

CHAPTER 2

A Critique of Current Taxonomy

2.1. THE ILLS OF MODERN TAXONOMY

Recent years have witnessed increasing dissatisfaction with the principles and practices of biological classification. New species are constantly being described and many new characters discovered in known species by modern studies in such fields as cytology, biochemistry, and behavior, as well as by more penetrating employment of traditional morphological methods. As our knowledge of the organic world increases there are continuing stresses and strains in the frame of the taxonomic system to accommodate these new discoveries, and the inadequacies of the present system become ever more apparent.

The thirties and forties of this century witnessed a considerable revision of our interpretation of evolutionary phenomena. Advances in genetics, cytology, and geographic variation had prepared the way for a "New Systematics," the advent of which might conveniently be dated with the appearance of Julian Huxley's book of the same name (1940). Considered *avant-garde* by taxonomists at first, this material slowly became assimilated, so that today terms such as cline, Rassenkreis, or gene pool are commonly used by most of them. However, a survey of the literature will soon make it apparent that most of the progress has been made at the species and infraspecies levels, and that there has in fact been little increase in our understanding of the nature and evolution of the higher categories. Books such as those by Rensch (1947), Schmalhausen (1949), and Simpson (1953) deal with this topic, but they contain little more than descriptive generalizations of it. The systematics

of the higher categories, and indeed problems of classification in general, have not benefited particularly from the "New Systematics," which is in fact rather a poorly chosen term, since the subject concerns primarily infraspecific variation and mechanisms of speciation. A recent paper by Blackwelder (1962) discusses the failure of the "New Systematics" to provide an adequate basis for animal taxonomy.

During the last few years a number of publications have appeared which attempt a re-evaluation of the logical bases of taxonomy. This has involved a separation of the various functions which the science is trying to fulfill. Cain, in a series of papers (1956, 1958, 1959a,b,c; and see Cain and Harrison, 1960b), has performed this task admirably, and there have been similar independent contributions by Gilmour (1937, 1940, 1951, 1961b), Michener (1957), Michener and Sokal (1957), Sneath (1957a), Sokal (1960), and Sneath and Sokal (1962).

There is remarkable agreement among these various authors to the effect that the present system of taxonomy attempts to fulfill too many functions and as a consequence does none of them well. It attempts (1) to classify, (2) to name, (3) to indicate degree of resemblance (affinity), and (4) to show relationship by descent—all at the same time. We shall show in separate sections below that it is impossible not only in practice but also in theory for the current system to perform these tasks adequately. Section 2.2 will concern itself with the problems of the natural system; Section 2.3 will deal with the difficulties of assigning phylogenetic interpretations to expressions of phenetic resemblance.

It may be advantageous at this stage to outline an important logical fallacy underlying current taxonomic procedure. It is the self-reinforcing circular arguments used to establish categories, which on repeated application invest the latter with the appearance of possessing objective and definable reality. This type of reasoning is, of course, not restricted to taxonomy—but it is no less fallacious on that account. Let us illustrate this point. An investigator is faced with a group of similar species. He wishes to show relationships among the members of the group and is looking for characters which will subdivide it into several mutually exclusive taxa. A search for characters reveals that within a subgroup **A** certain characters appear constant, while varying in an uncorrelated manner in other subgroups. Hence a taxon **A** is described and defined on the basis of this character complex, say *X*. It is assumed that taxon **A** is a monophyletic or a "natural" taxon. Thus every member of **A** (both known and unknown forms) is expected to possess *X*; conversely, possession of the character complex *X* defines **A**.

Henceforth group **A**, as defined by *X*, assumes a degree of permanence and reality quite out of keeping with the tentative basis on which it was established. Subsequently studied species are compared with **A** to establish their affinities; they may be within **A**, close to it, or far from it. It is quite possible that a species not showing *X* would be excluded from **A**, although it was closer overall to most of the members of **A** than some were to each other. It may be said that such problems would arise only when **A** was an "artificial" group erected on the basis of "unsuitable" characters. However, except in long-established taxa or those separated by very wide gaps from their closest relatives, the effect of the last classification carried out with a limited number of characters is quite pervasive. The circular reasoning arises from the fact that new characters, instead of being evaluated on their own merits, are inevitably prejudiced by the prior erection of taxon **A** based on other characters (*X*). Such a prejudgment ignores the fact that the existence of **A** as a natural (or "monophyletic") group defined by character complex *X* has been *assumed* but *not demonstrated*.

The taxonomist will recognize and define taxa on three kinds of evidence, which are not often clearly separated either in his mind or in publication. (1) Resemblance—those entities which phenetically resemble each other more than they do others form a taxon. (2) Homologous characters—a taxon is formed by entities sharing characters of common origin. (3) Common line of descent—membership in a common line of descent will define a taxon. Since (3) is rarely if ever known, it is usually inferred from (1) or (2). Conclusions on homologies (evidence of type 2) are often deduced from phylogenetic speculations (evidence of type 3). Thus taxonomists often reason facilely back and forth among these criteria without stopping to think how slender the evidence is on which their arguments are based. Their sound knowledge is usually restricted to an estimate of resemblance of a number of organisms without any knowledge of phylogeny and hence the common origin of their characters. Their estimates are usually based on few characters and thus do not reflect the overall similarity which could be obtained when more of the phenotype is considered. Once it is looked at critically, it becomes evident how much of taxonomic procedure is circular reasoning and extrapolation.

Phylogenetic interpretations of systematic relationships have remained in fashion ever since Haeckel. Modern emphasis on the "New Systematics" has attempted to embellish such conclusions with speculations on the evolutionary mechanisms likely to have brought about the sup-

posed systematic relationships under study. The authors are familiar with practices in several university departments where "plain" taxonomic theses are not welcomed or recognized for the Ph.D. degree and where a discussion of phylogeny and evolutionary principles illustrated by the revised taxon is mandatory. In the absence of experimental and fossil evidence such accounts are usually purely speculative and involve much tortuous dialectic. It seems to us rather absurd to indulge in speculations of this sort when for most groups the basic facts of phenetic resemblance are still to be obtained. There is, of course, no harm in speculation per se. The points we wish to make are that phylogenetic speculations should not be involved in the classificatory process and that any such speculations should be based upon a sound phenetic classification.

It is current practice to employ only homologous characters in creating taxonomic groups (the word is used here in the sense of characters having a common ancestral origin). The aim of this practice is to create only phyletic groups, not phenetic ones, since phenetic resemblance is based on all characters, including convergent ones. We are therefore faced by the question of which kind of classification we wish to make. Whatever our wishes, a number of points are clear. First, at the infra-specific level it is often impossible to decide if a feature is homologous. For example, the independent occurrences of repeated albino mutations in a species, as in mice, are not homologous in the sense used above (though the *tendency for such mutations to occur* may be homologous if it is an inherited tendency of the species). Second, in major phyla there is seldom any doubt about whether most character complexes are homologous, taken as a whole. In the intermediate ranks these decisions are especially difficult to make. Third, it is doubtful if taxonomists are in practice prepared to make decisions on the homologies of the thousands of characters which the organisms possess. We may conclude that as a general principle restricting classification to using homologous characters only is not feasible even if it were always desirable. This is doubly true when we consider how confused the concept of homology has become (see Lam, 1959).

Giving different weight to different characters according to their presumed importance is still orthodox teaching today. We believe that such weighting is unsound, and it will be discussed at length in Section 5.7.

An undesirable procedure of taxonomy, amounting to improper weighting, is the way in which certain characters will be used to differentiate the species of one genus, while being ignored in the next genus.

This is not necessarily because the character has been found unsatisfactory. The resulting harm to taxonomy is far-reaching. Not only is the diagnostic value of the character diminished, but it carries the danger that such arbitrariness may yield arbitrary and artificial taxa. The practical need for simple methods of identification, generally by means of diagnostic keys, has also had an undesirable influence on methods of classification; prominent single characters are valuable for such keys, and hence they have commonly been unconsciously assumed to be very important for creating the taxa.

In discussing the ill of modern systematics, mention must be made of several other contributing factors. It is discouraging to see how fragmented the study of systematics has become. In many biological disciplines there is an unconscious arrogance which assumes that the methods and definitions of other disciplines are useless or misleading. Few zoologists discuss whether their systematic principles can be applied to botany, since for them taxonomy is zoological taxonomy, and if the zoological concepts do not fit plants this is of no interest. The reverse view is held by some botanists. The term species is employed almost exclusively for a genetic concept applicable only to sexually reproducing populations; its protagonists often seem to consider other creatures too aberrant to be worthy of serious taxonomic study. Very few writers have shown much appreciation of the scope of systematics, though some (for example Blackwelder and Boyden, 1952) have emphasized that it is a discipline central to all biology. In fact its basic concepts are not restricted to biology, for they are applicable to many other disciplines (see Section 10.6).

Contemporary taxonomy, while progressive in the establishment and revision of taxa, is decidedly conservative in its practices and philosophy. Published systematic work at the generic level and above is little different from that at the turn of the century. We have already examined one reason for this—the lack of a “new systematics” for the higher categories. A second, equally important cause, which for reasons of tact is rarely discussed, is the inadequate training and background of so many taxonomists. It is surely a reflection on the state of the science that the description and classification of organisms is today one of the few fields of biology to which amateurs without sound theoretical and practical training are still able to contribute. We do not wish to disparage the efforts of these amateurs, some of whom have made valuable contributions to biology in the past. Yet in taxonomy the bad work of both professional and amateur cannot be as readily ignored as it would be in other sci-

ences, because of the rules of priority and synonymy. As more sophisticated ideas and techniques percolate through systematics, amateurs not trained in modern taxonomy will of necessity find their sphere of activity progressively more circumscribed. It may be argued that amateurs will still be needed for purely descriptive (so-called *alpha*) taxonomy. However, as the methods of synthetic taxonomy become more elaborate the proper collection of data will become more involved, whether the trend of systematic research is in the biochemical or the biostatistical direction.

Militating against progress in taxonomy is the deplorable fact that success and prestige in biology seem to be largely associated with experimental work in fields such as genetics, biochemistry, or radiation biology. Hence the best brains and the best facilities in biological institutions (with a few exceptions) are not devoted to systematic work. This state of affairs is reflected in the dearth of systematists in the most prestigious academies. It may be that taxonomists are too overwhelmed with the sheer bulk of the material confronting them that still requires description and classification to have the time to work on a theory of taxonomy. However, we hardly feel that this is a case where *tout comprendre, c'est tout pardonner*. Without stopping to examine the cause-and-effect relationships involved, we may summarize this portion of our comments by stating that taxonomy (at least of the higher categories) has become in some ways a straggler in the progress of biological science.

Another deplorable feature of systematic research is the involvement of personal feelings in nomenclatural procedures. While some subjective bias of this sort is present in almost any science, taxonomy, alas, has more than its proper share. The problems that exist in this connection have been often discussed and are well known to systematists. Indeed, some notable improvements in their attitudes concerning this matter have taken place in recent years. Suffice it here to say that the prospects of gaining fame or notoriety (of however illusory a nature) by the naming or renaming of a category involve personal and subjective considerations which should be alien to scientific procedures.

One may ask how it is that taxonomists using such questionable methods have developed a body of knowledge which is admittedly useful and in certain groups consistent to a high degree? While the facts of the case are beyond dispute, its causes bear further examination.

The majority of taxa are definable because of the discontinuities arising in phyletic lines as by-products of the evolutionary process. (This point is discussed in greater detail in Chapter 8.) Relationships are recog-

nized and organisms are more or less correctly classified because taxonomy, like other sciences, is self-correcting, proceeding in a manner analogous to iteration in numerical analysis. A wrong initial evaluation of a set of characters *A* by taxonomist **X** is modified by further consideration of characters *B* by taxonomist **Y**, followed by taxonomist **Z** studying characters *C*. In this manner, as a natural and logical consequence of the taxonomic procedure, more and more characters are being considered and a given single character or set of characters assumes progressively less importance. Consequently misevaluations and wrong classificatory judgments are amended by subsequent work. Time, expressed through the activities of successive taxonomists, becomes an equalizer of character weight. This self-correcting nature of taxonomy is an extremely important feature and justifies to a large degree the dependence of the worker in the field on prior studies.

However, taxonomic statements at a given time may be strongly biased by the last revision of the taxa concerned. Should the revision be carried out with less than usual care and competence, considerable confusion may pervade the system for a long period. Many times the results of a study cannot be confirmed independently because of the inaccessibility of material or the lack of personnel or for both these reasons. Thus the baneful effect of poor taxonomic work may be protracted and far-reaching. It is for this reason that Michener and Sokal (1957) wrote:

Taxonomy, more than most other sciences, is affected by subjective opinions of its practitioners. Except for the judgment of his colleagues there is virtually no defense against the poor taxonomist. Above the species level there are no experiments that can be repeated and shown invalid, no mathematical or symbolic reasoning that can be demonstrated to be in error.

It is the hope of numerical taxonomy to arrive at judgments of affinity based on multiple and unweighted characters without the time and controversy which seem necessary at present for the maturation of taxonomic judgments.

2.2. THE NATURAL SYSTEM

Great difficulties have always accompanied attempts at defining a natural system. Danser (1950) realizes the difficulty of defining natural groups but is not able to state any exact or scientific definition for them, ending with the hope that “. . . some day systematics will arrive at a

more exact stage, but this does not alter the fact that already now we are entitled to face its problems, be it for the moment in a more intuitive but nevertheless scientific manner." Simpson (1961, p. 57) agrees that "in fact much of the theoretical discussion in the history of taxonomy has, beneath its impersonal language and objective façade, been an attempt to find some theoretical basis for these personal and subjective results."

The nature of a taxonomy depends on its purpose, as Gilmour (1937, 1940, 1951, 1961b) has emphasized. We could arrange living creatures in many ways, but we choose one way because we think it is best for some purpose. If the purpose is restricted, then the classification is a special classification, often called "arbitrary." Such a classification conveys less information than a general or "natural" one. For example, we can divide mammals into carnivores and herbivores for the purpose of ecology; then the designation "carnivore" only tells us the kind of food they eat. We hold the view with Gilmour that a "natural" taxonomy is a general arrangement intended for general use by all scientists.

The earliest attempt at systematics were based, as Cain (1958) has shown, on Aristotelian logic. This was the method used by early systematists such as Cesalpino and even largely by Linnaeus. The Aristotelian system as applied to taxonomy consisted in the attempt to discover and define the *essence* of a taxonomic group (what we may somewhat loosely think of as its "real nature" or "what makes the thing what it is"). In logic this essence gives rise to properties which are inevitable consequences: for example, the essence of a triangle on a plane surface is expressed by its definition as a figure bounded by three straight sides, and an inevitable consequence is that any two sides are together longer than the third. Such logical systems are known as systems of *analyzed entities*, and early systematists supposed that biological classifications could be of this kind. The terms *genus* and *species* had technical meanings in logic, and these were taken over into taxonomy. These points are well discussed by Thompson (1952) and Cain (1958, 1962). Aristotelian logic does not, however, lend itself to biological taxonomy, which is a system of *unanalyzed entities*, whose properties cannot be inferred from the definitions—at least not if the taxonomy is to be a natural one.

It is to the credit of John Ray and Caspar Bauhin that they were less bound by the iron fetters of Aristotelian logic than other early workers. They had a strong intuitive sense of what natural taxa were, although

they did not express themselves clearly. This is what de Candolle (1813, p. 66) aptly called "groping" (*tâtonnement*), though he only attributes this to later authors such as Magnol. According to de Candolle, Magnol claimed to have a clear idea of a natural family of plants even though he could not point to any one character which was diagnostic of the family.

This comment by Magnol and a similar comment by Ray (quoted by Cain, 1959c) were among the first admissions that it might not be possible to find *any* single diagnostic character for a natural taxonomic group. This is a point of the very greatest importance, which can scarcely be overemphasized. Michener (1957) supports this view for genera; others, among them Sneath (1957a), state that it is also true of natural taxa of *any* rank. While "artificial" or "arbitrary" taxa can indeed be defined by a single character, this is not necessarily true of natural taxa. Every systematist knows of instances where a character previously considered to be diagnostic of a taxon is lacking in a newly discovered organism which clearly belongs to the taxon. A striking example is the lack in some species of fish of red blood corpuscles (Ruud, 1954), hitherto considered to be an invariable attribute of all vertebrates. Fortunately, as Michener (1957) says, natural taxa generally do possess some distinctive characters in practice, although they need not do so in theory.

Biologists owe a debt of gratitude to Beckner (1959) for the first clear enunciation known to us of one important concept of natural taxa, a concept which Beckner calls "polytypic." Since this term and its converse, "monotypic," have meanings already well established in systematics, Sneath (1962) has suggested that "polythetic" and "monothetic" are better names (from *poly*: "many," *mono*: "one," *thetos*: "arrangement.") Simpson (1961, pp. 41-57) has given a discussion of the implications of these concepts in taxonomy.

The ruling idea of monothetic groups is that they are formed by rigid and successive logical divisions so that the possession of a unique set of features is both sufficient and necessary for membership in the group thus defined. They are called monothetic because the defining set of features is unique. Any monothetic system (such as that of Maccacaro, 1958, or in ecology that of Williams and Lambert, 1959) will always carry the risk of serious misclassification if we wish to make natural phenetic groups. This is because an organism which happens to be aberrant in the feature used to make the primary division will inevitably be moved to a category far from the required position, even if it is identical with its natural congeners in every other feature. The disadvantage

of monothetic groups is that they do not yield "natural" taxa, except by a lucky choice of the feature used for division. The advantage of monothetic groups is that keys and hierarchies are readily made.

A polythetic arrangement, on the other hand, places together organisms that have the greatest number of shared features, and no single feature is either essential to group membership or is sufficient to make an organism a member of the group. This concept was stated many years ago (for example, by Jevons, 1877, pp. 682-698). For its formal expression we cannot do better than to quote Beckner's definition (1959, p. 22):

A class is ordinarily defined by reference to a set of properties which are both necessary and sufficient (by stipulation) for membership in the class. It is possible, however, to define a group K in terms of a set G of properties f_1, f_2, \dots, f_n in a different manner. Suppose we have an aggregation of individuals (we shall not as yet call them a class) such that:

1) Each one possesses a large (but unspecified) number of the properties in G .

2) Each f in G is possessed by large numbers of these individuals and

3) No f in G is possessed by every individual in the aggregate.

By the terms of (3), no f is necessary for membership in this aggregate; and nothing has been said to either warrant or rule out the possibility that some f in G is sufficient for membership in the aggregate.

He then goes on to say that a class is polythetic if the first two conditions are fulfilled and is fully polythetic if condition (3) is also fulfilled. He points out that taxonomic groups are polythetic classes, but that polythetic concepts are by no means restricted to taxonomy or even to biology, for Wittgenstein emphasized their importance in ordinary language and especially in philosophy—the concepts of "meaning," "referring," "description," and so on. There is a close parallel between Wittgenstein's "family resemblance" and taxonomic affinity. As we have noted above, natural taxa are usually not fully polythetic, since one can usually find some characters common to all members of a taxon. It is possible that they are never fully polythetic because there may be some characters (or genes) which are identical in all members of a given taxon; even if there are many alleles or pseudoalleles of a gene, there may well be parts of the gene which are identical in all members. Nevertheless, we must regard a taxon operationally as being possibly fully polythetic, since we cannot be sure that we have observed any characters that are common to all members.

Beckner points out also the importance of condition (2). If, for exam-

ple, the various f 's are found in only one individual of the aggregate, then each individual will possess a unique subset of the f 's and will share no f 's with any other individual. Such a situation does not yield a polythetic class: for example, individuals 1, 2, 3, and 4, with the respective f 's ABC, DEF, GHI, and JKL, do not form a polythetic class. If, however, individual 1 possesses ABC, individual 2 possesses BCD, individual 3 possesses ABD, and individual 4 possesses ACD, the class of 1 + 2 + 3 + 4 is polythetic (and in this instance is also fully polythetic, since no one character is found in all the four individuals). This may be seen in an arrangement such as this one:

		Individuals					
		1	2	3	4	5	6
A				A	A		
B			B	B			
C			C		C		
Characters			D	D	D		
						F	F
						G	G
						H	H

Individuals 5 and 6, however, form a monothetic group.

One of the difficulties of Beckner's definition is that in natural taxa we do commonly have f 's which are not possessed by large numbers of the class. Furthermore, we cannot test whether any given f is possessed by large numbers of the class before we have made the class, and therefore we cannot decide whether to admit this f into the set G . This difficulty can be avoided by defining class membership in terms of common (or shared) attributes. Polythetic groups can of course themselves be arranged polythetically to give higher polythetic groups, as is done in building a hierarchy in the natural system. The advantages of polythetic groups are that they are "natural," have a high content of information, and are useful for many purposes. Their disadvantages are that they may partly overlap one another (so that hierarchies and keys are less easy to make than with monothetic groups) and that they are not perfectly suited for any single purpose.

An important practical difference between "classification from below" (the grouping of species into genera, genera into tribes, tribes into families, and so on) and "classification from above" (the division of the kingdoms into phyla, phyla into classes, and so on) is that the latter

process is inevitably largely monothetic. Classification from above therefore carries the risk that the divisions do not give "natural" taxa, yet it is necessary in practice in order to isolate a group of organisms of a manageable size for study. The important point is that a group under study may be incomplete or very heterogeneous; that is, some of its closest relatives may have been omitted, either through ignorance or because the forms have been misclassified.

A thorough early re-evaluation of systematics was made by Adanson (1727–1806), a botanist of independent and original views. His experiences in Senegal, where he found many new forms which would not fit into the then current taxonomic system, led him to seek the true basis of a natural classification. This he claimed to have found. He rejected the a priori assumptions on the importance of different characters (which were a consequence of Aristotelian logic); he correctly realized that natural taxa are based on the concept of "affinity"—which is measured by taking all characters into consideration—and that the taxa are separated from each other by means of correlated features (Adanson, 1763, pp. clv, clxiv). The method he used was very cumbersome. He made a number of separate classifications, each based on one character, and examined them to find which classifications divided up the creatures in the same way. These classifications he took as indicating the most natural divisions, which were, of course, therefore based on the maximum correlations among the characters. The important corollary was that by treating every character in the same way he was in effect giving them equal weight. His contemporaries attacked him on this very point (see de Candolle, 1813, pp. 70–72), without realizing that their own beliefs on the relative importance of various characters, far from being based on a priori assumptions as they imagined, were in fact a posteriori deductions from intuitive taxonomies of precisely the kind Adanson was recommending (Sneath, 1957a; Cain, 1959a,b). This is an example of the self-correcting nature of intuitive taxonomy (see Section 2.1). Adanson's earliest work in this direction was on molluscs. In his treatment of this group, in *Histoire naturelle du Sénégal*, he says that since the previous arbitrary systems are unsatisfactory he will use a different method, first drawing up careful descriptions of all parts of the shell and of the body of the mollusc. He then makes the following observation (1757, p. xi):

Je me contenterai de rapprocher les objets suivant le plus grand nombre des degrés de leurs rapports & de leurs ressemblances . . . Ces objets ainsi réunis, formeront plusieurs petites familles que je réunirai encore

ensemble, afin d'en faire un tout dont les parties soient unies & liées intimement.

He gave a number of tables in which the molluscs were arranged according to a score or so of characters (Preface, pp. xxix–lxxxxviii), which foreshadowed the more elaborate tables in his botanical work of 1763. He was already aware that one could not choose diagnostic characters of genera while they were incompletely known, since new species might prove to be exceptional (Adanson, 1757, Pref., p. xx), a point which was also realized by Linnaeus. No other workers, except perhaps Vicq-d'Azyr (1792) and Whewell (1840) seem to have followed up Adanson's ideas until recently.

Vicq-d'Azyr clearly enunciates Adanson's principles (though without mentioning him) in the introduction to his work on comparative anatomy. He says that a natural class is the result of assembling species which have a greater number of resemblances to each other than they have to species in other classes, and continues “. . . d'ou il résulte qu'il seroit possible qu'une class fût très-naturelle, et qu'il n'y eût pas un seul caractère commun à toutes les espèces qui la composent.” Whewell's idea of taxonomic affinity was evidently close to Adanson's, since he makes the point that natural classes must possess many correlated common features, although he repeats de Candolle's criticisms of equal weighting.

We may ask why Adanson's method, though excellent in theory, was a failure in practice. Stearn (1961) considers that the material available in Adanson's day was too limited to allow of success, and we may add that such methods were quite impracticable before the advent of computing machines. Nevertheless, as de Candolle admitted, Adanson's taxa were for the most part more natural than earlier arrangements.

In the pre-evolutionary days of systematics it had been found empirically that a nested, hierarchical system gave the most satisfactory and “natural” arrangement of the data. Such a system could generally be constructed on the basis of a few characters. The art of the practice lay in finding suitable characters, to prevent the classification from creating strange bedfellows, clearly incongruous as judged by their great differences in other characteristics. There was little attempt either to understand why this should happen or to discover the rational method of choosing the “right” characters. We discuss below the development of the understanding of what it is which makes taxonomic groups “natural” and how it is possible *after creating such natural taxa* to discover characters

which are suitable for discriminating between them. Even at the time of Adanson the system must have been inadequate, since it did not provide for an effective evaluation of resemblance among the various forms. It is true that an approximate idea could be obtained from the classificatory scheme, since members of the same genus were more likely to resemble one another than were members of different genera. However, without an unmanageable proliferation of taxonomic categories, which were in any event of dubious validity at that time, the fine shades of difference and resemblance could not be shown.

Until the impact of the theory of evolution, the subsequent development of systematics took place largely in France (de Candolle, A. L. de Jussieu, Cuvier, and Lamarck) and was in the direction of greater sophistication on the theme of the coordination of characters into a harmonious whole. This was carried even to the point of implying that a whole animal could be reconstructed from one bone. One can, of course, identify a known animal from one bone, but to reconstruct from it a new animal with all its soft parts is a feat of a different order, as Simpson (1961, p. 44) points out.

The advent of the theory of evolution changed the practice of systematics very little, although the professed philosophical basis of systematics was radically altered. Natural classifications were considered to be those established on the basis of monophyletic taxa. (This whole question is discussed in detail in Section 2.3.)

The present century has witnessed a re-examination of the validity of the evolutionary basis of natural taxonomies. Empirical classifications have been proposed in several quarters. The conceptual basis of natural taxonomies was discussed from the standpoint of logic in a classic paper by Gilmour (1937) and expanded in later works (Gilmour, 1940, 1951, 1961b). He pointed out that logicians have long realized that the central idea underlying "natural" groupings is the great usefulness of a method which can group together entities in such a way that members of a group possess many attributes in common. Indeed, we maintain that the elusive property of naturalness is simply the degree to which this principle obtains. The idea of overall similarity follows from this and is a function of the proportion of attributes shared by two entities. As Gilmour points out, natural classifications are not restricted to biological ones (see Section 10.6). In addition, intermediate situations can occur between the highly natural (such as the class *Mammalia*) and the wholly artificial (such as creatures whose generic names begin with the letter "A"). An

example of a partly natural group is the group which gardeners call "Alpines."

The reason for the great usefulness of natural classifications is that when the members of a group share many correlated attributes, the "implied information" or "content of information" (Sneath, 1957a) is high; this amounts to Gilmour's dictum, that a system of classification is the more natural the more propositions there are that can be made regarding its constituent classes. Remane (1956, p. 4) tries to show that the predictive value of taxonomic groups is only true of natural taxa, not of artificial ones. It is obvious that artificial groups established on a single character are of low predictive value. Nevertheless, such groups may by chance prove to be partly natural, since such a single character may be highly correlated with the other characters of the taxa in question. It would be possible to devise a measure of the extent to which this is true of any character in any given taxonomic system. The techniques of Rescigno and Maccacaro (1960) are steps in this direction. A natural classification can be used for a great variety of purposes, while an artificial one serves only the limited purpose for which it was constructed. As Sneath has emphasized (1958), natural or "general" classifications can never be perfect for all purposes, since this is a consequence of the way we make natural groupings. By putting together entities with the highest proportion of shared attributes, we debar ourselves from insisting that they shall share any particular attribute, as a very simple trial would show. This is the reason for emphasizing the historical importance of the realization that natural taxa do not necessarily possess any single specified feature. This spelled the doom of the Aristotelian concept of an *essence* of a taxon, for natural groups are in logic *unanalyzed entities* (see Cain, 1958). Simpson (1961) rejects as illogical the contention by Gilmour (1951) that a classification serving a large number of purposes will be more natural than one which is more specialized and that the most useful and generally applicable classification will be the most natural one. We feel that Gilmour's usage corresponds to the intuitive sense of naturalness which taxonomists have possessed since even before Darwin. Gilmour's dictum—that a system of classification is the more natural the more propositions there are that can be made regarding its constituent classes—admits of objective measurement and testing, in contradistinction to Simpson's natural system. Furthermore, Gilmour's system has powerful predictive properties; it is therefore to be recommended. We believe that it will eventually be shown that, with few

exceptions, monophyletic taxa will also be most natural in the sense of Gilmour and that therefore the two concepts will emerge as substantially identical. If this is so, phylogenetic conclusions may eventually be drawn from a demonstration of naturalness, *sensu* Gilmour.

2.3. PHYLOGENETIC CONSIDERATIONS

We have seen in the previous section how the pre-Darwinian biologists interpreted a natural system. With the publication of the *Origin of Species*, however, the entire problem was viewed from a different perspective. Suddenly the reason for the existence of natural systematic categories became apparent: *their members were related because of descent from a common ancestor!* A taxon was now interpreted as a monophyletic array of related forms.

Science always receives a new impetus when a well-known and ordered body of facts can be given meaning by a new, consistent hypothesis. The present instance was no exception, and under the influence of the evolutionists, led by Haeckel, the fields of systematics and comparative morphology were reoriented phylogenetically. It has, however, been frequently pointed out (as by Bather, 1927, and Remane, 1956) that this change of philosophy did not bring with it a change in method. Taxonomy proceeded as before; only its terminology had changed. Remane (1956) quotes Naef (1919, pp. 35–36):

. . . und was Haeckel und die Phylogenetiker zunächst getan haben, war nichts anderes als die Übersetzungen der speziellen Einsichten, die sich an diese Lehre früher geknüpft hatten, in eine Sprache durch Anwendung einer neuen Terminologie, ohne doch die Lehre selbst einer Vertiefung zuzuführen oder einer kritischen Betrachtung zu unterwerfen. Auch die—wenig abgeklärten—Grundbegriffe der alten Morphologie wurden von Haeckel einfach in die neue Sprache übersetzt, die dem Wesen nach eine genealogische war. Dabei wurde dann

aus Systematik	Phylogenetik,
aus Formverwandtschaft	Blutsverwandtschaft,
aus Metamorphose	Stammesentwicklung,
aus systematischen Stufenreihen	Ahnenreihen,
aus Typus	Stammform,
aus typischen Zuständen	ursprüngliche,
aus atypischen	abgeänderte,
aus niederen Tieren	primitive,
aus atypischer Ähnlichkeit	Konvergenz,
aus Ableitung	Abstammung usw. usw.

It is also noteworthy that T. H. Huxley, writing at the height of his protagonism of Darwin's evolutionary theory, was perceptive enough to avoid confusing phylogeny with classification (Huxley, 1869).

The difficulty with the use of a phylogenetic approach in systematics emerged after the first wave of enthusiasm for it had subsided and has remained apparent to perceptive observers ever since. *We cannot make use of phylogeny for classification, since in the vast majority of cases phylogenies are unknown.* This is one of the statements most commonly heard at meetings of taxonomists, yet it is most consistently ignored. Let us restate it in other words for emphasis. The theoretical principle of descent with modification—phylogenetics—is clearly responsible for the existence and structure of a natural system of classification; we may even agree with Tschulok (1922) that the natural system can be considered as proof of the theory of evolution. However, since we have only an infinitesimal portion of phylogenetic history in the fossil record, it is almost impossible to establish natural taxa on a phylogenetic basis. Conversely, it is unsound to derive a definitive phylogeny from a tentative natural classification. We have described this fallacy of circular reasoning earlier.

In recent years three comprehensive analytical studies of systematic principles have been published in books by Hennig (1950), Remane (1956), and Simpson (1961). It is especially regrettable that the earlier two books, published in German, have been almost entirely ignored in the English and American literature. Hennig's book presents the issues with particular clarity and objectivity, and there is considerable truth in Kiriakoff's (1959) statement that a number of controversies of the last decade published in the United States are in a sense outdated and could have been guided into more productive channels if Hennig's thoughts had been available to the disputants.

All three authors mentioned above are fully aware of the dilemma of circular reasoning inherent in systematic procedure. They are not satisfied with solutions based on "groping." Simpson (1961) thinks that taxonomy is an evolutionary science, and he attempts to outline a series of phylogenetic principles on the basis of which taxonomic evidence should be examined to yield evolutionary interpretations and classifications. We shall examine these principles in detail later in Chapter 8. However, Simpson nowhere in his book is able to present a logical and consistent defense for the circularity of reasoning inherent in such procedures. By calling the process of classification an art, rather than a science, he defines the problem out of existence.

Hennig (1950) describes the dilemma in even greater detail. He

defends the circularity of reasoning by the "method of reciprocal illumination." By this he means that some light is thrown from one source of logical illumination onto a natural situation, kindling another, brighter light in the latter, which in turn will throw added illumination onto the first source. Thus, in a self-reinforcing, positive feedback type of analysis, the relationships under study are eventually clarified. Hennig feels that phylogenetic relationships are the entity of systematics whose parts consist of morphological, ecological, physiological, and zoogeographic similarities. Each of these parts mirrors phylogenetic relationships, which are to be investigated by the method of reciprocal illumination. But we cannot see how the principle of reciprocal illumination differs from the much-condemned vertical construction of hypothesis upon hypothesis.

Remane (1956), in spite of his fundamentally phylogenetic orientation, has also realized that phylogenetic reasoning cannot serve as the basis for erecting a natural system. He similarly rejects affinity (based on few characters) as the basis of a natural system. He considers that while both of these approaches enter on occasion into the techniques practiced by "good systematists," the exclusive application of only one of them is likely to lead to misclassification. Affinity or resemblance when based on one or few characters can lead the systematist astray, Remane claims, as he would be too easily deceived by chance convergences resulting from poor sampling of the characters. Remane attempts escape from the *circulus vitiosus* by basing his taxonomy on nonphylogenetic criteria of homology.

In spite of the differences in fundamental outlook among these three authors, their actual approach to classification is quite similar and is akin to that employed by the majority of competent taxonomists today (or 100 years ago, for that matter). The procedure is difficult to define and delineate; attempts to do so of necessity illuminate the faults of the system. The method will often lead to roughly correct (that is, reasonable) results because of happily correct intuition or the self-correcting features of the classificatory process already discussed.

The basis of the classical method of establishing taxa is commonly held to be the recognition of homologies. Similarly, the recognition of analogies will serve to separate artificial taxa. Success in distinguishing homologies from analogies will therefore reflect success in systematics. In determining which characters are homologous (of common descent) and which have been independently evolved, the systematist has to express a judgment on the relative probability of the independent origin of

different character complexes. Sometimes, as in color patterns, knowledge of the genetics of the character in the forms (or in related forms) may be invoked. More often the worker has to weigh independently the relative improbabilities of various phenetic rearrangements in order to arrive at a working hypothesis. Thus, indigo (as a glycoside) must be of independent origin in such widely scattered families of plants as Leguminosae, Cruciferae, Acanthaceae, Orchidaceae, and Polygonaceae. This seems overwhelmingly more probable than that these taxa together form a natural phenetic and monophyletic taxon—the family “Indigoaceae.” Luminescence has presumably arisen repeatedly and independently in algae, bacteria, fungi, crustacea, insects, myriapods, tunicates, coelenterates, annelids, molluscs, and fishes, with apparently very similar biochemical mechanisms in each case so far examined. Chlorophyll in bacteria has presumably arisen independently of chlorophyll in higher plants, and hemoglobin has arisen in annelids, molluscs, crustacea, and insects as well as in vertebrates. Entomologists have decided that social organization arose independently in the termites, bees, wasps, and ants, and evidence is accumulating that it arose independently several times in the bees. Social parasitism in bees and slave-making in ants appear to have arisen independently a number of times. But the occurrence of symbiotic flagellates in the guts of termites and the roach *Cryptocercus* appears to be traceable to an ancestor carrying such flagellates. Even judgments of this kind rely heavily on estimates of similarity between the organisms or structures concerned.

Classifications are only as good as the homologies of the characters on which they are based. Furthermore, decisions on homologies of certain characters are based upon the validity of the classification of the groups involved in the argument; this classification in turn is based upon homologies of other characters used to establish the classification *ab initio*. When the circular arguments are interrupted we are left with much uncertainty. We feel therefore that the operational homology of Adansonian taxonomy involving fewer assumptions (see Section 5.3.4) is to be preferred to the classical methods.

Further difficulties are presented by cases of convergence—that is, by the resemblances of certain subordinate taxonomic groups in *different* higher taxa. If no account is taken of this convergence and the classification is based on descent alone, the biologist inspecting the nomenclature or a family tree of the group will have no idea of the phenetic closeness of the taxa concerned. Such resemblance, while considered “superficial” by the phylogeneticist, should not be lightly dismissed. It

does represent important genetic changes and presumably important physiological and ecological similarity. But if the resemblance of the two groups is considered in establishing the classification, erroneous conclusions may be reached about common ancestry or recency of separation of the stems.

As soon as phylogenetic considerations were added to systematics, three new questions arose. What are the phylogenetic relationships, or which stem branched off where? When in geological or relative evolutionary time did a given branching take place? How rapid was the evolutionary rate of a given line in a given time period? Classificatory theory and procedure, already inadequate, could not simultaneously accommodate these differing aspects of the problem.

The first of the new questions raised was emphasized the most. Undoubtedly more utter rubbish has been written since the time of Haeckel on supposed phylogenies than on any other biological topic. The fact is that we have a reasonably correct picture of the phylogenies of only a very few taxa and these entirely on the basis of paleontological evidence. Even in paleontology the proportion of fact to speculation is not too high. We quote a recent critic in this field (Challinor, 1959): "Works which refer to the fossil evidence of evolution usually cite a few of the well-known cases of evolutionary series as if they were merely representatives of a host that might have been quoted, instead of stressing the fact that records of such cases are rare."

Speculation on phyletic relationships based on neontological evidence is very questionable indeed. Sporne (1956) has discussed this in detail in the case of plant phylogeny, particularly the pitfalls in interpreting continuous series of characters. The well-known law of recapitulation, by which the successive embryonic states of an animal are said to mirror its phylogeny, is now realized to be open to similar misinterpretations, since there is no way of being certain which embryonic features do and which do not reflect the actual phylogeny (see George, 1933). Similar difficulties apply to other laws, such as Dollo's law, and to chromosomal karyotypes (Simpson, 1953, pp. 245-259, 310; Stebbins, 1950, pp. 442-475, 498). While both positive and negative statements of a very general nature can be made with a reasonable probability of correctness—as that vertebrates did *not* descend from spiderlike ancestors, that mammals evolved from reptilelike ancestors—detailed establishment of relationships by descent are likely to be quite fanciful. What evidence on the development of piston engines would be yielded by a present-day comparative study of such varying machines as an airplane engine, a

diesel engine, and a car motor? Would we be able to reason back to the early steam engines? It may be argued that "living fossils," which are often transitional forms, help us in establishing a possible reconstruction of the chain of events. In the analogy employed by us, these living fossils would be old steam and other engines found in abandoned mine shafts or still working in primitive regions of the world. Such evidence would surely help in raising the confidence which we could place in any given evolutionary construct. However, the improvement brought about by living fossils is essentially one of degree. Thus an unusual number of relict forms would have to occur before we could establish our hypothesis with certainty. We would first have to assure ourselves that these were indeed relicts: it is no easy matter to recognize a living fossil before any fossils are known. Even then it is unlikely that we would understand the detailed derivation of every form.

Hennig (1950), aware of the above criticisms of the phylogenetic approach to taxonomy, justifies such a procedure on the basis of three main arguments. The first is that the phylogenetic system is the most meaningful of all possible systems because all other types of classifications, such as ecological, zoogeographic, or morphological, can be derived and explained through the phylogenetic system. In a sense this argument is analogous to that of Gilmour, who states that a natural system is the one of most general application. Indeed, none of the special classifications could occupy such a central and all-explanatory position as does a phylogenetic system. The theory of evolution is the most adequate, most unitary, and indeed simplest hypothesis to which a great variety of biological phenomena—geographic distribution, physiological adaptation, morphological similarity, or biocoenotic complexity—can be related. Phylogeny can thus be seen as the central cause of much biology, yet it cannot be used for an explanatory concept, as it is not known in the vast majority of instances. Hence an empirical classification, although it may not be able to explain the above-mentioned biological phenomena, is at least a self-sufficient, factual procedure and may in most cases be the best classification that we can get.

Hennig's second reason for preferring a phylogenetic taxonomy has been negated by the development of numerical taxonomy. He thought that phylogenetic relationships are at least in principle measurable, but that similarities are not. The very opposite opinion is held today.

Third, Hennig claims that there is no exact correspondence between morphological similarity and phylogenetic relationship and that convergence may mask phylogenetic relationships. He feels that of the two

relationships the second is the more important. It remains to be demonstrated which of the two relationships would be more pertinent to the taxonomist (in this connection, see Sneath and Sokal, 1962). In any case this is a field unexplored in theory or practice. Until it can be shown through plausible models to what degree phylogenetic relationship can differ from morphological (phenetic) relationship or until a case of known phyletic history can be used to explore quantitatively the correspondence of morphological with phylogenetic relationships, judgments on such issues should be suspended in favor of research upon them.

Careful thought must also be expended on the meaning of rate of evolution (or of divergence). An overall rate of divergence based on some multivariate summary or abstraction of the phenotype can be given, if the time elapsed to achieve this divergence is known. In many cases divergence may be very rapid but may only involve a single organ or organ system. The overall divergence of the descendent forms from the ancestral stocks may not be impressive; yet the actual rate of change for the organ system concerned may be very high. Such situations can very easily give rise to errors of estimate regarding times of branching. As judged by its feathers, *Archaeopteryx* was a bird; as judged by its skeleton, it was still a reptile. Evolution of new food preferences or color adaptations to different backgrounds may be in this category.

Difficulties arise often when organisms exist in phenetically different life forms. These could be immature forms and adults (as in many insects), males and females (many birds), or different morphotypes occurring in alternation of generations (aphids, Sporozoa, plants) or through social differentiation (termites, ants). Sometimes a satisfactory classification can be based on one of these forms but not on the other. More disturbing are cases where separate classifications are possible but do not match. Such examples are conventionally explained away by bad choice of characters and by nonrecognition of homologies (Remane, 1956). In view of the unreliability of classifications based on few characters, it seems to us surprising that more such incongruences, as Weismann termed the phenomenon, do not occur.

Customary procedure in dealing with incongruences is objectionable: the taxonomy of the later described stage is constantly compared with the earlier classified one. Thus almost inevitably the earlier classification will influence the later one. A thorough investigation of this phenomenon has yet to be made, but it is our belief, defended in Chapter 5, that similar classifications would result from an analysis of sufficient numbers of characters in different life forms of identical taxa, at least as a rule.

If it is the aim of systematics to structure the organic world phenetically and phyletically, then we must regretfully conclude that no currently available taxonomic method can achieve this simultaneously. Michener and Sokal (1957) have said that classification cannot describe both affinity and descent in a single scheme, and this point must be re-emphasized.

Verbal or written descriptions of relations among organisms have proved quite inadequate. For this reason a variety of mnemonic and didactic aids have been developed, most of them graphic. These are largely different forms of trees of relationships (or phylogenetic trees). Mayr, Linsley, and Usinger (1953) have called these drawings *dendrograms*, which seems a suitable term without any implication about the nature of the relationship.

A dendritic description of the taxonomic system has much to recommend it and seems in many ways to be the "natural" way of illustrating relationships and descent. Other schemes of presentation have been tried repeatedly but unsuccessfully (Remane, 1956). A vague, general agreement on the interpretation of diagrams of relationship exists among taxonomists, yet when a given diagram is subjected to detailed, critical scrutiny we rarely find consistency of meaning within it. The interpretation of the basic facts which a diagram offers is likely to be based on varying degrees of certainty in different parts of the tree. No generally accepted conventions for constructing such diagrams exist; hence, seemingly similar diagrams may have quite different meanings, which are often not clearly enunciated by the author of the tree. Cain and Harrison (1960b) have shown in an illuminating discussion on the different components that can be included in phylogenetic relationships, that often an author has not himself a clear idea of the meaning of a diagram of relationships presented by him.

The following symbolisms have been used most frequently.

(1) The vertical axis (or radius in circular dendrograms)—to indicate time, either in absolute units or in relative evolutionary ones (most frequently unspecified).

(2) Furcations—to indicate branches in the phyletic sequence in order to indicate the relationships between the forms based on the lineages alone (not considering their phenetic similarities); that is, the *cladistic* affinity (Cain and Harrison, 1960b).

(3) Location and relative position of tips of branches with respect to each other—to indicate (phenetic) relationships.

(4) Location of furcation along a vertical axis (which now designates resemblance)—to indicate closeness of relationships between taxa represented by stems issuing from the furcation. Symbolisms (3) and (4) are often used in combination to indicate what part of the relationships is due to convergence and what part is due to inheritance of characters of a common ancestor (*patristic* affinity, Cain and Harrison, 1960b).

(5) Levels of tips along a vertical axis (or along the radius of a circular dendrogram)—to indicate whether the forms are recent or extinct, and also to give some estimate of the time scale of the extinction.

(6) Levels of tips along a vertical axis (or along the radius of a circular dendrogram)—to indicate degree of perfection or complexity of form. This convention, related to the *scala naturae* of an earlier day, is largely out of fashion, although some of its ideas and its vocabulary are still employed occasionally. Thus Rensch (1947) uses the term *Höherentwicklung*.

(7) The angle between stems—to represent velocity of (phenetic?) differentiation.

(8) Thickness of stems—to represent abundance at a given point in time. Abundance is usually measured by the number of species or taxa contained within the stem, but occasionally represents the number of organisms supposedly extant.

It is easily seen that (1), (2), and (5) can be combined into a single diagram. Unless the rate of evolution has been constant, (4) cannot also be included. It is generally impossible to represent phenetic relationships on a two-dimensional graph; hence (3) is bound to be a distorted representation. The use of the angle to indicate velocity of evolutionary change (7) is never very successful except in the simplest diagrams. Abundance (8) can usually be added to most diagrams, although the results are often not very esthetic. The basic difficulty is the graphic representation of phenetic resemblances and phenetic change. These are multidimensional relationships and cannot satisfactorily be compressed into a two-dimensional diagram. An interesting contribution by Hayata (1921, 1931), discussed at some length by Du Rietz (1930), emphasizes the multidimensional nature of taxonomic resemblance, with the added implication that evolution may be much more reticulate than is commonly thought. The only possible way of transmitting the various types of information listed in points (1) through (8) is by three separate graphs for (a) time and branches (cladistic relationships), possibly combining symbolisms 1, 2, 5, 7, and 8; (b) phenetic relationships between junc-

tions of stems only, as customarily employed in dendrograms in numerical taxonomy—symbolism (4); (c) complexity of form or organization—symbolism (6). Phenetic relationships among tips of branches (symbolism 3) can usually not be represented in a two- or three-dimensional space. Proper representation would require two-dimensional cross sections through the hyperspace which is necessary to represent such relations properly (see Section 6.2.3). We shall consider the mechanics of such a presentation in greater detail in Section 7.4.

As regards aspect (2)—phyletic sequence—it is obvious that a diagram can be constructed only if phylogenetic evidence can be obtained from fossils or in some other reliable fashion. We have already pointed out the dangers of deductive reasoning in tracing phylogenies. The sequences in phyletic lines are often much more uncertain than authors wish to admit. Some authors indicate probable descent by dotted lines. If there are many such dotted sections, the chances of the diagram's being substantially and seriously misleading may be very high indeed. Unfortunately there seems to be no study on this point to tell how misleading earlier phylogenies have been when compared with later detailed and convincing fossil evidence. Such a study might be illuminating. It is true that many authors, quite properly, disavow any phylogenetic significance of their diagrams and caution readers against considering them to be in any way reflections of evolutionary history. We ourselves follow conventional practice in arranging taxa by a system of hierarchic, nested categories which roughly give an indication of point (4). These can easily be represented by a table (such as Table 1 in Michener and Sokal, 1957); a dendrogram is another acceptable form and has been used by Michener and Sokal (1957, Figures 5 through 8; Figures 12 through 15 in the same paper do not fall into this category, being attempts at phyletic interpretations). The form of dendrograms, together with the intellectual traditions of present-day zoologists, makes it very difficult to view them without some evolutionary interpretations.

The criticisms of phylogenetic taxonomy enumerated above (and many more) have been voiced repeatedly and by a large number of writers during the past hundred years. An adequate summary of them would require a volume thicker than the present one. In return, these criticisms have prompted defenses of the phylogenetic approach, among the most recent being those by Hennig (1950, 1957) in Germany and by Simpson (1961) in the United States. One might expect that after a subject has been discussed for so long a period of time some agreement would have been reached on the relative merits of the various points of

view put forward. This does not appear to be the case, however. In our view, a major difficulty in which the critics of the phylogenetic method have found themselves in the past is that though being able to criticize the position of the phylogeneticist on valid grounds, they have been unable to suggest a consistent and workable alternative procedure. Such a goal is now within reach, as we hope to demonstrate in the chapters that follow. For this reason we have kept our criticism of current, largely phylogenetic principles of taxonomy relatively brief and shall fill the greater part of this book with a positive statement of our views, in the belief that the inherent faults of the phylogenetic method will thereby be most clearly shown. Further discussion of phylogenetic aspects of systematics can be found in Sections 4.6, 4.7, and 5.5 and in Chapter 8.

2.4. PROBLEMS OF TAXONOMIC RANK

2.4.1. Criteria for taxonomic rank

Phenetic as well as genetic criteria for taxonomic rank are commonly used. At and below the level of biological species they may be in conflict. If a genetic criterion is used to define species, there may be a variable species whose members interbreed freely; yet an adjacent group of equal variability may be split by genetic isolating mechanisms into several distinct biological species of much smaller internal variability (for example, sibling species). The phenetic ranks of the species are very different in the two groups. In the absence of data on breeding and in all apomictic groups (which include the great majority of practical applications in systematics), the species are based on the phenetic similarity between the individuals and on phenotypic gaps. These are assumed to be good indices of the genetic position, although they need not be. The rank of higher categories must perforce depend on phenetic and not genetic criteria. The intrusion of an entirely different criterion for taxonomic rank in those few situations where genetic or phyletic relations are known with certainty, seems to us to be a needless source of confusion. Alternative terminologies have been suggested for genetic and also ecological entities (especially in botany, where these problems are most acute); these terms have not been widely used, possibly because of the prestige attached to the term "species" (compare Gilmour and Gregor, 1939; Camp and Gilly, 1943; Gilmour and Heslop-Harrison, 1954). For reasons of clarity it is desirable that the meaning in which taxonomic rank is used should be specified. In this book it will be used

in the sense of phenetic rank, unless otherwise indicated. We have not attempted to define and distinguish different usages of the term species; where it is germane to our argument we have qualified the term to avoid confusion.

It is undesirable for the rank of a group to be affected by the number of contained subgroups. There is a modern tendency to make each family contain only a few genera and each genus only a few species; in some works most genera are monotypic. The rank should be based on affinity alone. In our view it is better to introduce new rank categories (such as subfamily or superclass) than to use the number of contained subgroups as an arbiter of rank.

There has indeed been a great elaboration of such new categories of rank, largely independent of the cycles of "lumping" and "splitting" which Simpson (1945) has commented upon. This practice has not always been justifiable, for although the traditional categories of phylum, class, order, family, genus, species may be inadequate, it seems labored to use forty or fifty new categories without any numerical justification; such new categories have contributed little other than being handy containers for speculative views on minutiae. Although careless creation of new categories of rank is undesirable, we believe that when they are necessary numerical taxonomy will afford good evidence on which to base them.

2.4.2. Limits of taxa

Limits of taxonomic groups can be considered from two points of view. One can trace the change of taxa with time, looking at the phylogenetic tree in its entirety. This is what Simpson (1945) has called vertical classification. One can also look at a cross section of the tree and obtain the relationships among taxa at a given point in time (horizontal classification, Simpson, 1945). We shall first consider the problems in vertical classification, which can only be practiced on fossil material.

It is obvious that when one phyletic lineage evolves into a new form there can be no sharp division between the ancestral and the descendant species, other than an arbitrary one, except in the case of allopolyploids and other forms of hybrid origins. This has been well treated by Bather (1927), who discusses the various ways in which the division can be made. It is inevitable that the accidents of discovery of fossil forms should affect classificatory decisions, since the divisions will at first be placed where there are gaps in the fossil record. As the gaps are closed by new

discoveries, the most common practice is to choose for the dividing line some prominent, but commonly arbitrary, evolutionary step—for example, the change in jaw structure in the evolution of reptiles into mammals. So long as the arbitrariness is clearly realized, these methods are unobjectionable and are matters of convenience. The choice of such an arbitrary step is not without some danger, however, for it may lead to incongruous situations.

Yet a better plan, commonly advocated in paleontology when a relatively full fossil record is known, is to place the divisions at places where abrupt changes in the rate or direction of evolution make for rational and convenient groups or where phyletic lines branch. The demerit of this course is that the divisions are then through parts of the lineages which are of special interest for students of evolution. Nevertheless, the bulk of the total material will be grouped in a convenient way, and it is standard practice in analogous situations to make the divisions on the same principle (such as where a small amount of hybridization occurs between living species, even if it is the hybrids which are of most interest to us). An acceptable nomenclature for the borderline forms, such as *X-Y* intermediate forms, or *X-Y* hybrids, is then the main problem.

As is discussed later (Section 8.2.1), the development of numerical taxonomy may allow us to find in the fossil record the points of abrupt evolutionary change and diminish this dependence on arbitrary evolutionary steps. Yet in practice there are few fossil series of such completeness that they warrant altering the rather simple treatment mentioned above.

To turn to horizontal classification, much of the difficulty lies in the definitions which we adopt or in the kinds of grouping we recognize. A phenetic taxonomic group may not always be identical with a phyletic group. For example, the appearance of a sterility barrier will at once divide a normal genetic species into two sibling species. Yet for many generations (until the two sibling species have accumulated sufficient genetic differences in the course of their independent evolution) they may remain one single phenetic group because the differences which cause the sterility barrier (plus the few other accumulated differences) will be insignificant in comparison with the many variable attributes of other kinds which the two sibling species will share. It may be unwise to call this phenetic group a single taxon without qualifying this latter term, but it will certainly be a single phenetic taxon, and it is also a single natural taxon, where the word “natural” has the restricted

meaning discussed in Section 2.2. Without the use of a more precise definition of the term species, the situation is certain to become confused. Such situations are evidently not uncommon; closely similar species may overlap in all observed features, as suggested by investigations such as those of Ehrlich (1961c) on butterflies and Lack (1947, pp. 82–86, 88–89) on Darwin's finches in the Galapagos Islands. It is notorious that many birds are classified almost only upon their skins, and the acknowledged success of the taxonomist is, as Lack points out, in large measure because the species recognition marks, which enable the birds to recognize mates from their own species, are usually visual features which the taxonomist also uses to distinguish these species; where visual marks are absent (as in some warblers which recognize one another by song) the classification is more difficult (Lack, 1947, pp. 16–54).

If a difficulty in assigning limits to taxa can occur when the facts are not in dispute, it is no wonder that confusion is common when the facts are uncertain. Numerical taxonomy may not solve the genetic problems, but by making precise the phenetic groupings it will help their solution.

One of the more obvious principles of delimiting taxa is that we place divisions at places in the taxonomic scheme which are empty—that is, where there are no known creatures. This is a corollary of the concepts underlying natural taxonomies, for it is these gaps in the universe of possible character combinations which give us the correlations between features on which the natural taxonomic groupings are based. But this itself creates its own difficulties. As Michener (1957) has emphasized, the gaps may only be gaps in our knowledge of living or extinct forms, and no consistent treatment has been developed to deal with this problem.

It is widely acknowledged that it is unwise to recognize taxa which are only differentiated on one or two features, though in some branches of biology systematists find this temptation hard to resist. Even if it is practicable to name endless varieties of this sort, the fact that they are usually established on a few features raises the suspicion that they are quite arbitrary taxonomic groups. How many geographical races, color variants, or other forms in mammals, birds, and butterflies can stand critical examination in this regard? This is the substance behind the criticism of Wilson and Brown (1953), and most of the current systematics of bacteria and yeasts is unsound for the same reason.

Low taxonomic ranks may be difficult to define and to arrange hierarchically. Such groups would appear in a numerical taxonomic study as contiguous and indistinct clusters of individuals. Methods for

best dividing the organisms into clusters have only recently been considered, and at present there is no consistent practice among taxonomists. Procedures such as "the 75% rule" are often ambiguous (Pimentel, 1959), since different results may be obtained using different features. If a single variable feature is employed the divisions are, in general, not natural taxa, and the divisions may not correspond at all to phenetic clusters.

2.4.3. The hierarchy of characters

One fallacious argument, happily now on its way out, is the theory of the hierarchy of characters. By this is meant the claim that one can lay down a priori rules as to which sorts of characters separate species, which sorts separate genera, which sorts separate families, and so on. This is the antithesis of the a posteriori method of discovery. The latter *finds* those features which do in fact separate the previously recognized natural taxa. It is, we maintain, the correct procedure. The hierarchic claim is another form of the theory of unequal and a priori weighting. It is very old; in fact, as Cain (1958) has shown, it springs from the Aristotelian theory of "essences." He cites Cesalpino, who stated that nutrition was of paramount importance to plants, and hence, a priori, the form of root was the first subdividing feature for plants, yielding the major division of herbs and trees. De Candolle (1813, pp. 73-89) called it the principle of subordination of characters, attributing it to Bernard de Jussieu, and listed the importance of botanical characters in the following order: those of the embryo, those of the stamens and pistil, those of the envelopes of the embryo, and those of accessory floral structures. However, he qualified these rules with so many exceptions that their validity is very doubtful on his own showing, and it is clear that in practice it is the correlation among various characters which is the real basis of his theories.

We believe that no such hierarchies can be made a priori, and it is well known that in practice quite different hierarchies of characters are used for different taxonomic groups. Characters which separate the species of one genus usually do not do so in the next genus; where they do appear to do so, the taxonomies are commonly arbitrary and unnatural. Many systematists have believed that certain classes of characters will separate species within genera (for example, physiological characters) while other classes (such as morphological characters)

separate genera within each family. It is easy to find exceptions in most natural classifications, and in practice some physiological and some morphological characters are found both to separate species and to separate genera. The difficulty is to decide which physiological characters and which morphological characters are those obeying the rule and which are the exceptions; in any event, many of them can be plausibly regarded as both morphological and physiological. Similar objections apply to all classes of characters and to all taxonomic ranks. It would indeed be curious if evolution, which is responsible for the natural hierarchy, should be so obliging as to operate only on certain classes of characters at specified taxonomic ranks. Again, it is clear that in practice it is correlations among characters, whatever their nature, which decide the issue, for otherwise it is impossible to explain why diametrically opposite views are held by different systematists of undoubted competence. Frequently the problem of classes of characters arises because taxonomists are aiming at two mutually incompatible purposes, as when, for example, adaptive features are used to estimate the degree of evolutionary convergence, while nonadaptive features are used to estimate phylogenetic relations; discrepancies between the schemes are then inevitable.

2.4.4. Adaptive characters

Overemphasis on adaptiveness of characters is another fault of modern systematics. Some taxonomists prefer to base their classification on what they suppose to be nonadaptive characters. However, modern genetics is showing us that few if any characters can be considered nonadaptive. The converse view—that taxa should be based on adaptive characters (Inger, 1958)—is quite impracticable, as has been shown by Sokal (1959).

2.5. NOMENCLATURE

It is not our intention to enter upon an extensive discussion and criticism of present-day practices in nomenclature. First of all, such an undertaking would require considerable space to do it justice; more important, however, we have no constructive revision to offer in connection with our proposals for numerical taxonomy. An excellent discussion of the problems of nomenclature is that of Simpson (1961,

pp. 28-34). The phenon system of nomenclature which we do propose (in Section 9.1.1) is not suggested as a substitute for the existing system of nomenclature but is designed to be used alongside it.

An ideal system of nomenclature would contain within it information serving to distinguish at least symbolically its lowest units, presumably species. It would also contain information linking a given species to other species and give some idea of its affinity with its neighbors. A third requirement of an ideal nomenclatural system would be to locate the lowest unit correctly within the hierarchy of taxa in nature; that is, the name should inform about the phylum, class, order, family, and any other necessary taxon to which the organism belongs. A fourth requisite of such a system would be that it serve as a ready and internationally accepted handle for recognizing and dealing with the species.

Systematists have from time to time attempted to make binominal nomenclature serve all these functions. It is of course easily recognized that it can do none of them at all well. Since the time of Linnaeus no attempt has been made to summarize species differences in a single word. There are by now so many genera that most generic names are quite unfamiliar to all but a few specialists, and generic names in different categories convey different ideas of affinity to persons working with them. Since no familial or other information of higher category is contained in the binomen, it cannot serve as a marker of the species' place in the system of nature; finally, instability of nomenclature has restricted much of the usefulness of the binomen as a handle or label.