue Cuvier, 75005 Paris, France.
Email: sapo421@gmail.com

Abstract

Describing new species is a fundamental work for the knowledge of the endangered biodiversity of our planet, a large proportion of which is still unknown to science. To be really useful to all other comparative disciplines of biology, this work must be carried out in a professional manner. This requires following a strict methodology for the taxonomic recognition of species and for their nomenclature. The taxonomic work must be based upon actual specimens, kept in permanent collections, and on phenetic and cladistic analyses and comparisons based on their characters, attributes and relecters. Different “kinds of species” (bisexual panmictic, parthenogenetic, gynogenetic, etc.) must be distinguished and characterized. For the progress of taxonomic knowledge, revisionary works of supraspecific taxa are much more important than mere descriptions of “new species”. Descriptions and diagnoses must be carried out in a standardized manner. As for the nomenclatural methodology, taxonomists should strictly follow the rules of the *Code*, in particular regarding its three-level structure (distinguishing availability, allocation and validity of nomina), the principles of coordination, of nomenclatural foundation, of onomatophores and of priority. No new nomen should be created if an available one exists, possibly “hidden” in a synonymy, for the species recognized by modern work. More attention and care should be paid by taxonomists to the problems related to the etymology, aspect and length of nomina: for a proper communication with all other biologists and non-biologists, the latter should be short, euphonious, clearly distinct and original.

Key words: Methodology, Taxonomy, Taxa, Specimens, Descriptions, Diagnoses, Nomenclature, Nomina

PART 1 – Taxonomic methodology

The *taxonomic urgency* is the new paradigm created by the combination, in our *century of extinctions*, of the *taxonomic gap* and of the *biodiversity crisis*

(Dubois, 2010a, b). It requires an increased effort by biologists to discover in the field, collect, store in permanent collections, study, describe and name the

still unknown living species of the planet, which are counted in millions, before they are extinct. To be beyond criticism, this work must be carried out in a professional manner. As nowadays many young taxonomists have received no academic training, or individual teaching from an experienced taxonomist, it may be useful to publish some advice in this respect.

Dayrat (2005) presented some useful general comments on the discipline of taxonomy. In fact, what he calls “integrative taxonomy” is not a new way to look at our discipline, but mostly an explicit statement of some of the general concepts and methods that have long been used and that should be used by all serious taxonomists. Below will be found some additional comments and advice concerning general guidelines for taxonomic work, many of which are very simple and “basic”, as it appears that a number of working taxonomists nowadays are not following them, because the author’s research field is amphibians, examples given below are borrowed mostly from this group, but the general guidelines provided apply to all zoological groups.

Specimens

The time of typological taxonomy is over. It is impossible to properly study, characterise and describe a species on the basis of a single or just a few specimens. A series of specimens from the same population, collected randomly, is indispensable for a proper assessment of intra-sample and inter-sample variation, and thus for meaningful comparisons with other samples, permitting the discovery of significant differences between them that may be taxonomically relevant. For statistical comparisons, it is indispensable to treat males and females separately, especially in groups such as amphibians and reptiles, which often show sexual dimorphism in many or some characters. Males and females randomly collected in a single locality constitute independent samples that allow a double test of the differences between two or more samples. If similar conclusions are obtained for both sexes, this is a strong support for their validity, in the same way as similar results obtained from the repetition of an experiment in experimental scientific disciplines. For reasons of constraints in the use of (preferably nonparametric) statistical tests, the ideal sample to characterise a population is between 12 and 20 specimens of each sex, and every effort should be made to obtain at least 8 specimens of the most common sex (usually male),

as below this figure some statistical comparisons cannot be carried out (Dubois, 1984).

Collecting specimens of living species for taxonomic research is not a pure accumulation of samples like a stamp collection. Biological data are part of our assessment of the taxonomic status of organisms or collective entities of organisms. When collecting specimens, as much other information should be obtained as possible. This includes precise information on the date, hour, weather and temperature, the location including elevation (using GPS and precise maps), observations on the habitat, on the behaviour of the organisms studied, including their localisation in the habitat (distance from water or from soil, substrate, etc.) and interactions between themselves and other organisms, on their vocalisations if relevant (with recording of their calls), on their colours in life (with photographs and/or movies), and if possible on their chromosomes (as squash preparations can be made in the field). Development pattern is part of the characteristics of a species, and if possible it is useful to collect and fix specimens of the taxon at various stages (eggs, larvae in amphibians). Allocation of these stages to the species should be done either by direct evidence (eggs directly obtained from a pair of adults, or series of tadpoles collected in the field, some of which are fixed while others are raised until metamorphosis) or by indirect evidence (presence of morphological diagnostic characters, barcoding).

Before fixation of a voucher specimen (usually with formalin) for permanent storage in collection, some tissue (blood, muscle) should be obtained from it and kept deep-frozen or in ethanol. Fixation of the specimen should be done carefully, arranging it in a nice symmetrical position, with the arms and digits expanded so as to permit their easy and accurate measurement. The specimen should be carefully labelled with numbers in indelible writing or perforations on cardboard or plastic tags, either attached to the specimen by a solid thread, or placed in the container with the specimen. For specimens that cannot be labelled individually (eg. tadpoles kept in a mixed solution of ethanol and formalin), a label should always be inserted in the container, even if another label is stuck on its outside, as this external label always runs the risk, through transportation and contacts, to be erased or unglued. A very important point is that the same number should always be given to the specimen, its tissues, chromosomes, call recordings, movies, photographs, field notes and observations. Good quality

taxonomic research cannot be carried out on the basis only of observations, drawings, photographs or movies, even if DNA samples (obtained from tissues) are available, as DNA sequences do not tell us anything about the morphology and anatomy of an organism. All these side-data must be related to a complete specimen, the voucher, without which many important pieces of information cannot be obtained (Dubois & Nemésio, 2007; Dubois, 2009d).

The specimens used for the description of new species or for comparisons with them are of particular taxonomic importance. They have a double function: a taxonomic one and a nomenclatural one. From a taxonomic point of view, all the individuals used by a taxonomist to characterise, compare, describe and diagnose a new species, i.e., the *hypodigm* (Simpson, 1940, 1961), provide an objective reference material for this taxonomic act. These specimens have a function of *semaphoronts* (Hennig, 1950, 1966), i.e., of “sign-bearers”. It is on the basis of their *characters* that the taxonomic decision was taken to erect a new taxon, either for phenetic or for cladistic reasons. They should be permanently kept in a well-curated collection, as, many decades after the original description, another taxonomist may need to compare them again with other specimens, considered to belong either in the same or in a different taxon. The second function of the vouchers, the nomenclatural one, will be considered below.

Characters, attributes, relacters and analyses

The number of characters that can be observed on a specimen is virtually unlimited, as they cover all the dimensions of the semaphoront: they include morphological, anatomical, molecular, cytological, behavioural, bioacoustic and many other “intrinsic” characters that are coded by the genotype or by interaction between the latter and epigenetic factors. Additionally, some data cannot properly be called characters, as they are caused not by genetic or epigenetic factors, but by historical facts: this is the case for example of the geographical distribution of an entity, which is only in part determined by the physiological or other needs and constraints of the species, but also by purely contingent “extrinsic” factors such as the geological and ecological history of the region where it lives. Such data should rather be called *attributes* than *characters* as they do not pertain to the organisms themselves but to their history or environment.

Three main kinds of analyses useful for taxonomic research can be carried out on the basis of characters: phenetic ones (diagnostic characters), cladistic ones (apognotic characters) and relational ones (relacters).

Phenetic analyses rely on detailed descriptions of specimens and on comparisons between them. All kinds of characters can be used for such analyses. They allow us to define a taxon through a *diagnosis*, i.e., a set of diagnostic characters, unique to the taxon and distinguishing it from all other closely related taxa. They also allow some evaluation of the divergence between taxa, through measures of “distances” between them based on various sets of data. Provided some methodological precautions are taken, molecular distances can give some clues on the approximate age of the cladogenesis which has separated two lineages observed nowadays.

Cladistic analyses are based on Willi Hennig’s (1950, 1966) methodology, using the distinction between plesiomorphic and apomorphic characters. They allow us to formulate hypotheses on cladistic relationships between entities, which can be helpful for taxonomic decision. So-called “clades” mentioned in many recent publications are not “facts”, but just hypotheses about the real (but unknown) evolutionary clades, and these hypotheses often change from one analysis to another, so that the use of Ernst Mayr’s (1997) term *cladon* to designate a taxon based on such a hypothesis would be more appropriate. Taxa recognized on the basis of such a hypothesis can be defined on the basis of an *apognosis* (Dubois, 2006d), i.e., a set of uniquely derived characters not shared with other taxa, or apognotic characters. In taxonomy, such data are mostly useful at supraspecific levels. Even if phylogeographic patterns of differentiation and dispersal can be disclosed between different populations or subspecies of the same species, these patterns are not properly cladistic, because reticulate relationships between these entities, involving hybridization and gene flow, are often involved.

As a matter of fact, a crucial set of information regarding the species problem concerns hybridization phenomena. Long considered rare by zoologists, natural hybridization between different animal species is in fact a rather common phenomenon, particularly in amphibians and reptiles. Contrary to what some believe, the existence of natural hybrids between two entities, or even of a hybridization zone connecting their distributional areas, is not evidence that these

entities are the same species. The taxonomic interpretation and decision must rely on the *dynamics* of the hybrid zone. Although intermediate situations do exist, schematically two extreme situations can be opposed (Dubois, 1998). In the first, after a period of separation due to climate or geological events, two sets of populations of common origin come into contact again. During the isolation period, they may have diverged in several respects (molecules, morphology, behaviour, etc.), but still retained the ability to successfully breed together when they meet again. If crossing between them, and with their hybrids, is unrestricted, a bidirectional gene flow will become established, thus potentially leading to a widening of the hybrid zone, and progressively to re-unification and re-homogenization of the two entities as one single entity, possibly with a remnant geographic differentiation testifying both to their previous isolation and to their possible adaptation to the local conditions in force in their respective distribution areas. In such a case, both entities must be treated as the same species, sometimes subdivided into subspecies if the parapatric entities can be unambiguously recognized and diagnosed. In the second situation, although hybridization occurs when the two entities meet again, the hybrids are poorly viable, or sterile, or dominated by the two parental stocks in competition for food or breeding. In such a case, no gene flow occurs between the two entities, the hybrid zone does not expand, and the two sets of organisms must be treated taxonomically as distinct species, despite the existence of isolated hybrids or even of a wide and stable hybrid zone between them. In the long run, these two species may become sympatric over part or all of their ranges, thus being found together in the same habitats but without breeding together, an indisputable criterion for the recognition of two species as two isolated or at least *protected* gene pools, i.e., sets of individuals which are not connected by a free bidirectional gene pool, even if some exceptional hybridization may sometimes occur between them.

In such situations, the taxonomic interpretation does not rely on characters or even on attributes (e.g., geographic distribution) taken *per se*, but on the *relative distributions* of the two entities and on the *kind of relationships* between them in their contact zone. The crucial information used for the taxonomic decision is provided here by *retracters* (Dubois, 1988, 2004b), i.e., particularities of the interactions between the two sets of organisms, not of any of them taken separately. Relative

geographic distribution (sympatry, parapatry or allopatry), competition, crossability, or dynamics of the hybrid zone, are *retracters*. They play an important role in the recognition of species under the “biological species concept” (BSC) or *mayron* (Dubois, 2007a, 2008b, 2008g, 2009a), which was explained, illustrated and defended by Ernst Mayr in many of his works (eg. 1969, 1981, 1982, 1995, 1997; Mayr & Ashlock, 1991), and by his followers.

Such criteria are useful in sympatry and parapatry but useless in allopatry. Of course, because they never meet in nature, organisms belonging to allopatric entities cannot be connected by gene flow, and the *mayron* concept cannot apply to them. In such cases, for practical purposes of classification, it is justified to use “by default” another species concept, the “phylogenetic species concept” (PSC), “lineage species concept” or *simpson* (Dubois, 2007a, 2008b, 2008g, 2009a): any set of individuals that displays fixed autapomorphic characters (either molecular, morphological or other) can be viewed as a separate lineage, and afforded the status of species. The criterion of the amount of divergence (genetic, morphological or other distance) between them can also be taken into account, especially if it allows hypotheses about the age of the cladogenesis from which the two lineages originated. “Small” distances can be used to recognize the two entities as subspecies, and “large” ones as species. The choice of the threshold values in such cases is however entirely arbitrary. Although *mayrons* and *simpsons* are treated taxonomically and nomenclaturally in the same way, they are not equivalent entities, as only the first corresponds to a consistent biological model of speciation, the second being merely a practical device for biologists.

Species and kyons

It is not justified to seek and struggle for a “unified species concept”, as widely different evolutionary situations exist in nature. Trying to refer them to a single model stems from a “reductionist” approach of biology that evolutionary biologists should not support. In amphibians and reptiles, many “strange species” or *kyons* (Dubois, 2008b) are known which correspond neither to the *mayron* nor to the *simpson* concepts. These entities are of hybrid origin, they may be polyploid and unisexual, but not always. Some of these entities, the *klonons* (Dubois, 1991), are composed only of females that reproduce by parthenogenesis and have a clonal inheritance. Other entities, the *kleptons* (Dubois & Günther, 1982), depend at each generation on another taxon

for their reproduction. They are of two main kinds (Dubois, 1991, 2008b, 2009a). In *gynokleptons*, the eggs are stimulated for development by an alien sperm from another taxon (mayron or klepton) that does not enter the egg: inheritance is therefore clonal. In *zygokleptons*, a particular gametogenesis (elasopietic metameiosis, sometimes improperly called “hybridogenesis”; see Dubois, 2008b) results in the production of gametes that are homogenous for the chromosomes of one only of the original hybridized species. These gametes undergo normal fertilization by gametes from another mayron or klepton, therefore restoring in each generation the original hybrid genotype, or producing a new genotype through ploidy elevation (up to pentaploidy, i.e., 5 n chromosomes). These strange phenomena, and others not considered here, have been observed in many zoological groups worldwide, and are quite numerous in amphibians and reptiles. Klonons are well-known in several groups of lizards from Europe and South America, and in a few snakes, and are probably more widespread. Kleptons have been much studied in two amphibian groups, the European green frogs of the genus *Pelophylax* and the North-American salamanders of the genus *Ambystoma*. It is quite likely that similar (or still different) phenomena exist in other herpetological genera. This possibility should be seriously considered by taxonomists whenever in some populations they find biased sex-ratios, unusual character distributions, or mixtures of phenotypes difficult to interpret through usual models of panmictic bisexual diploid populations with Mendelian inheritance of characters.

Comparisons

It is quite rare to discover nowadays a completely new kind of organism, one that is part not only of a new species but also of a new genus, family or still higher taxon. Most new species that are now discovered and described are part of a well-known genus, subgenus or species-group. A new species cannot be characterised alone, in an “absolute” way, but only in a relative way, through detailed comparisons with all other closely related or similar species. In fact, all diagnoses and apognoses of taxa are relative to one another. Therefore, the first step before embarking on a new description is to identify the related and similar species, and then to make careful comparisons with them, using all available characters. Comparisons should always be made between *specimens*, not between specimens and descriptions. Many descriptions, especially ancient ones, are incomplete or partly wrong, and even if of good quality they may not mention characters that

may be useful for a comparison with a newly discovered species. Description of a new species requires comparison with voucher specimens of all closely related species and all species likely to be confounded with it, even if not closely related, including, but not only, the hypodigms of these species. Except in a few cases of species with striking diagnostic characters, taxonomic journals should not accept for publication any descriptions of new species that do not include a consistent list of “*Material examined*” belonging to all related or similar taxa. This often requires the describer to obtain specimens on loan from several museums or, when this is impossible, to visit these institutions. Grants for travels for this purpose are available from many museums and other institutions worldwide, and should be called upon when necessary.

Revisions

Much more than isolated descriptions of new species, taxonomy needs revisionary works dealing with species groups, genera or other taxa. Such revisions should not only provide descriptions of new taxa, but also comprehensive reviews of all the information available about all the species of the group, critical evaluation of the accepted synonymies, standardised descriptions of all taxa, identification keys, phylogenetic hypotheses about relationships within the group and with related groups, biogeographical and evolutionary interpretations, and discussions. Rather than multiplying isolated descriptions of species in various genera and families, professional taxonomists, and all serious amateur taxonomists, should specialize themselves in one or a few groups, and work regularly for years or decades on them in order to know well their morphological and other characters, the specimens that are available in various collections worldwide, the literature dealing with the group, etc. This highly specialised work will help the researcher to avoid falling into the various traps of taxonomy (misinterpretation of characters and homoplasies), and will allow the production of long-lasting works that will be useful to all biologists dealing with the organisms at stake.

Descriptions and diagnoses

Once characterised, a taxon of nomenclatural rank species (whether a mayron, a simpson or a kyon) must be described and diagnosed. Although both terms are often used interchangeably, they apply to different objects: strictly speaking, a *description* only applies to a *specimen* (or, provided some precautions are taken, a series of specimens, eg. giving measurements as ranges, means and standard

deviations, or including variability in coloration or other characters), whereas a *diagnosis* applies to a taxon.

In any revisionary work comparing specimens or taxa, descriptions or diagnoses of different entities must use the same plan, the same characters and the same definitions of character states. Whenever describing the onymophoronts of several species of the same genus or family, it is recommended to number the characters and describe them using the same terminology for all taxa. The exact same term must be used in different descriptions for different specimens or taxa showing the same character state. A given extension of webbing cannot be described as “half-webbed” in one specimen and “incomplete webbing” in another one. Standard descriptive terms must be used in all cases. A specimen cannot be described as “small” if it appears small compared to the other ones of its genus, and “large” if it appears large compared to those of another genus: categories of sizes must be quantitatively defined and used consistently. An important point is that any character used in the description of one species must also be used in the description of another, even if it is absent. It is not appropriate to state “vocal sacs lateral” for one species, and nothing for this character in another species that lacks vocal sacs: “vocal sacs absent” must be written in full, as otherwise any subsequent writer may wonder whether the absence of this mention means that the character was absent or that no observation was made about it. Characters common to all species of a group should be described only once, at the beginning, as a “common factor” of all species of the group, but any character that may have at least two character states must be described individually for each specimen, even if only a single specimen shows a difference from the others.

Identification keys are practical devices allowing allocation of specimens to taxa. They are not diagnoses. Diagnostic characters are not only an aid for the identification of taxa, but they must also provide information on the characters of all the taxa being diagnosed. In any comparison between two or more taxa, diagnostic characters should be given for both terms of the comparison, not only for one of them. It is not acceptable to define three closely related species as follows: (1) species A has a blue coloration; (2) species B breeds in April; (3) species C reaches a size of 35 cm. This imaginary example is not drastically different from some real situations found in some publications. All diagnoses of taxa being compared in a work must provide information

on the same characters for all taxa. This also applies to apognoses: if a state of character considered autapomorphic is given for a taxon, the state of character of its sister-taxon must be given, even if it is plesiomorphic. Diagnoses and apognoses can be provided together, by simply writing differently (e.g., in bold), the character states considered autapomorphic, as exemplified in the following apodiagnoses of the two genera of the frog subfamily *LEPTOLALAGINAE* as recognized by Delorme *et al.* (2006) and Dubois (2007a):

Genus *Leptobrachella* Smith, 1925

Apognosis and diagnosis. – (1) **Adult of very small size**; (2) glands on eyelids absent; (3) axillary glands small and tit-like; (4) **finger II much longer than finger I**; (5) **tips of digits with lanceolate disks**; (6) supernumerary tubercles absent under toes and fingers; (7) **rostrodont of tadpole weakly keratinised, flexible**; (8) keratodonts of tadpole absent.

Genus *Leptolalax* Dubois, 1980

Apognosis and diagnosis. – (1) Adult of small to medium size; (2) glands on eyelids present; (3) axillary glands wide and flat; (4) fingers II and I subequal; (5) tips of digits rounded; (6) **supernumerary tubercles present under toes, absent under fingers**; (7) rostrodont of tadpole very keratinised, breakable; (8) keratodonts of tadpole present.

PART 2 – Nomenclatural methodology

The Code

Taxonomists have so far described and named about 2 million species, and millions remain to be described. Given this gigantic amount of taxa, it is indispensable to have a simple, straightforward, universal and automatic nomenclatural system for their designation, so that any two taxonomists on both sides of the planet can find by themselves, immediately and unambiguously, the valid scientific name or *nomen* of a taxon without having to apply to a board, committee or court to settle “disagreements”. Such an automatic system exists: it is provided by the *International Code of Zoological Nomenclature* (Anonymous, 1999; “the Code” below), which was first implemented more than one century ago and which has the force of law for all zoologists worldwide. Proposals for alternative nomenclatural systems such as the *Phylocode* are justified neither on theoretical nor on practical grounds and would require renaming many taxa, thus diverting the time and energy of

taxonomists from their urgent work of exploration of the planet to collect and describe the unknown species before their extinction (Dubois, 2005). The *Code* is certainly not perfect and should be drastically improved to make it more useful for taxonomists (Dubois, 2008f, 2010c), but not abandoned for another system.

The *Code* is a system which, although simple in its foundations, requires from zootaxonomists some time and effort to understand its philosophy, basic principles and rules (Dubois, 2000, 2005). This is nothing special or shocking: every other scientific or technical domain requires the same commitment before starting to use it. The requests by some to make it “simpler” and “easier to use” are largely unrealistic as would be similar requests to other scientific or technical domains. But the Rules of the *Code* should be explained and taught to university students, which is still far from being the case in many institutions and countries. This would avoid the regular publication of many misinterpretations and nomenclatural mistakes.

Nominal-series and the principle of coordination

First of all, it is important to distinguish the three “groups of nomina” or better *nominal-series* recognized by the *Code*: the *species-series* (species group) for nomina of species, subspecies, and groups of species and subspecies, the *genus-series* (genus group) for nomina of genera and subgenera and the *family-series* (family group) for nomina of families, superfamilies, subfamilies, tribes and subtribes, and possibly of taxa at additional ranks. The *principle of coordination* states that, within a nominal-series, the same nomen is available for any taxon at any rank that includes its onomatophore (see below): so a genus *Rana*, if subdivided into subgenera, must include a “nominotypical” or *hyponymous* (Dubois, 2006d) subgenus *Rana*, or a subdivided family *RANIDAE* a hyponymous subfamily *RANINAE*. Such coordinated nomina or *eponyms* are not different, homonymous, nomina, but different avatars of the same nomen, having the same author, date and onomatophore.

Although based on the same principles and similar in many respects, the nomenclatural rules are quite different in the details in the three nominal-series. As we are mostly concerned here with the description of new species, we will mostly, but not only, consider the rules in the species-series.

The three-level structure of the Code

In each nominal-series, the nomenclatural process follows a three-level system, like a three-storey house where each storey must be visited before accessing to the next. This structure has been in force since the beginning of the *Code* in the early 20th century, but never clearly identified as such, a deficiency which has been responsible of many mistakes by zootaxonomists of the past and even of today. The current plan of the *Code*, which mixes these three levels in the order of its chapters, should be modified in order to make this structure much clearer to users of that book (Dubois, 2010c).

First level: nomen availability and the nomenclatural founder principle

The first level is *availability*. To be usable in zoological nomenclature, a nomen must first have been *published* in a permanent medium (paper), i.e., not online or on a CD-Rom or DVD, and *associated* with some information (diagnosis, description or “indication”) that shows that the nomen is based on a real specimen or series of specimens, and not on a “hypothetical taxon”. Nomina published without such information, or *nomina nuda*, cannot be used in zoological nomenclature. Once created, and except for a few very special cases, a nomen cannot be changed, whether by its original author or by others; this can be known as the *nomenclatural founder principle*. This principle is not recognized as such in the *Code*, but it should be so. It is mentioned, as the *first-reviser principle*, only in a particular situation, when a choice has to be made between two possibilities left open in the original publication (e.g., choice of precedence between two synonyms published simultaneously). Because of the founder principle, care should be taken at the creation of new nomina to make them short and euphonious (see below).

Second level: nomen allocation and the onomatophore principle

The second level is *allocation* of the nomen to one or several taxa. This is not made through any “definition” of the taxon (i.e., the diagnosis or description used in the previous level do not play any role in this process), but by the use of a special tool, the *onomatophore* (Simpson, 1940, 1961) or “name-bearing type”. The use of the term “type” and other terms including this root should in my opinion progressively be extirpated from zoological nomenclature, as it carries the misleading impression that taxonomy nowadays is still “typological”, i.e., non-evolutionary (Dubois, 2005). In the species-series, onomatophores are voucher

specimens, the *onymophoronts* (Dubois, 2005). Unlike the semaphoronts used in the taxonomic process described above, the role of onymophoronts is not to *show characters* but to *bear the nomen*, i.e., to objectively *tie* it to a natural population of animals. In zoology, they can be of four kinds, *holophoronts* (called “holotypes” in the *Code*), *symphoronts* (“syntypes”), *lectophoronts* (“lectotypes”) or *neophoronts* (“neotypes”). These specimens are usually also part of the hypodigm, thus having a double function, both taxonomic and nomenclatural. In the other nominal-series, onomatophores are not specimens, but nomina, or, more exactly, “nominal taxa” or *taxomina* (Dubois, 2000, 2005).

Third level: nomen validity and the priority principle

The third level of the nomenclatural process is *validity*. Quite often, several nomina are available for the same taxon: these are *synonyms*. Among them, only one can be valid. The basic criterion for establishing the valid nomen among synonyms is *priority of publication*. This is an *objective* and *automatic* criterion, based on a unique simple unilinear factor, *time*, and which avoids having to enter into endless discussions about “the best” among several nomina. The recent tendency to challenge priority in favour of “usage” should not be encouraged, as it amounts to the replacement of universal stringent rules, in front of which all zootaxonomists worldwide are equal, by a system of polls, pressure groups or “lobbies”, i.e., in fact to a “principle of authority”, which has nothing to do with science (Dubois, 2010b). Priority also allows the establishment of the valid nomen among several *homonyms*, i.e., nomina having the same spelling, or spellings very similar in a few special cases. Although this is not stated as such in the *Code*, the *principle of priority* is the general one, which applies to two “subprinciples”, that of *synonymy* (a given taxon cannot bear validly two different nomina), which in fact is not mentioned in full words in the *Code*, and that of *homonymy* (two different non-coordinated taxa cannot bear the same nomen).

Variety-series and class-series nomenclature

The *Code* does not regulate the nomina of higher taxa (orders, classes, phyla: *class-series* nomina) or of taxa below subspecies (varieties, forms: *variety-series* nomina). This is a real problem at low levels because the *Code* does not regulate the nomenclature of infrasubspecific entities, which is therefore left to disciplines like phylogeography and

conservation biology (Dubois, 2006b). This is also most detrimental in higher nomenclature. Many new higher taxa are currently being recognized as results of molecular phylogenetic analyses. In the absence of rules, each author feels entitled to use the nomina he likes for the taxa he recognizes, often without regard for their original contents, thus in full contradiction with the untold nomenclatural founder principle. One hopes that this problem will be solved soon, before complete chaos is installed in higher nomenclature (Dubois, 2006a, 2006c, 2008f, 2009b-c).

In the meanwhile, taxonomists should at least try to follow two “implicit rules”, not mentioned in the *Code*, but which directly derive from the rules of this text and are just a matter of common sense. First, nomina of the class-series should not be liable to be confounded with nomina of the family-series (e.g., SALAMANDROIDEA) or, even worse, of the genus-series (TERRARANA): such ill-formed nomina should be emended, a change that any taxonomist is entitled to do since the *Code* does not exclude this possibility. Second, nomina similar to those of one nominal-series should not be used for taxa subordinate to taxa belonging to a lower nominal-series, such as “Cathopliogenici” (a nominative plural) as a subdivision of the insect genus *Carabus* (see Dubois 2006b) or “Allocentrolenidae” (not based on a generic substantive) as a subdivision of the frog superfamily *HYLOIDEA* (see Guayasamin *et al.* 2009). Such nomenclatural practices are not compatible with the *Code* as they entail an overlap between nominal-series or at least a lack of respect for the coordination principle, thus precluding the use of nominal-series for a transcription of hypothesised phylogenetic relationships in the form of nomenclatural hierarchies (see Dubois, 2007a, 2008g). Such nomina should therefore be considered nomenclaturally unavailable, and taxonomists should be encouraged to replace them by nomina compatible with the rules of the *Code* and in particular with the principle of coordination.

Revalidation of synonyms

Whenever describing a new species, any taxonomist should first enquire into the possibility that a nomen is already available for this taxon, but “hidden” in a synonymy, which is quite frequent. In such a case, rather than coining a new nomen, he should “resurrect” that already available nomen, i.e., remove it from the synonymy where it had been placed by mistake. Otherwise, subsequent taxonomists will do this sooner or later, which will entail nomenclatural instability. This requirement

completes the statement above that the description of a new species cannot be an isolated fact, but must be done in the context of comparisons with all the related and similar species. Ascertaining the status of synonyms cannot rely only on their descriptions, especially if these are ancient, as publication standards have much improved in the recent decades. It requires examination of the onymophoront(s) of this nomen. If this is impossible, specimens from the same locality (*onymotope*, usually called “type locality”) as the latter, or *topophoronts*, can be used by default. This is because the function of a nomen-bearing specimen is not to provide characters, but to tie the nomen to a natural biological population. It is among the duties of referees and the editor of a taxonomic paper proposing a new species nomen to check that the author(s) indeed provided evidence that no nomen for this taxon already exists in the literature, which requires checking the validity of the synonymies of all the related and similar species.

Etymology, aspect and length of nomina

This point is rarely discussed in taxonomic papers and books, because most taxonomists consider themselves as the creators, and often the “owners”, of the new nomina they publish, and demand in fact “full freedom” for the aspect and length of these nomina, feeling very “shocked” if someone questions their choice. However, the nomina they publish are not only for their personal pleasure and use, but must be at the service of a discipline (taxonomy), of a community (taxonomists) and in fact of all users of scientific nomina, which in some cases (laymen, farmers, lawyers, custom officers, etc.) are very far from biology. Some guidelines at least, if not rules, should exist regarding the characteristics of new nomina. As this is a poorly explored domain, I will here devote a rather detailed discussion to it. These seldom discussed questions were posed in a recent paper (Dubois & Raffaëlli, 2009), part of which is summarised below.

Scientific nomina are not an aim in themselves, but *tools* that are used in various contexts. They are not descriptions, diagnoses, statements on the characters, distribution or other characterizations of the taxa they designate. They are not models, evolutionary, phylogenetic or genetic theories about the hypothesized origin of these taxa. They are not paeans of praise for their authors, for the discoverers of the taxa or for the persons to whom they may be dedicated. They are just neutral *labels* meant at designating *unambiguously* and *universally*

a given taxon *within the frame of a given taxonomy*, i.e., allowing the automatic designation of the taxa recognized by taxonomists at a given stage of their research. As repeatedly stated by Ernst Mayr and others, these labels allow storage and retrieval of the information accumulated in taxonomies, but they are not intended to express this information by themselves. As such, nomina are completely *meaningless* and should remain so. This is why the *Code* expressly states that availability of nomina “*is not affected by inappropriateness*”, and allows a new generic or specific nomen to be “empty of meaning”, as for example they can be “*an arbitrary combination of letters provided this is formed to be used as a word*”. It is certainly praiseworthy for an author to have taken care for a new nomen to be “full of meaning”, derived from an identified etymology, and “strictly formed” from a grammatical point of view, but this is much less important than the nomen being grammatically correct and short, euphonious in all languages and “*easily memorable*”.

Because of the nomenclatural founder and priority principles, a nomen, once created, cannot be changed by subsequent authors. If it is the first one available for the taxon it designates, it will have to be used by all authors who will deal with this taxon. Once coined, a new nomen will appear not only in taxonomic and phylogenetic publications, but also in all the scientific and non-scientific literature, in titles, official documents and lists, etc., published and distributed over the whole planet, that will deal with the organisms it designates. Whenever they are used in such non-specialized literature, long and complex nomina are certainly not good “publicity” for taxonomy, especially in these times when this scientific discipline is facing difficulties.

The *Code* provides a few guidelines for coining the specific (*epithets*) and generic (*substantives*) nomina of new species and genera. Following these guidelines requires to have at least a basic knowledge of the Latin grammar, or to seek help from someone having such knowledge.

Epithets must be words of two or more letters, that can either provide some information on the characters of the taxon (morphology, behaviour, colour, etc.), or on its geographic distribution, or dedicate it to a person or several, or even be devoid of “meaning”. They may be terms of various kinds, including adjectives (*Bufo viridis*) or participles in the nominative singular agreeing in gender with the generic substantive (*Rana clamitans*), invariant

nouns in the nominative singular in apposition to the substantive (*Pleurodeles waltl*), or nouns or adjectives in the genitive (*Rana boulengeri*). The *Code* recommends that new nomina “be euphonious and easily memorable”, but this is quite imprecise, especially as the term “euphonious” is not defined.

The same term may sound more or less “euphonious” according to the language spoken by a person. It does however seem clear that a brief nomen composed of simple syllables with only two or three letters each (one or two consonants and a single vowel) will be considered “simple and euphonious” by all, whereas more complex structures may not. The number and succession of letters, and the number of syllables, contribute to the “euphony” of nomina. In classical Latin, all vowels were pronounced separately, so that a nomen like *Hyalinobatrachium*, which contains 8 vowels, must be considered to consist of 8 different syllables (*Hy-a-li-no-ba-tra-chi-um*). Probably the longest generic nomen ever published in zoological nomenclature was *Siemienkiewicziechinogammarus* Dybowski, 1926 (29 letters, 14 syllables), which was fortunately invalidated by the International Commission on Zoological Nomenclature (see Ng, 1994), but other long nomina were published and must be used as the valid ones for the taxa they designate.

When coining new nomina, zootaxonomists should therefore ensure that they are short and simple. These considerations were clearly not taken into account by some authors who created long nomina in recent decades.

Tables 1–8 present some data on the length (number of letters) of 924 specific and 230 generic amphibian nomina published in various years from 1758 to 2008. As these samples are not of similar composition or size, they do not allow for carrying out statistical analyses, but they clearly suggest some trends.

Tables 1–4 show that, in our sample, the number of letters of specific epithets varies from 3 to 18, with a mean of 8.94 ± 2.61 , a mode of 8 and a median of 9 (medians not shown in tables). They tend to be longer in some regions of the world than in others, being longest in tropical Asia and neighbouring areas (Oriental region) and shortest in Europe and neighbouring areas (Palearctic region). The mean number of letters of all nomina published in the sampled years is higher than 10 for species described in two Asian countries, India and

particularly China. Among the longest nomina published in the sampled years (table 4), 15 out of 26 (i.e., 57.7 %) were from the Oriental region, although this region only accounts for 30.6 % of the total number of nomina (283 over 924). In contrast, although nomina for the Palearctic region account for only 6.1 % of the total number (56 over 924), they represent 14.3 % (3 over 21) of the shortest nomina (table 3).

The length of nomina is in part due to the way they are coined, i.e., their etymology. The length is shorter for nomina based on patronyms and on modern terms, higher for nomina based on classic (Greek and Latin roots) and on geographical terms, and dramatically higher for nomina based on geographical terms combined with the ending *-ensis* (in the masculine or feminine) or *-ense* (in the neuter). This is simply because the *-ensis* ending always increases the length of the word by 4 or 5 letters, making them 8–17 letters long, with a mean of 11.60 ± 1.75 letters, in our sample. Nomina in *-ensis* are clearly unnecessarily long, and they have another drawback: as they share this long ending, they often tend to resemble each other, so being liable to cause confusions among them. This is well shown by the following 15 epithets of Chinese species of the frog genus *Xenophrys* (Megophryidae) (see Fei *et al.*, 2008): *baolongensis*, *binchuanensis*, *binlingensis*, *huangshanensis*, *jingdongensis*, *kuatunensis*, *mangshanensis*, *medogensis*, *nankiangensis*, *sangzhiensis*, *shapingensis*, *shuichengensis*, *wawuensis*, *wuliangshanensis* and *wushanensis*. Scientific nomina should be an aid to communication, not a barrier to it. Taxonomists should seriously consider stopping coining epithets in *-ensis*, and using instead geographic names themselves, unchanged or slightly modified, as nouns in apposition to the generic substantive, or as adjectives agreeing in gender with it, or as names in the genitive (see below).

The 230 generic substantives surveyed in tables 5–8 cover the same range of variation in number of letters (3–18), but they tend to be longer than specific epithets: their mean is 10.62 ± 2.80 , their mode and median are 10. Nomina of genera described in African, Neotropical and Palearctic regions tend to be longer than those of the other three regions. The main cause of long substantives is the use of a classic (Latin or Greek) etymology, only followed secondly by nomina based on geographic terms. This is also shown by the fact that generic nomina proposed for fossils tend to be

longer than those for recent species, a trend not observed for specific nomina (data not shown), and these substantives are mostly of classic etymology. No one knows, when a scientific nomen is coined, whether it will appear only a few times in very specialized taxonomic publications or will become a very well-known nomen, used in hundreds or thousands of publications in all fields of research and even popular literature. Taxonomists should always keep this fact in mind when coining new nomina. For a species widely used in experimental research and whose nomen will appear in thousands of publications, it is much better to have a short and euphonious nomen like *Mus musculus* than *Caenorhabditis elegans*, which no one knows how to spell. Particular care should be taken not to coin long and unpalatable nomina in two situations: (1) when a taxon is “exceptional” or “extraordinary” in some respect, and thus likely to become famous and to be quoted hundreds of times in textbooks, in non-specialized literature, on the web and in various other media; in this respect, a nomen like *Karsenia* (8 letters, 4 syllables) is certainly preferable to *Nasikabatrachus* (15 letters, 6 syllables); (2) in the case of genera, whenever their nomen is, or is likely to become, the root of a family-series nomen, which will require the addition of three to five letters to it (*-INI*, *-IDAE*, *-OIDEA*, etc.); in this regard, a nomen like *Paa* (3 letters, 2 syllables; basis of *PAINI*, 5 letters, 3 syllables) is certainly preferable again to *Nasikabatrachus* (basis of *NASIKABATRACHIDAE*, 17 letters, 8 syllables).

We certainly do not need in zoological nomenclature specific nomina like *thoracotuberculatus* (19 letters, 8 syllables), *acanthidiocephalum* (18 letters, 8 syllables), *christianbergmanni* (18 letters, 6 syllables), *pseudomalabaricus* (17 letters, 8 syllables) or *tchabalmaboensis* (17 letters, 6 syllables), generic nomina like *Amphignathodontoides* (20 letters, 8 syllables), *Saevesoederberghia* (18 letters, 9 syllables), *Palaeosalamandrina* (18 letters, 9 syllables), *Pseudotyphlonectes* (18 letters, 7 syllables) or *Cryptobranchichnus* (18 letters, 5 syllables), familial nomina like *PSEUDOPHLEGETHONTIIDAE* (22 letters, 10 syllables) or *CALYPTOCEPHALELLIDAE* (20 letters, 9 syllables), or higher taxa nomina like *HYDATINOSALAMANDROIDEI* (22 letters, 11 syllables) or *PALAEOBATRACHOMORPHA* (20 letters, 9 syllables). Taxonomists should also certainly avoid coining particularly highly repetitive nomina like *Ogalalabatrachus* (16 letters, 7 syllables). Although such nomina are indeed a very small

minority among the many available nomina of zoology, there seems to be a trend for them to become more and more common, at least in some taxonomic groups. This can be exemplified by the curves published by Dubois & Raffaëlli (2009) concerning the mean increase of length over years of the generic nomina of the families *BUFONIDAE* and *SALAMANDRIDAE*, and the same is true in many other groups. For the sake of communication with the whole community of zoologists and non-zoologists, this increase should not be encouraged, and future nomina to be coined should be short and simple.

Dubois & Raffaëlli (2009) proposed the following rule of thumb regarding specific and generic nomina: they should include a maximum of 8–12 letters (preferably less) arranged in 4–5 syllables as defined above (preferably less), the latter being mostly composed of one or two consonant(s) and one vowel, as this is more likely to be euphonious in all or most languages. This should probably not become a rule of the *Code*, but it would be a useful addition to its recommendations. This rule of thumb can be used as a guideline by all taxonomists working nowadays. If we apply this rule of thumb to the 924 epithets surveyed here, this would mean that 94 of them (i.e., 10.2 %) are too long. If the bar is placed above 14 letters instead of 12, still 26 of them (i.e., 2.8 %), those listed here in table 4, are too long. Probably nobody can seriously argue that these 26 nomina were well-coined for fluent usage and easy memorization! Applying the same standards to the 230 generic substantives surveyed here, 56 (i.e., 24.3 %) would be too long with a bar above 12 letters and still 24 (i.e., 10.4 %) with a bar above 14. This shows that taxonomists, possibly because they want them to be “meaningful” and because they use classic etymologies for many of them, tend to coin generic nomina which are too long particularly if some of them are or will become the roots for family-series nomina. The number of letters of the expanded Latin alphabet (26), the use of which is allowed by the *Code*, offers enough possible combinations of letters to still allow the coining of numerous non-preoccupied euphonious nomina of 3 to 8 or even to 6 letters, as shown in tables 3 and 7.

There are several devices for coining short and simple nomina, without completely losing their etymology and “meaning”. Five recommendations can be given in this respect:

(1) The use of more than two complete roots for a nomen should be avoided, as this always results in

long unpalatable nomina (*Allomesotriton*, *Brachytarsophrys*, *Pseudhymenochirus*).

(2) For coining nomina based on two or more different roots, nothing in the *Code* requires one to combine the *complete* roots. Such nomina can validly be created by combining *parts* only of the roots, as exemplified by many generic nomina of amphibians (e.g., *Afrana*, *Grobina*, *Kurixalus*, *Megophrys*, *Telmalsodes*). Generic nomina like *Lyciasalamandra*, *Nasikabatrachus* or *Paramesotriton* are unnecessary long. Virtually the same nomina would aptly have been coined as “Lyciandra”, “Nasikus” or “Paratriton” (none of which is preoccupied).

(3) Among several roots that carry the same message, preference should be given to the shortest and simplest one: e.g., “rana” instead of “batrachus” or “bufo” instead of “phrynus”.

(4) An efficient way to reduce the length of nomina is to avoid adding long, useless endings to their basic root: thus, a specific nomen based on the name of a locality, region or country can well be coined by simply using the name of this place as it is, placed in apposition to the generic nomen, hence invariable. This avoids adding long endings in *-ensis*, *-ense*, *-cola*, *-icus*, *-ica*, *-icum*, *-ianus*, *-iana*, *-ianum*, etc. Additionally, this avoids high similarity between several epithets referred to the same genus and this precludes potential grammatical mistakes of agreement in gender in case of transfer of the species to another genus. Dubois & Raffaëlli (2009) suggested that this should become a recommendation of the *Code*, and that its current recommendation that “*An unmodified vernacular word should not be used as a scientific name*” should be deleted. This recommendation in fact amounts to encouraging the creation of long unpalatable nomina. The recent increase in the number of specific nomina ending in *-ensis*, especially in some countries, provokes real indigestion in people who are sensible to the aspect and length of nomina, and this should certainly change. Until now, specific nomina based on local geographical terms were rarely very short (compare tables 3 and 4), but they can be: short nomina like *Rana rara* (4 letters), *Rana diuata* (6 letters), *Aubria masako* (6 letters), *Hyperolius viridiflavus nimbae* (6 letters), *Avitabatrachus uliana* (6 letters), *Polypedates afghana* (7 letters), *Phrynopus carpish* (7 letters), *Rhacophorus laoshan* (7 letters), *Colostethus roraima* (7 letters) or *Bufo siculus* (7 letters) are based on geographic terms, and such

nomina should be preferred to unpalatable ones like those listed in table 4. All biologists and other users of zoological nomina would certainly appreciate such a move from zootaxonomists.

(5) The *Code* regulates the use of nomina many of which are just “Latin-like” but not genuine Latin terms. At the beginnings of zoological nomenclature, most zoologists had a good knowledge of the Latin language, but this has not been the case now for a long time. The “Latin” used in the *Code* is a very special, mostly technical, language that has very little to do with classical Latin. Latinists consider it as “dog Latin”, but it may be simply qualified as “*Code Latin*” (Dubois, 2007b). Only a small minority of all available nomina of animal taxa are real classical Latin terms. The *Code* only requires the use of the “Latin alphabet” (in fact expanded to include the letters *j*, *k*, *w* and *y*, absent in classical Latin), and a nomen may be derived from any language (even not using an alphabet), or even “*be an arbitrary combination of letters providing this is formed to be used as a word*”. The use of other languages offers an unlimited source of roots for scientific nomina that has been underexploited so far, although more and more taxonomists tend to use them. They allow in many cases the coining of very short, euphonious nomina that bring new blood into the old discipline of zoological nomenclature, such as *Rhacophorus kio* (3 letters), *Leptodactylus coca* (4 letters), *Batrachoseps kawia* (5 letters), *Leptolalax kecil* (5 letters), *Colostethus wayuu* (5 letters), *Plethodon cheoah* (6 letters), *Proceratophrys cururu* (6 letters), *Telmatobius huayra* (6 letters), *Platymantis isarog* (6 letters) or *Oreophryne wapoga* (6 letters). The recourse to such unusual etymologies should be strongly encouraged.

As a matter of fact, although important, the length of nomina is not the only element that should be taken into account. An important aspect is euphony or “pronounceability”, which should discourage the publication of nomina, even very short ones, that cannot be articulated in most languages, such as *Philautus crnri* (see Dubois, 1999). Another strong quality for a nomen is its originality. It often allows memorization of the nomen, and avoids potential confusion with other nomina, or potential homonymy in cases of generic reallocation. Many taxonomists seem to lack imagination and to be prisoners of academic traditions, as testified by the plethora of species bearing nomina like *viridis*, *maculatus*, *vulgaris*, *monticola* or *sinensis*. Using either classic or modern etymologies, it is possible

to coin original, poetic, suggestive nomina like *campanisona*, *cavernibardus*, *cuneirostris*, *orchymelas* or *pipilodryas*, which, although longer than those listed above, are to be encouraged as they are likely to remain associated with a unique, well identified species. Taxonomists should feel free to use their imagination and good taste to coin unprecedented nomina rather than always using the same old terms.

Conclusion

Humanity could probably do without taxonomy. It could also do without many other cultural, scientific and technical knowledge and skills. After all, human beings survived as wandering groups seeking temporary shelter in caves for much more considerable periods than as civilizations with agriculture, cities and books. Humanity could also probably survive on a much depauperate planet with a much reduced biodiversity where only a few species directly “useful to man” would have avoided extinction (but also probably several inescapable “harmful to man” ones!). Is this what we want? Do we just want to “survive”? Or do we want to live on a beautiful planet, with a diverse biodiversity, a rich culture, deep scientific knowledge about the world we live in, its evolution and history? If so, we absolutely need taxonomy, and we need to put much more emphasis on this discipline in the present period than before, because of the urgency caused by the biodiversity crisis: many species that will not be collected and studied in the coming decades will be forever unavailable for study later, as they will vanish without even leaving fossils.

But because the need for taxonomy is still not understood by many decision-makers, and even by many scientists including biologists, including many “evolutionary biologists” and “conservation biologists”, taxonomy needs to be an *exemplary* discipline. It cannot continue to work by approximation. Taxonomists cannot continue to create new synonyms, mostly in the hope to “attach their names” to new taxa (Dubois, 2008d), or to create unpalatable nomina, or nomina liable to be confused with others, because they are unable to think that scientific nomina are means of communication, not only within their small group of specialists, but with all other biologists and with society as a whole. They must use much better and stricter standards for their discipline, not only as regards their taxonomic methodology (Dayrat, 2005), but also concerning their nomenclatural approach, which should finally get rid of the

“principle of authority”, to follow stringent and universal rules (Dubois, 2010b). Only at the expense of such efforts will this discipline be able to be recognized as an important one among sciences.

Acknowledgements

I am grateful to Annemarie Ohler (Paris) for help in the quantitative analyses presented in the tables.

Further reading

A complete bibliography on the questions tackled above would occupy more than a full issue of *Taprobanica*. For the readers interested, additional comments and many references can be found in the following works, some of which only are cited in the text above:

Taxonomic methodology, species concept, strange species

Dayrat, B., 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85: 407–415.

Delorme, M., A. Dubois, S. Grosjean and A. Ohler, 2006. Une nouvelle ergotaxinomie des Megophryidae (Amphibia, Anura). *Alytes*, 24 (1–4): 6–21.

Dubois, A., 1984. Sample-size constraints in the use of nonparametric Mann-Whitney *U* test for the comparison of two independent samples: consequences in Anuran Amphibians systematics. *Alytes*, 3 (1): 20–24.

Dubois, A., 1988. The genus in zoology: a contribution to the theory of evolutionary systematics. *Mémoires du Muséum national d'Histoire naturelle*, (A), 140: 1–123.

Dubois, A., 1991. Nomenclature of parthenogenetic, gynogenetic and “hybridogenetic” vertebrate taxons: new proposals. *Alytes*, 8, 61–74.

Dubois, A., 1998. List of European species of amphibians and reptiles: will we soon be reaching “stability”? *Amphibia-Reptilia*, 19 (1): 1–28.

Dubois, A., 1999. South Asian Amphibia: a new frontier for taxonomists. Invited editorial / Book review. *Journal of South Asian natural History*, 4 (1): 1–11.

Dubois, A., 2003. The relationships between taxonomy and conservation biology in the century of extinctions. *Comptes rendus Biologies*, 326 (suppl. 1): S9–S21.

- Dubois, A., 2004a. Developmental pathway, speciation and supraspecific taxonomy in amphibians. 1. Why are there so many frog species in Sri Lanka? *Alytes*, 22 (1-2): 19–37.
- Dubois, A., 2004b. Developmental pathway, speciation and supraspecific taxonomy in amphibians. 2. Developmental pathway, hybridizability and generic taxonomy. *Alytes*, 22 (1-2): 38–52.
- Dubois, A., 2008a. Handicap taxinomique et crise de la biodiversité: un nouveau paradigme pour la biologie au XXI^e siècle. In: D. Prat, A. Raynal-Roques and A. Roguenant (Eds.), *Peut-on classer le vivant? Linné et la systématique aujourd'hui*, Paris, Belin: 141–160.
- Dubois, A., 2008b. Drôles d'espèces. Hybridation, perturbations de la méiose et spéciation dans le règne animal: quelques points délicats de terminologie, d'éidonomie et de nomenclature. In: D. Prat, A. Raynal-Roques and A. Roguenant (Eds.), *Peut-on classer le vivant? Linné et la systématique aujourd'hui*, Paris, Belin: 169–202.
- Dubois, A., 2008c. Un nouveau paradigme pour la biologie au XXI^e siècle. In: M. Veuille, J.-M. Drouin, R. Deleporte and J.-F. Silvain (Eds.), *Linnaeus – Systématique et biodiversité*, *Biosystema*, 25: 127–142.
- Dubois, A., 2009a. Qu'est-ce qu'une espèce animale? In: *Aller à l'espèce: illusion ou nécessité*, *Mémoires de la Société entomologique de France*, 8: 9–48.
- Dubois, A., 2010a. Editorial. Taxonomy in the century of extinctions: taxonomic gap, taxonomic impediment, taxonomic urgency. *Taprobanica*, 2(1): 1-5.
- Dubois, A. and R. Günther, 1982. Klepton and synklepton: two new evolutionary systematics categories in zoology. *Zoologische Jahrbucher, Abteilung für Systematik, Ökologie & Biologie der Tiere*, 109: 290–305.
- Hennig, W., 1950. *Grundzüge einer Theorie der phylogenetischen Systematik*. Berlin, Deutscher Zentralverlag: i–vii+1–370.
- Hennig, W., 1966. *Phylogenetic systematics*. Urbana, Chicago & London, University of Illinois Press: i–vii+1–263.
- Mayr, E., 1969. *Principles of systematic zoology*. New York, McGraw-Hill: i–xiii+1–428.
- Mayr, E., 1981. Biological classification: toward a synthesis of opposing methodologies. *Science*, 214: 510–516.
- Mayr, E., 1982. *The growth of biological thought*. Cambridge, Massachusetts & London, Belknap Press: i–xiii+1–974.
- Mayr, E., 1995. Systems of ordering data. *Biology & Philosophy*, 10: 419–434.
- Mayr, E., 1997. *This is biology. The science of the living world*. Cambridge, Massachusetts & London, Belknap Press: i–xvii+1–327.
- Mayr, E. and P. D. Ashlock, 1991. *Principles of systematic zoology*. Second edition. New York, McGraw-Hill: i–xx+1–475.
- Simpson, G. G., 1940. Types in modern taxonomy, *American Journal of Science*, 238: 413–431.
- Simpson, G. G., 1961. *Principles of animal taxonomy*. New York, Columbia University Press: i–xii+1–247.
- Nomenclatural methodology**
- Anonymous [International Commission on Zoological Nomenclature], 1999. *International code of zoological nomenclature*. Fourth edition. London, International Trust for zoological Nomenclature: i–xxix+1–306.
- Dubois, A., 2000. Synonymies and related lists in zoology: general proposals, with examples in herpetology. *Dumerilia*, 4 (2): 33–98.
- Dubois, A., 2005. Proposed Rules for the incorporation of nomina of higher-ranked zoological taxa in the *International Code of Zoological Nomenclature*. 1. Some general questions, concepts and terms of biological nomenclature. *Zoosystema*, 27 (2): 365–426.
- Dubois, A., 2006a. Proposed Rules for the incorporation of nomina of higher-ranked zoological taxa in the *International Code of Zoological Nomenclature*. 2. The proposed Rules and their rationale. *Zoosystema*, 28 (1): 165–258.
- Dubois, A., 2006b. New proposals for naming lower-ranked taxa within the frame of the *International Code of Zoological Nomenclature*. *Comptes rendus Biologies*, 329 (10): 823–840.
- Dubois, A., 2006c. Incorporation of nomina of higher-ranked taxa into the *International Code of Zoological Nomenclature*: some basic questions. *Zootaxa*, 1337: 1–37.

- Dubois, A., 2006d. Naming taxa from cladograms: a cautionary tale. *Molecular Phylogenetics & Evolution*, 42: 317–330.
- Dubois, A., 2007a. Phylogeny, taxonomy and nomenclature: the problem of taxonomic categories and of nomenclatural ranks. *Zootaxa*, 1519: 27–68.
- Dubois, A., 2007b. Genitives of species and subspecies nomina derived from personal names should not be emended. *Zootaxa*, 1550: 49–68.
- Dubois, A., 2007c. Naming taxa from cladograms: some confusions, misleading statements, and necessary clarifications. *Cladistics*, 23: 390–402.
- Dubois, A., 2008d. A partial but radical solution to the problem of nomenclatural taxonomic inflation and synonymy load. *Biological Journal of the Linnean Society*, 93: 857–863.
- Dubois, A., 2008e. Authors of zoological publications and nomina are signatures, not persons. *Zootaxa*, 1771: 63–68.
- Dubois, A., 2008f. Le *Code international de nomenclature zoologique*: présentation, philosophie, règles majeures, problèmes actuels. In: D. Prat, A. Raynal-Roques and A. Roguenant (Eds.), *Peut-on classer le vivant? Linné et la systématique aujourd'hui*, Paris, Belin: 355–402.
- Dubois, A., 2008g. Phylogenetic hypotheses, taxa and nomina in zoology. In: A. Minelli, L. Bonato and G. Fusco (Eds.), *Updating the Linnaean heritage: names as tools for thinking about animals and plants*, *Zootaxa*, 1950: 51–86.
- Dubois, A., 2009b. Miscellanea nomenclatorica batrachologica. 20. Class-series nomina are nouns in the nominative plural: Terrarana Hedges, Duellman and Heinicke, 2008 must be emended. *Alytes*, 26 (1-4): 167–175.
- Dubois, A., 2009c. Incorporation of nomina of higher-ranked taxa into the *International Code of Zoological Nomenclature*: the nomenclatural status of class-series zoological nomina published in a non-latinized form. *Zootaxa*, 2106: 1–12.
- Dubois, A., 2009d. Endangered species and endangered knowledge. *Zootaxa*, 2201: 26–29.
- Dubois, A., 2010b. Zoological nomenclature in the century of extinctions: priority vs. “usage”. *Organisms Diversity & Evolution*, 2426: 1–42.
- Dubois, A., 2010c. Registration as a fourth level of the nomenclatural rules. *Bulletin of zoological Nomenclature*, 67 (1): 11–23.
- Dubois, A. and A. Nemésio, 2007. Does nomenclatural availability of nomina of new species or subspecies require the deposition of vouchers in collections? *Zootaxa*, 1409: 1–22.
- Dubois, A. and J. Raffaëlli, 2009. A new ergotaxonomy of the family Salamandridae Goldfuss, 1820 (Amphibia, Urodela). *Alytes*, 26 (1-4): 1–85.
- Ng, P. K. L., 1994. The citation of species names and the role of the author’s name. *Raffles Bulletin of Zoology*, 42 (3): 509–513.

Other references cited above

- Dubois, A., R. I. Crombie and F. Glaw, 2005. *Amphibia Mundi*. 1.2. Recent amphibians: generic and infrageneric taxonomic additions (1981–2002). *Alytes*, 23 (1-2): 25–69.
- Fei, L., S. Hu, C. Ye and Y. Huang (Eds.) 2008. *Amphibia*. Volume 2. *Fauna Sinica*, Beijing, Science Press: [i–vii]+i–xiii+1–957. [In Chinese].
- Frost, D. R. (Ed.) 2010. *Amphibian Species of the World*. [http://research.amnh.org/herpetology/amphibia]. [Accessed on 10 March 2010].
- Guayasamin, J. M., S. Castroviejo-Fisher, L. Trueb, J. Ayarzagüena, M. Randa and C. Vilà, 2009. Phylogenetic systematics of glassfrogs (Amphibia: Centrolenidae) and their sister taxon *Allophryne ruthveni*. *Zootaxa*; 2100: 1–97.

DESCRIBING NEW SPECIES

Table 1. – Some data on the number of letters in 924 amphibian epithets published from 1758 to 2008. These include all the epithets published in the years 1758, 1808, 1858, 1858, 1808, 1958 and 2008 according to Frost (2010), and all the nomina published from 1981 to 2002 listed by Dubois *et al.* (2005). The species described during these years came from 71 countries, here distributed in 6 main regions. Detailed data are given for the 14 countries represented by 15 or more new species. Five categories of etymology are recognized: “Patronym”, for any scientific nomen based on the patronym of one or several person(s); “Modern”, for any scientific nomen derived from a term in modern language or of unknown origin; “Classic”, for any scientific nomen based on Latin or Greek roots or on a combination of both; “Geographic varia”, for any scientific nomen based on the name of a locality, region or country, but not ending in “-ensis” or “-ense”; “Geographic with -ensis”, for any scientific nomen, based on the name of a locality, region or country, followed by the ending “-ensis” or “-ense”.

Criterion	Category	Number of nomina	Number of letters: range	Number of letters: mean ± standard deviation
All nomina	–	924	3 – 18	8.94 ± 2.61
Regions	Palaearctic	56	4 – 13	8.39 ± 2.12
	Oceania	9	5 – 12	8.56 ± 2.74
	Nearctic	43	5 – 14	8.63 ± 2.47
	Africa	103	4 – 17	8.64 ± 2.34
	Neotropical	430	3 – 16	8.75 ± 2.55
	Oriental	283	3 – 18	9.53 ± 2.81
Countries	Madagascar (Africa)	27	4 – 12	7.89 ± 1.78
	Colombia (Neotropical)	57	5 – 15	8.02 ± 2.50
	Indonesia (Oriental)	24	5 – 14	8.08 ± 2.41
	Brazil (Neotropical)	47	3 – 16	8.60 ± 3.13
	Mexico (Neotropical)	27	6 – 14	8.70 ± 1.94
	Bolivia (Neotropical)	16	5 – 13	8.81 ± 2.43
	USA (Nearctic)	31	5 – 14	8.94 ± 2.76
	Papua New Guinea (Oriental)	31	5 – 15	8.97 ± 2.04
	Philippines (Oriental)	15	6 – 15	9.07 ± 2.96
	Vietnam (Oriental)	15	5 – 15	9.20 ± 3.10
	Peru (Neotropical)	46	6 – 13	9.96 ± 2.05
	Venezuela (Neotropical)	33	5 – 16	9.97 ± 2.79
	India (Oriental)	20	5 – 17	10.10 ± 3.19
	China (Oriental)	70	3 – 16	10.70 ± 2.68
Etymology	Patronym	267	3 – 18	7.79 ± 2.21
	Modern	114	4 – 17	7.86 ± 2.32
	Classic	379	4 – 16	9.18 ± 2.55
	Geographic varia	40	6 – 15	9.32 ± 1.98
	Geographic with -ensis	124	8 – 17	11.60 ± 1.75

Table 2. – Numbers of letters in the 924 epithets surveyed in table 1, according to the five categories of etymology defined there. In each column, the number in **bold** is the mode of the distribution.

Number of letters	Patronym	Modern	Classic	Geographic varia	Geographic with -ensis	Total
3	3	0	0	0	0	3
4	9	3	6	0	0	18
5	20	12	16	0	0	48
6	39	19	29	3	0	90
7	59	20	52	4	0	135
8	55	26	61	5	2	149
9	35	12	58	13	9	127
10	21	8	50	8	25	112
11	10	6	34	2	30	82
12	8	3	31	4	20	66
13	2	1	17	0	23	43
14	3	3	11	1	7	25
15	2	0	11	0	5	18

16	0	0	3	0	2	5
17	0	1	0	0	1	2
18	1	0	0	0	0	1
Total	267	114	379	40	124	924

Table 3. – The 21 shortest epithets (3–4 letters) among the 924 surveyed in tables 1 and 2. LE: number of letters of the epithet. GS (grammatical status of epithet): AN, adjective in the nominative, in gender agreement with substantive; NG, invariable noun in the genitive; NN, invariable noun in the nominative, in apposition. Etymology (see legend of fig. 1): C, Classic; M, Modern; P, Patronym.

Binomen	LE	Year	GS	Country	Region	Etymology
<i>Adenomus dasi</i>	4	1998	NG	Sri Lanka	Oriental	P
<i>Alexteroon jynx</i>	4	2000	NN	Cameroon	Africa	C
<i>Atelognathus ceii</i>	4	1998	NG	Cuba	Neotropical	P
<i>Bolitoglossa tica</i>	4	2008	AN	Costa Rica	Neotropical	M
<i>Bufo jimi</i>	4	1999	NG	Brazil	Neotropical	P
<i>Chiasmocleis jimi</i>	4	2001	NG	Brazil	Neotropical	P
<i>Eleutherodactylus zeus</i>	4	1958	NN	Cuba	Neotropical	C
<i>Eleutherodactylus zugii</i>	4	1958	NG	Cuba	Neotropical	P
<i>Hyla rhea</i>	4	1999	NN	Brazil	Oriental	C
<i>Hylodes uai</i>	3	2001	NG	Brazil	Neotropical	P
<i>Hyperolius fuca</i>	4	2008	AN	Taiwan	Oriental	C
<i>Leptodactylus coca</i>	4	2008	NN	Bolivia	Neotropical	M
<i>Leptolalax liui</i>	4	1991	NG	China	Oriental	P
<i>Mantidactylus enki</i>	4	2002	NG	Madagascar	Africa	P
<i>Paa (Feirana) yei</i>	3	2002	NG	China	Oriental	P
<i>Rana bufo</i>	4	1758	NN	Sweden	Palaearctic	C
<i>Rana hyla</i>	4	1758	NN	Switzerland	Palaearctic	C
<i>Rana lini</i>	4	1999	NG	China	Oriental	P
<i>Rana osca</i>	4	1993	AN	Italy	Palaearctic	C
<i>Rana pipa</i>	4	1758	NN	Surinam	Neotropical	M
<i>Stefania coxi</i>	4	2002	NG	Guyana	Neotropical	P

Table 4. – The 26 longest epithets (15–18 letters) among the 924 surveyed in tables 1 and 2. LE: number of letters of the epithet. GS (grammatical status of epithet): AN, adjective in the nominative, in gender agreement with substantive; NG, invariable noun in the genitive; NN, invariable noun in the nominative, in apposition. Etymology (see legend of fig. 1): C, Classic; GE, Geographic with *-ensis*; GV, Geographic varia; P, Patronym.

Binomen	LE	Year	GS	Country	Region	Etymology
<i>Amolops spinapectoralis</i>	15	1999	AN	Vietnam	Oriental	C
<i>Amolops tuberodepressus</i>	15	2000	AN	China	Oriental	C
<i>Atelopus monohernandezii</i>	15	2002	NG	Colombia	Neotropical	P
<i>Austrochaperina novaebritanniae</i>	15	2000	NG	Papua New Guinea	Oriental	GV
<i>Bolitoglossa guaramacalensis</i>	15	2002	AN	Venezuela	Neotropical	GE
<i>Centrolene papillahallicum</i>	15	2000	AN	Guyana	Neotropical	C
<i>Colostethus caeruleodactylus</i>	16	2001	AN	Brazil	Neotropical	C
<i>Crotaphatrema tchabalmaboensis</i>	17	2000	AN	Cameroon	Africa	GE
<i>Eleutherodactylus angustilineatus</i>	15	1998	AN	Colombia	Neotropical	C
<i>Eleutherodactylus rhyacobatrachus</i>	15	2000	NN	Costa Rica	Neotropical	C
<i>Hyalinobatrachium guairarepanensis</i>	16	2001	AN	Venezuela	Neotropical	GE
<i>Hyla pseudomeridiana</i>	15	2000	AN	Brazil	Neotropical	C
<i>Leptolalax ventripunctatus</i>	15	1991	AN	China	Oriental	C
<i>Litoria christianbergmanni</i>	18	2008	NG	Indonesia	Oriental	P
<i>Megophrys wuliangshanensis</i>	16	1995	AN	China	Oriental	GE
<i>Micrixakus magnapustulosus</i>	15	1958	AN	Thailand	Oriental	C

DESCRIBING NEW SPECIES

<i>Nyctibatrachus dattatreyaensis</i>	15	2008	AN	India	Oriental	GE
<i>Odorrana exiliversabilis</i>	15	2001	AN	China	Oriental	C
<i>Philautus erythrophthalmus</i>	16	2000	AN	Malaysia	Oriental	C
<i>Phrynopus lechriorhynchus</i>	15	2008	AN	Peru	Neotropical	C
<i>Platymantis sierramadrensis</i>	15	1999	AN	Philippines	Oriental	GE
<i>Proceratophrys concavitympanum</i>	15	2000	NN	Brazil	Neotropical	C
<i>Rana multidenticulata</i>	16	1997	AN	Taiwan	Oriental	C
<i>Rhacophorus chuyangsinensis</i>	15	2008	AN	Vietnam	Oriental	GE
<i>Rhacophorus pseudomalabaricus</i>	17	2000	AN	India	Oriental	C
<i>Scutigera mokokchungensis</i>	15	2000	AN	India	Oriental	GE

Table 5. – Some data on the number of letters in 230 generic substantives published from 1758 to 2008. These include all the nomina published in the years 1758, 1808, 1858, 1858, 1808, 1958 and 2008 according to Frost (2010), all the nomina published from 1981 to 2002 listed by Dubois *et al.* (2005) and all those of the families *BUFONIDAE*, *RANIDAE* and *SALAMANDRIDAE* surveyed in Dubois & Raffaëlli (2009). Four categories of etymology are recognized: “Modern”, for any scientific nomen derived from a term in modern language or of unknown origin; “Patronym”, for any scientific nomen based on the patronym of one or several person(s); “Classic”, for any scientific nomen based on Latin or Greek roots or on a combination of both; “Geographic”, for any scientific nomen based on the name of a locality, region or country.

Criterion	Category	Number of nomina	Number of letters: range	Number of letters: mean ± standard deviation
All nomina	–	230	3 – 18	10.62 ± 2.80
Regions	Oceania	1	9	9
	Oriental	70	3 – 14	9.81 ± 2.42
	Nearctic	19	7 – 15	10.00 ± 2.60
	Palaearctic	67	4 – 18	10.66 ± 2.95
	Neotropical	47	6 – 18	11.15 ± 2.96
	Africa	26	7 – 15	12.23 ± 2.49
Fossil / recent	Recent	175	3 – 18	10.38 ± 2.79
	Fossil	55	7 – 18	11.38 ± 2.70
Etymology	Modern	14	6 – 14	9.43 ± 2.62
	Patronym	35	3 – 15	9.83 ± 2.97
	Geographic	21	7 – 15	10.52 ± 2.62
	Classic	160	4 – 18	10.91 ± 2.76

Table 6. – Numbers of letters in the 230 generic substantives surveyed in table 5, according to the four categories of etymology defined there. In each column, the number in **bold** is the mode of the distribution.

Number of letters	Patronym	Modern	Geographic	Classic	Total
3	1		0	0	1
4	1		0	2	3
5	0		0	0	0
6	1	2	0	3	6
7	2	2	1	6	11
8	8	1	5	22	36
9	5	3	4	17	29
10	5	2	2	29	38
11	2	1	2	20	25
12	2	1	2	20	25
13	3	0	1	14	18
14	2	2	1	9	14
15	3	0	3	7	13
16	0	0	0	5	5
17	0	0	0	3	3
18	0	0	0	3	3
Total	35	14	21	160	230

Table 7. – The 21 shortest substantives (3–7 letters) among the 230 surveyed in table 5. LS: number of letters of the epithet. F / R: fossil / recent. Etymology (see legend of fig. 5): C, classic; G, geographic; M, modern; P, patronym.

Binomen	LS	Year	F / R	Country	Region	Etymology
<i>Amerana</i>	7	1992	R	USA	Nearctic	G
<i>Amo</i>	3	1992	R	Malaysia	Oriental	M
<i>Amolops</i>	7	1865	R	Unknown	Oriental	C
<i>Ansonia</i>	7	1870	R	Malaysia	Oriental	C
<i>Babina</i>	6	1912	R	Japan	Oriental	M
<i>Bishara</i>	7	1997	F	Kazakhstan	Palaearctic	M
<i>Bufo</i>	4	1758	R	Sweden	Palaearctic	C
<i>Chaunus</i>	7	1828	R	Brazil	Neotropical	C
<i>Conraua</i>	7	1908	R	Cameroon	Africa	M
<i>Cynops</i>	6	1838	R	Japan	Oriental	C
<i>Eburana</i>	7	1992	R	Japan	Oriental	C
<i>Huia</i>	4	1991	R	Malaysia	Oriental	P
<i>Meinus</i>	6	2009	R	Spain	Palaearctic	M
<i>Pingia</i>	6	1935	R	China	Oriental	P
<i>Rafinus</i>	7	2009	R	USA	Nearctic	P
<i>Rana</i>	4	1758	R	Sweden	Palaearctic	C
<i>Rhaebo</i>	6	1862	R	Colombia	Neotropical	M
<i>Rugosa</i>	6	1991	R	Japan	Oriental	C
<i>Taricha</i>	7	1850	R	USA	Nearctic	C
<i>Tigrina</i>	7	1990	R	India	Oriental	C
<i>Twittyta</i>	7	2009	R	USA	Nearctic	P

Table 8. – The 24 longest substantives (15–18 letters) among the 230 surveyed in table 5. LS: number of letters of the substantive. F / R: fossil / recent. Etymology (see legend of fig. 5): C, classic; P, patronym.

Binomen	LS	Year	F / R	Country	Region	Etymology
<i>Atelophryniscus</i>	15	1989	R	Honduras	Neotropical	C
<i>Cryptobranchichnus</i>	18	1941	F	Italy	Palaearctic	C
<i>Cryptophyllobates</i>	17	2000	R	Peru	Neotropical	C
<i>Dendrophryniscus</i>	16	1871	R	Brazil	Neotropical	C
<i>Gigantobatrachus</i>	16	1958	R	Argentina	Neotropical	C
<i>Heteroclitotriton</i>	17	1903	F	France	Palaearctic	C
<i>Lyciasalamandra</i>	15	2004	R	Turkey	Palaearctic	C
<i>Melanophryniscus</i>	16	1961	R	Argentina	Neotropical	C
<i>Nectophrynoides</i>	15	1926	R	Tanzania	Africa	C
<i>Nimbaphrynoides</i>	15	1987	R	Guinea	Africa	C
<i>Palaeopleurodeles</i>	17	1941	F	Germany	Palaearctic	C
<i>Palaeosalamandra</i>	16	1949	F	Germany	Palaearctic	C
<i>Palaeosalamandrina</i>	18	1949	F	Germany	Palaearctic	C
<i>Paratelmatobius</i>	15	1958	R	Brazil	Neotropical	C
<i>Peratosauroides</i>	15	1983	F	Uzbekistan	Palaearctic	C
<i>Poyntonophrynus</i>	15	2006	R	South Africa	Africa	P
<i>Pseudopelobates</i>	15	1958	R	South Caucasus	Palaearctic	C
<i>Pseudotyphlonectes</i>	18	1986	R	Colombia	Neotropical	C
<i>Psychrophrynella</i>	16	2008	R	Peru	Neotropical	C
<i>Rubicacaecilia</i>	15	2001	F	Morocco	Palaearctic	C
<i>Septentriomolge</i>	15	2001	R	USA	Nearctic	C
<i>Spinophrynoides</i>	15	1987	R	Ethiopia	Africa	C
<i>Vandijkophrynus</i>	15	2006	R	South Africa	Africa	P
<i>Wolterstorffina</i>	15	1939	R	Cameroon	Africa	P