TAXONOMY AND SEMANTIC CONTRAST

PAUL KAY

University of California, Berkeley

The notions 'taxonomy' and 'semantic contrast' have played an important role in ethnographic semantics, but they are not so simple as is sometimes supposed. In particular, 'level of contrast' is an ill-defined and misleading term. When one formulates the notion 'taxonomy' in a precise way, as is done here, light is shed on several questions of lexical semantics involving the notion of semantic contrast. For example, several different types of semantic contrast relation, all appearing to have empirical reality, can be formally defined.

0. INTRODUCTION. The concept of taxonomy is becoming one of increasing importance in the fields of ethnography and semantics. This paper offers a formal treatment of the subject and an application of the formalism to several problem areas in the fields of semantics, ethnography, and cognition.[1]

The expression 'taxonomy' has been used by anthropologists and linguists in a variety of ways, some of which bear little relation to the subject of this essay. Two such usages should be mentioned briefly, in order that possible confusions may be minimized. First, some people have used 'taxonomy' to refer to any system of classification and naming, regardless of its structure. Used in this way, 'taxonomy' is effectively synonymous with 'lexical domain'. Given the long tradition in biology of applying 'taxonomy' only to a particular structural type of classification system, such a use of the term is unfortunate. 'Taxonomy' and 'taxonomic structure' are defined here so as to be applicable only to a particular subclass of lexical domains whose members display certain formal properties.

Second, 'taxonomy' is sometimes employed to refer, not to a hierarchy of sets, but rather to an arrangement of properties (semantic features), presumably one felt to be consonant with a hierarchy of sets (Lounsbury 1964). For example, in the standard usage followed here, if one were to consider the English words *person*, *man*, and *woman* to be involved in taxonomic relations, the elements of the taxonomic structure would be the set of humans, the set of men, and the set of women. On the other hand, in the view suggested by Lounsbury, the basic elements of the taxonomic structure are the properties (features) HUMAN, MALE, FEMALE, and so on. It would perhaps be possible to construct a formal account of taxonomic structure based on the property or feature approach. I think, however, that there are reasons against adopting a feature approach at the outset, although this is not the place to present those arguments in detail (see Kay 1966,

Gregg 1968). In any case, the basic elements of taxonomic structure as they are defined below are sets, not properties.[2]

The intuitive notions of taxonomy and taxonomic structure formalized here are similar to those of biosystematics. Biological taxonomy, however, involves an order of structure beyond that set up here as taxonomic—namely, that of the series of absolute categories SPECIES, GENUS, FAMILY, ..., KINGDOM, and the definition of these categories as sets of taxa. Our taxonomic structure is simply a hierarchy of inclusion relations among sets. Formal treatments of biological taxonomies have been given by Gregg 1954, 1967, 1968; Beckner 1959; Sklar 1964; and van Valen 1964.

Some anthropologists (e.g. Frake 1961, 1962; Conklin 1962a,b, 1964; Berlin 1969a,b; and Berlin, Breedlove, & Raven 1966, 1968) have found taxonomic structures in the cultures of non-literate peoples, and it appears likely that important parts of the lexicons of all natural languages are organized taxonomically. The initial discoveries in the modern era that extensive and precise taxonomies exist among illiterate primitives originally occasioned surprise bordering on incredulity in some quarters. But it is increasingly recognized that the similarity to Linnaean taxonomy of the folk taxonomies discovered by ethnographers and ethnobiologists need not cause surprise, since Linnaean taxonomy is simply the particular folk taxonomy with which Western Europeans are most familiar. Linnaeus did not invent the principles of taxonomy; he simply employed, and made explicit, those which were implicit in his own culture—which, as it turns out, for the most part represent universal principles of classification and nomenclature, found in all human cultures and languages (see Berlin 1969a,b with regard to universals in taxonomic nomenclature). The present study is thus an attempt to summarize as explicitly as possible certain empirical findings of classical biosystematics and modern semantic ethnography which may now be considered to represent formal universals of human mental structure.

Section 1 introduces the formal definition of taxonomic structure and sketches the major outlines of this kind of mathematical object. The presentation here is informal. References to Appendix I, which contains an axiomatic treatment of the subject, are included in square brackets. The mathematical object called a taxonomic structure in Appendix I is quite similar to Gregg's (1954) notion of a taxonomy. The chief differences are: (a) the present formulation is simpler, (b) the present formulation contains nothing corresponding to Gregg's CATEGORIES,

---

[2] It is not clear that there is a particular formal structure definable for the properties (or features) of lexical items involved in a taxonomy. Indeed, one is inclined to wonder at the tacit acceptance by linguists of the undocumented assumption that denotative meanings can, in general, be constructed by finite logical formulae using semantic features as atomic predicates. If we do assume lexical meanings to be expressible in terms of semantic features and the machinery of logic, it does not follow that there is a particular feature structure characteristically associated with taxonomies, although such an association seems likely. The area appears worthy of empirical investigation. If the distinctions drawn below between taxonomic structure, taxonomic nomenclature, and the mapping that relates the two are of any value, then they may prove useful in the empirical investigation of the characteristic feature-structures of taxonomies.

and (c) Gregg's formulation contains nothing corresponding to the types of contrast relations defined here.[3]

Section 2 introduces some theoretical problems relating to taxonomy in ethnographic and semantic contexts, and shows how the present formulation applies to, and perhaps clarifies somewhat, these problems. In particular, the notion of semantic contrast in the context of taxonomy is examined in some detail. Section 3 introduces the notion of taxonomy, and examines the nature of the mapping which governs the realization of conceptual taxa by lexical items.

1. The purely formal entity which, under certain empirical conditions, underlies a taxonomy is called a taxonomic structure. This section is concerned primarily with taxonomic structures. Detailed discussion of what must be added to a taxonomic structure to make a taxonomy is deferred until §2. Briefly, the distinction is this: a taxonomic structure is concerned with sets (or classes, or segregates) and the relations among them; it is not concerned with the names these sets may or may not have. In keeping with standard usage, we call the sets (or classes, or segregates) involved in a taxonomic structure TAXA (singular TAXON). A taxonomy, on the other hand, always includes both a taxonomic structure and also a set of names, and a mapping involving the set of taxa and the set of names. We return to this subject in greater detail in the following section.

A taxonomic structure is a relational structure that has two components and that satisfies two axioms. The first component is a finite set T of taxa. Each taxon is itself a non-null set, i.e. a set which has some members. Hence T is a set whose members are non-null sets. Examples of taxa are: the set of all plants, the set of all trees, the set of all oaks—but not the English words *plant(s)*, *tree(s)*, *oak(s)*. (In what follows I will use 'oak' synonymously with 'the set of all oaks' and *oak* synonymously with 'the word *oak*'.) Let us call the number of taxa involved in a given taxonomic structure n; we may then enumerate the set of taxa: $T = \{t_1, t_2, \cdots, t_n\}$. This set of taxa is the first component of a taxonomic structure.

The second component is a relation, namely the relation STRICT-INCLUSION-OF-SETS restricted to the members of T. A set t_i strictly includes another set t_j just if every member of t_j is a member of t_i and there is at least one member of t_i which is not a member of t_j. The set of plants strictly includes the set of trees and also the set of oaks; the set of trees strictly includes the set of oaks. However, the set of *Quercus* does not strictly include the set of oaks, although it includes it, since the two sets have the same membership; that is, every set includes itself, but no set properly includes itself.

The phrase 'restricted to the members of T' indicates that we are concerned only with strict inclusion relations among the sets under consideration, not among all imaginable sets. If the members of T are, say, all the plant taxa, then the relation 'strict-inclusion-restricted-to-the-members-of-T' does not hold between, for

---

[3] Gregg 1967 takes a somewhat different approach in the matter of categories from his 1954 paper, treating them as basic to his definition of n-rank Linnaean structures; in 1954, categories were considered to be definable in terms of inclusion relations among taxa. Gregg's more recent approach is discussed in Appendix II to this paper. See particularly Fig. 11 for the revised treatment of categories.

example, animal and vertebrate, though it does hold between tree and oak. The second component of a taxonomic structure is thus the relation 'strict-inclusion-of-sets-restricted-to-the-members-of-T' [See Appendix, 1].

Given the relation of strict inclusion restricted to a set T, we define another relation, IMMEDIATE PRECEDENCE, in order to make it easy to express naturally the two axioms for taxonomic structures. Let t_i and t_j, each of which is a set, be distinct members of T. We say t_i immediately precedes t_j just if (a) t_i strictly includes t_j, and (b) there is no other set t_k in T such that t_i strictly includes t_k and t_k strictly includes t_j [Appendix, 4]. For example, 'tree' immediately precedes 'oak' because 'tree' strictly includes 'oak', and there is no other plant taxon which is strictly included in 'tree' and which also strictly includes 'oak.' Speaking loosely, one taxon immediately precedes another when the first is 'just above' the second in a taxonomic structure.

We symbolize the two components of a taxonomic structure as T (the set of taxa) and $\supset$ (the relation 'strict-inclusion-restricted-to-the-members-of-T'). Let us represent the ordered pair formed from these two components with the Greek letter tau, τ. That is, using angles to enclose an ordered pair, $\tau = \langle T, \supset \rangle$ [Appendix, 1].

A relational structure such as τ is a taxonomic structure just if it satisfies the following two axioms:

First, there is exactly one member of T which strictly includes every other member. This member is called the unique beginner. In a taxonomic structure of plants, 'plant' is the unique beginner; it strictly includes every other taxon, such as 'tree', 'oak,' 'grass,' 'bamboo' etc. [Appendix, 2].

The second axiom involves the notion 'partition'. A partition is a division of a set into subsets that places each member of the original set in exactly one of the subsets. The subsets are called cells of the partition. For example, suppose the prisoners in a jail are each assigned to a cell so that every prisoner is in some cell and no prisoner is in more than one cell (though different cells may contain varying numbers of prisoners). Then the cells of the jail are the cells of a partition of the set of prisoners.

Let $c(t_i)$ stand for the set of all those taxa immediately preceded by the taxon t_i. For example, if t_i is 'oak' then the members of $c(t_i)$ are 'live oak', 'jack oak', 'red oak', etc. By the second axiom, for any taxon t_i in T, if $c(t_i)$ has any members, than $c(t_i)$ is a partition of t_i. This axiom ensures, for a taxon such as 'oak' that HAS subclasses, that (a) each individual oak in the world is in some subclass, and (b) each individual oak is in just one subclass. The second axiom does, of course, allow for the possibility that a taxon has no subclasses in T. An example is 'live oak' (for me, at least) [Appendix, 3].

From this simple axiomatic definition, a number of consequences follow that are proved as theorems in Appendix I. These seem to agree rather well with standard notions regarding the formal properties of the structures which underlie taxonomies. In addition, several definitions can be made which appear to correspond to existing intuitive concepts. Let us begin with some of the latter.

(i) The set of all taxa immediately preceded by the same taxon constitutes a CONTRAST SET. That is, a contrast set is any non-null set $c(t_i)$ where t_i is a mem-

ber of T. In the previous example, all the immediate subclasses of oak—live oak, jack oak, etc.—constitute a contrast set [Appendix, 8].

(ii) A TERMINAL TAXON is one that strictly includes no other taxon; that is, a taxon t_i for which $c(t_i)$ is the empty set. For me, live oak is a terminal taxon [Appendix, 11].

(iii) The LEVEL OF A TAXON is defined as follows: The level of the unique beginner is 0. The level of a taxon immediately preceded by the unique beginner is 1. The level of a taxon immediately preceded by a taxon of level 1 is 2, and so on.[4] For example, if, in a taxonomic structure, the unique beginner, 'plant', immediately precedes 'tree', 'tree' immediately precedes 'oak', and 'oak' immediately precedes 'live oak', then 'plant' is at level 0, 'tree' at level 1, 'oak' at level 2, and 'live oak' at level 3 [Appendix, 13].

(iv) The DEPTH OF A TAXONOMIC STRUCTURE is the greatest (deepest) level attained by any taxon in that structure. In the running example, if no taxon has a level greater (deeper) than 'live oak', which is 3, then 3 is the depth of the taxonomic structure [Appendix, 18].

Some of the direct consequences of the above axioms and definitions are the following:

(v) Every contrast set is a proper subset of T and contains at least two members [Appendix, 9, 10].

(vi) Each taxon other than the unique beginner has exactly one immediate predecessor [Appendix, 12].

The following result is the fundamental theorem, as it were, of the theory of taxonomic structure:

(vii) If two distinct taxa have any members in common, then one of them strictly includes the other. That is, any two distinct taxa are either mutually exclusive or in the relation of strict-inclusion. The possible relations between two distinct taxa t_i and t_j in a taxonomic structure are pictured in the Venn diagrams of Figure 1a; the disallowed relation is shown in Figure 1b. Note that fact (vii) is not taken as axiomatic, but rather is a consequence of the axioms [Appendix, 16]. It corresponds to the intuitive notion that there is no partial overlap between taxa; either one taxon is totally included in the other, or they have no member in common. This result (vii) also expresses formally the intuition that taxonomic structures are strictly hierarchic and contain no element of cross-classification. This is the major formal basis for representing taxonomic structures in the standard tree or box diagrams.

(viii) If a taxonomic structure is of depth n, then it contains at least one taxon at each level from zero to n inclusive. For example, if the depth is five, then there is at least one taxon at each of the levels 0, 1, 2, 3, 4, and 5. In this sense, a taxonomic structure has no gaps [Appendix, 19].

(ix) The terminal taxa constitute a partition of the unique beginner. E.g., in a taxonomic structure of plants, each individual plant (i.e. each member of the set 'plant') belongs to exactly one terminal taxon (e.g. 'live oak', 'pitch pine' etc.) [Appendix, 20].

(x) Each taxon other than the unique beginner belongs to exactly one contrast

---

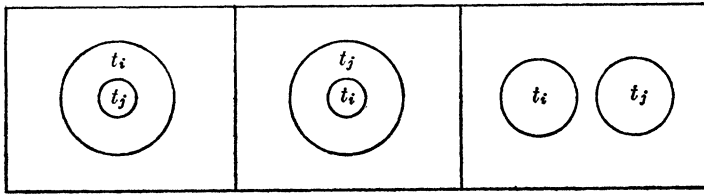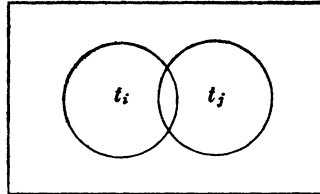[4] In general, the level of a taxon immediately preceded by a taxon of level n is $n+1$.

FIGURE 1a. Possible relations between distinct taxa t_i, t_j.

FIGURE 1b. Disallowed relation between distinct taxa t_i, t_j.
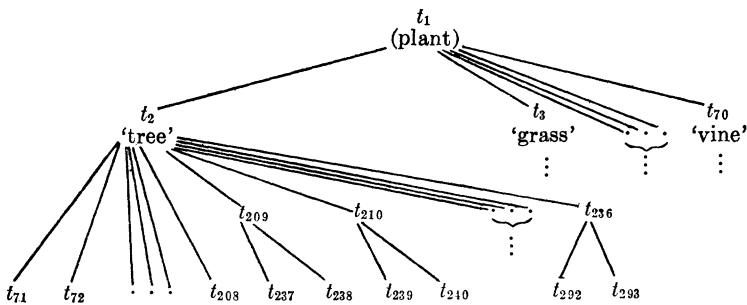
FIGURE 2. Taxonomic structure underlying Tenejapa Tzeltal plant taxonomy (one informant); adapted from Berlin, Breedlove, & Raven 1968. Single quotes indicate glosses of native lexemes. A sequence of three dots indicates omitted detail. The unique beginner 'plant' is not lexically realized in Tzeltal. Berlin 1969a,b notes that this circumstance is not at all exceptional but rather is characteristic of all but the most advanced developmental stages of natural biotaxonomies.

set. For example, 'live oak' belongs to the contrast set c(oak), 'pitch pine' to the contrast set c(pine) [Appendix, 27].

The various concepts introduced here are illustrated in Figures 2 and 3. The following observations, relating to the preceding numbered statements (i–x), may be made concerning Figure 2:

(1) The unique beginner t_1 is not in any contrast set, and is the only taxon at level zero.

(2) For any taxon t_i:

(i)   For $2 \leq i \leq 70$, t_i is in the contrast set $c(t_1)$ and is at level one.

(ii)   For $71 \leq i \leq 236$, t_i is in the contrast set $c(t_2)$.

(iii)   For $237 \leq i \leq 238$, t_i is in the contrast set $c(t_{209})$.

(iv)   For $239 \leq i \leq 240$, t_i is in the contrast set $c(t_{210})$.

(v)   For $292 \leq i \leq 293$, t_i is in the contrast set $c(t_{236})$.

FIGURE 3. Box diagram corresponding to Figure 2. Note that hatched areas indicate omitted detail. Comparison with Figure 2 reveals that box diagrams are somewhat less precise than tree diagrams in indicating omitted detail.
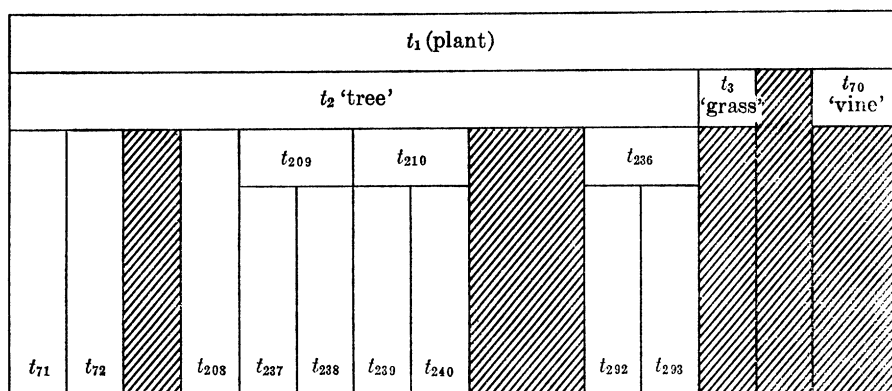
(vi)   For $71 \leq i \leq 208$ or $237 \leq i \leq 240$ or $292 \leq i \leq 293$, t_i is a terminal taxon. (See note after viii.)

(vii)   For $71 \leq i \leq 236$, t_i is at level two.

(viii)   For $237 \leq i \leq 293$, t_i is at level three. This statement, although true, is not directly inferable from Figure 2. Taxa t_{241}, ..., t_{291}, which are not pictured, are in fact all at level three, as each is immediately preceded by a taxon from the set of unpictured, level-two taxa $\{t_{211}, ..., t_{235}\}$.

(3) Not all terminal taxa are at the same level.

(4) Two distinct taxa at the same level may or may not be in the same contrast set.

(5) The depth of the taxonomic structure is not indicated in the figure. In fact the greatest depth in Tzeltal plant taxonomic structure is obtained within 'beans', an unpictured member of $c(t_{70})$.

Figure 3 illustrates the same structure in the type of box diagram often used by anthropologists.

2. The notion of taxonomy has been of use to ethnographers and cognitive theorists principally as a means of organizing relations of meaning among items in natural languages and cognitive systems. The organization of meaning relations necessarily entails the notion of (semantic) contrast, and 'contrast' has figured prominently in much recent work in folk taxonomy.

'Contrast' is obviously a relational term, but it has not always been clear in the literature what sort of entities are involved in relations of contrast, and what the various types of (semantic) contrast relations are. In particular, despite Conklin's admonition (1962b:121), the distinction between classification and nomenclature is sometimes ignored; consequently, it is not clear whether 'contrast' is a relation obtaining between taxa (concepts) or names of taxa (lexemes). Also the term contrast has been combined with 'level' in the phrase 'level of con-

trast', whose precise signification is far from clear (cf. Kay 1966:21). We return to these issues in greater detail below. Before doing so, however, it is necessary to specify what, in general, we mean by semantic contrast.

Probably we should take the same initial attitude to 'contrast' in semantics as is taken in linguistics generally, namely that one does not assume items to be in contrast with one another unless one finds positive evidence for it. The usual form this evidence takes is that of a frame in which (1) the informant allows substitution of either item, and (2) the informant judges the utterances resulting from the alternative substitutions to be different utterances. If the utterance is restricted in length to a single word, this is the method of minimal pairs. Such tests are usually held to be too strong, in the sense that passing them is a sufficient but not a necessary condition for saying that two items contrast. (The problem of deciding which of the informant's responses to take as indicating that two utterances are different is an important but not an insuperable one.) For semantics we might adopt the following criterion of 'same/different' as a test comparable to that of minimal pairs—that is, one that provides a sufficient, but not a necessary, operational diagnostic for semantic contrast:

    (1) Two lexical items can be said to contrast semantically if
        (a) there exists an assertion frame in which the informant allows substitution of either item, in the sense that with either alternative substitution he can easily judge the resulting assertion as true or false; and
        (b) the informant's truth judgments for the assertions resulting from the two substitutions are different, i.e. one true and one false.

Although we wish theoretically to define relations of contrast between units of content—in this case taxa—our operational criterion must deal with the overt expressions of these units, the lexemes through which the taxa are realized in actual speech. Test 1 applies in practice to taxa then, to the extent that we can establish empirically the particular taxon which is realized by each test lexeme (lexical item). I return to this problem below.

Speaking somewhat loosely, we may say that Test 1 boils down to this: two categories contrast semantically if an assertion using one elicits assent from the informant, while the otherwise identical assertion employing the other elicits dissent from the same informant. As we noted, this is probably too strong a test. For example the lexical items *morning star* and *evening star* might never be shown to contrast on Test 1 for some informants; yet we might still wish to think of them as contrasting semantically. It suffices to establish here that if two items pass a test like 1, then we have to say that they contrast semantically.

Note that above we have spoken about assertions, not sentences or even declarative sentences. Consider:

    (2) (a) That's not a MAILMAN, Johnny, it's a POSTMAN.
        (b) I am not an EYE-DOCTOR, Sir, I am an OPHTHALMOLOGIST.

The emphasized words in each case may appear to be in semantic contrast in the sense of Test 1. If, however, one considers the assertions being made rather than the sentences that are their realizations, one sees that these assertions do not concern mail carriers and medical practitioners, but rather express preferences about

the words *mailman, postman, eye-doctor,* and *ophthalmologist.* These are meta-
linguistic assertions in which the words *mailman, eye-doctor* etc. occur not in use
but in reference. In fact, most native speakers of English will probably accept
the following as a paraphrase of 2a:

> (3) Johnny, 'mailman' and 'postman' mean the same thing, but I prefer
>   that you USE 'postman'.

If one accepts the kind of test for contrast given in Test 1, then clearly, in
statements such as 'The categories plant and tree do not contrast, since all trees
are plants', 'contrast' is being used in a peculiar sense. In order to retain con-
tinuity between semantics and general linguistics, we prefer to keep 'contrast'
sufficiently general that at least any pair of items that pass Test 1 will be said to
contrast. Evidently, 'tree' and 'plant' will pass many versions of Test 1, for ex-
ample, all those of the form

> (4) All X's have Y

where X is a variable whose values are the taxa 'tree' or 'plant', and Y is any
characteristic (or set of characteristics) that distinguish trees from all other
plants.

In short, all pairs of items in a taxonomic structure are in contrast, since asser-
tion frames can always be created for use in Test 1 that focus precisely upon the
characteristics that distinguish one taxon from the other. Any two distinct taxa
must, of course, have different characteristics; otherwise they would have the
same membership.

Let us return to the practical problems engendered by a theoretical position
that defines contrast as a relation between taxa rather than a relation between
lexemes. Test 1, narrowly interpreted, cannot be applied to taxa directly, but only
to the lexemes that are their names. Nevertheless, the relations of contrast ob-
taining between taxa may be empirically determined, not only in cases where
taxa are lexically realized as polysemous lexemes, but even in those cases which
might be supposed a-priori to present insuperable obstacles—namely, those in
which there exist several closely related taxa that are not named at all, but which
nevertheless participate mutually in several different kinds of contrast relations.
Berlin, Breedlove, & Raven 1968 have recently demonstrated that such covert
categories may be unambiguously identified by several independent empirical
procedures. These empirical successes encourage us in the broad view of Test 1.
Accordingly, we take Test 1 to be literally applicable to taxa in principle, and
practically applicable to taxa to the degree that we are able to extend our range of
empirical methods for assessing informants' judgments of the truth or falsity of
an assertion.

It is theoretically crucial that contrast relations be originally defined upon
taxa rather than upon the lexemes that realize them. In effect, this essay consti-
tutes an argument in support of that assertion. It is based on the general assump-
tion that theoretically significant contrasts in linguistic content are psychologi-
cally independent of significant contrasts in linguistic expression, but not con-
versely. To put it in plain, if perhaps imprecise, terms: We speak in order to
communicate thought; we do not think in order to provide content for our
speech. Perhaps this belief is not totally dissimilar to that which animates a

number of contemporary approaches to grammar which, although quite diverse, share a common concern with meaning (e.g. those of Fillmore, Lakoff, McCawley, Ross, Chafe, and many others).

We consider now the frequently used term 'level of contrast' and various derivative expressions. Two kinds of confusion may result from the use of these terms. The first is a confusion between nomenclature (lexemes) and classification (taxa); the second is a confusion between the relations 'having the same level' and 'being in the same contrast set'.

Figures 4–5 illustrate the way expressions such as 'level of contrast' may lead to self-contradictory or nonsensical statements by confusing taxa with the lexemes that are their names. The two figures diagram the same set of semantic relationships. In speaking of situations such as that pictured here, it is sometimes said that 'MAN contrasts with ANIMAL at one level and with WOMAN at another level.' It is quite unclear, however, whether the expressions MAN, ANIMAL, and WOMAN are supposed to refer to taxa, or to the lexemes that realize taxa—or, in fact, what the statement means at all. Let us examine the facts of which note is implicitly being taken. These are (1) that the lexeme *man* is polysemous, being the realization both of t_2 and t_4 ; (2) that t_2 is in the same contrast set as t_3 (which is realized by the lexeme *animal*), while t_4 is in the same contrast set as t_5 (which is realized as *woman*); (3) that the contrast sets $\{t_2 , t_3\}$ and $\{t_4 , t_5\}$ are at different levels. Less technically, the lexeme *man* is polysemous, being the realization of two distinct taxa which are members of distinct contrast sets that also happen to be at different levels. The key fact being noted is the polysemy of *man*. (The polysemy of *animal* is incidental to the point.)

Now compare Figures 6–7. In this kind of situation it is also in keeping with common usage to say, 'WILD PEPPER contrasts with HOUSEYARD PEPPER at one level and with HOUSEYARD CHILI PEPPER at another level.' But again, the intended
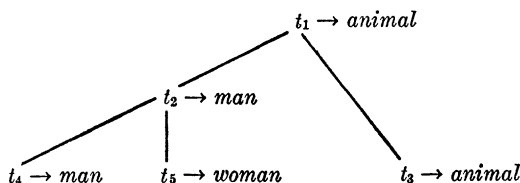

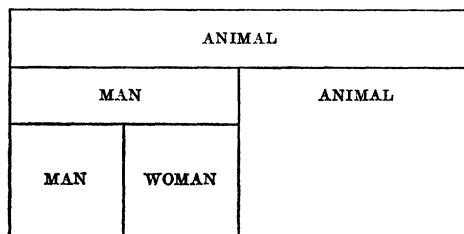FIGURE 4. Adapted from Frake (1961:117). Arrows connect taxa to the lexemes that are their realizations.

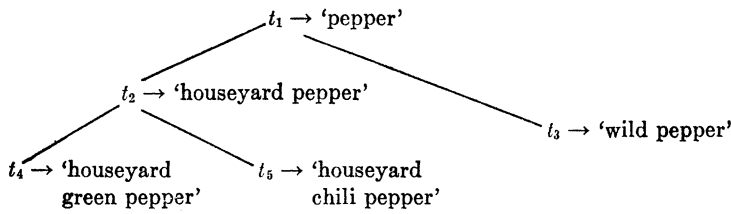FIGURE 5. Box diagram of the set of relationships shown in Figure 4.

FIGURE 6. Adapted from Conklin (1962b:131–3)

FIGURE 7. Box diagram of the set of relationships shown in Figure 6.

reference of the capitalized forms is not at all clear. Note particularly that, whereas the key factor in the previous example was the polysemy of the lexeme *man*, there is no polysemy in this case.

It is true that, in the box diagram of Figure 7, the box containing WILD PEPPER is visually 'on a level', so to speak, both with the box containing HOUSEYARD PEPPER and the one containing HOUSEYARD CHILI PEPPER. However, this is merely an observation on the visual properties of box diagrams, and reflects nothing about the taxonomic structure being pictured. Perhaps this accident of visual imagery has led people to imagine that there is some meaningful sense of taxonomic 'level' according to which a taxon (for example t_3—WILD PEPPER—in Figs. 6–7) may be said to have two distinct levels. If so, I have found no indication of what that sense might be. In any case, whatever the expression 'X contrasts with Y at one level and with Z at another level' means with respect to Figs 4–5, it cannot mean the same thing as it does with respect to Figs. 6–7; in the former case the reference is to polysemy, and in the latter there is no polysemy. In Figs. 4–5, 'the same linguistic form [*man*] designates segregates [the distinct taxa t_2 and t_4] at different levels ...'; while in Figs. 6–7 'a single unpartitioned segregate [the taxon t_3] contrasts with two or more other segregates [e.g., taxa t_2 and t_4] which are themselves at different levels ...' (Frake 1962:82).

The second confusion arises from the very expression 'level of contrast', because the existence of such a term has suggested to some that any two taxa at the same level are in the same kind of contrast relation. The literature abounds with uses of the expression 'level of contrast', which implies that all taxa at the same level are in the same contrast set. However, this is obviously not the case,[5] and it is unlikely that anyone would have been led to this confusion were it not for the currency of the unfortunate expressions 'level of contrast' and 'contrast at the

---

[5] Cf. Fig. 2 and the various numbered statements at the end of §1.

same (different) levels'. Knowing the levels at which two taxa occur gives us very little information about the kind of contrast relations that obtain between them. The level of a taxon merely says how many taxa occur between it and the unique beginner in the chain of immediate precedence that connects them. In particular, two distinct taxa that have the same level may or may not be in the same contrast set [Appendix, Remark following Theorem 17]. (Level 1 is unique in that all the taxa at this level do constitute a single contrast set.) The expression 'level of contrast', as it has been used in the anthropological literature, is at best ambiguous. In the context of the present formulation, it is meaningless.

In the hope of avoiding these and similar kinds of terminological confusions regarding the contrast relations between taxa, we introduce the following definitions of kinds of taxonomic contrast. As we have discussed, two taxa, one of which strictly includes the other (such as 'tree' and 'oak'), do contrast semantically. We name this kind of contrast relation INCLUSION CONTRAST [Appendix, 23].

A special contrast relation obtains between any two taxa which belong to the same contrast set. We call this relation DIRECT CONTRAST. That is, two taxa are said to contrast directly just in case they are in the same contrast set [Appendix, 21]. Direct contrast is perhaps what users of the expression 'contrast at the same level' have most often had in mind.

Two taxa which are in neither 'direct contrast' nor 'inclusion contrast' are said to be in INDIRECT CONTRAST (or to contrast indirectly) via the two taxa which include them and which are themselves in direct contrast. For example, in Figure 8, t_4 and t_6 contrast indirectly via t_2 and t_3 [Appendix, 23]. (Note that in defining indirect contrast, we speak of inclusion—not strict inclusion. Thus, for example, in Figure 8, t_4 and t_3 contrast indirectly via t_2 and t_3 .)

We define a special contrast relation which obtains among the terminal taxa: any two of these are said to be in TERMINAL CONTRAST. In Figure 8, t_4 and t_9 are in terminal contrast, as are also the pairs (t_4 , t_5) and (t_4 , t_6). In general, any two terminal taxa are in terminal contrast regardless of whether they are in the same contrast set or at the same level. In Figure 8, t_9 is at level one, while each of the other terminal taxa $t_4 , ..., t_8$ is at level two; nevertheless t_9 is in terminal contrast with each taxon $t_4 , ..., t_8$.

The basis of the intuition 'terminal contrast' is the fact that the terminal taxa collectively furnish the finest possible partition of the unique beginner [Appendix, 20]. That is, the terminal taxa collectively provide the finest available set of mutually exclusive and jointly exhaustive taxonomic categories for classifying an individual. Thus every terminal taxon has a special contrast relation with every other, regardless of its level or contrast set affiliation. To describe individuals in terms of terminal taxa is to slice the taxonomic pie as fine as possible.
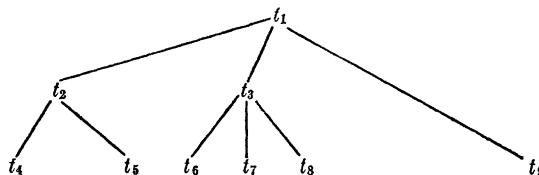

FIGURE 8.

The relation defined here as terminal contrast may be related to certain usages of the vague expression 'lowest level of contrast'. It is sometimes said, for example, that certain analyses are appropriately performed only on taxa that 'occur at the lowest level of contrast'. As we have seen, the 'lowest level of contrast' is meaningless, since the expression 'level of contrast' is itself undefined. Specifically, all terminal taxa are not necessarily at the same level. It cannot be emphasized too strongly that the levels at which two taxa occur bear only indirect relevance to the relation(s) of contrast obtaining between them, and in no way are sufficient to determine those relations.

Berlin has recently made a series of remarkable discoveries concerning a subset of taxa which he terms generic (1969 a,b; Berlin, Breedlove, & Laughlin 1969). Generic taxa appear to occupy a privileged status in all natural taxonomies. Substantively, in biotaxonomies they correspond frequently, although not invariably, to Western biological genera. (Although Berlin's researches have so far been restricted to biological taxonomies, I would speculate that generic categorization may turn out to be a fundamental process of human thought.) From a formal point of view, the generic taxa constitute a partition of the unique beginner. They may occur at any non-zero level and may or may not be terminal. As we have already noted, any individual in the taxonomic domain—that is, any member of the unique beginner—is a member of exactly one terminal taxon. Each terminal taxon is in turn linked to the unique beginner by a chain of taxa connected by immediate precedence. Hence, any object that may be characterized by membership in a terminal taxon may, alternatively, be conceptualized by the user of the taxonomic structure in terms of any of the taxa in the appropriate immediate precedence chain. For example, Lassie is a collie, and also a dog, a mammal, a vertebrate, ..., and ultimately an animal. Since the generic taxa partition the domain of relevant individuals, each chain of immediate precedence of this type contains exactly one generic taxon. One of Berlin's important substantive hypotheses deriving from the concept of generic taxon is that one taxon from each such chain is the most salient, the most frequently employed by actual persons in actual classifying events; and that this is the generic taxon. Roughly then, the generic taxa are the ones that partition the domain of individuals in the way that corresponds to the most obvious discontinuities in nature, furnishing a subset of taxonomic categories which are the most obvious, natural, and frequently employed. For example, 'dog' is a generic category in folk English, and a particular dog, say Lassie, is probably more often thought of as a dog than as a collie, or as a mammal, or as a vertebrate. (Certainly she is usually referred to by the lexeme *dog*.)

The above discussion gives only a suggestion of the substantive nature of Berlin's concept of generic taxon; the reader is referred to the works cited for full explanation of generic taxa, their relation to taxonomic nomenclature, and their role in the synchronic use and diachronic development of taxonomies. My purpose here is only to indicate the existence of empirical motivation for an additional type of contrast relation based on the psychological importance of the generic partition. GENERIC CONTRAST is defined simply as that special relation of contrast which holds between any two generic taxa [Appendix, 28, 29]. Thus, as terminal

contrast is based upon the finest partition of the domain of individuals available to the user, generic contrast is based upon the most natural and psychologically salient partition.

We have defined five special types of contrast relations. Three of these—inclusion, direct contrast, and indirect contrast—form a natural group in that any pair of taxa in a taxonomic structure are related to each other in exactly one of these ways [Appendix, 24]. The remaining two types of contrast relations—generic and terminal contrast—are also a natural set. Both are based on membership in special sets of taxa that constitute partitions of the domain of individuals, i.e. the terminal and generic partitions. Since the generic taxa may or may not be in terminal contrast (and conversely), two taxa in generic contrast may or may not also be in terminal contrast (and conversely) [Appendix, 30]. Two taxa that are in either generic or terminal contrast (or both) are never in inclusion contrast, and therefore they must contrast either directly or indirectly [Appendix, 24, 26, 30].

3. As mentioned in the introduction, the intuitive notion of taxonomy involves a taxonomic structure, a set of names (lexemes), and a mapping that relates the two. The image of a taxon under this mapping is its lexical realization. However, we do not feel that just any mapping of taxa onto lexemes will model in a revealing way the class of empirical objects we wish to call taxonomies. The problem is that just any mapping does not provide enough structure to capture our intuition of what empirical taxonomies are like. Figures 9–10 provide some hypothetical examples of mappings of taxa onto sets of lexemes that most people will, I am sure, not wish to call taxonomies.[6]

In Figure 9 a single name (the lexeme x_1) is the image of each taxon in the taxonomic structure; hence, none of the structural relations among the taxa are preserved under the mapping. A parallel case would occur in English if it were to contain a single word, say *fern*, to refer to every taxon in the plant world. *Fern* would thus have the various significations 'plant', 'tree', 'grass', 'oak', 'lichen', 'American Beauty rose' etc., for each plant category in our culture. It is preferable that our formal definition of taxonomy rule out such absurd cases.

Figure 10 pictures a different but equally unsatisfactory situation. Here 99 lexemes are employed to denote eight conceptual categories, and each of the names can refer to any of the categories. Each of the 99 lexemes is thus eight ways polysemous; further, each lexeme associated with a given taxon is in a relation of synonymy with every other lexeme associated with that taxon. Again the underlying taxonomic structure is totally obscured.

Absurdities such as these can be concocted at will so long as we introduce no constraints on the mapping of taxa onto lexemes. Clearly some constraints must be introduced if the formalism is to be narrow enough to reflect accurately the amount and kind of structure we feel intuitively to exist in real taxonomies. The

---

[6] I am here using 'mapping' in a loose sense for the sort of thing mathematicians sometimes call a correspondence or relation, when they wish to define 'mapping' as that special type of correspondence in which every member of the domain has an image. I believe the present usage is more in accord with the ordinary sense of 'mapping' as the word is used in non-technical discussions.
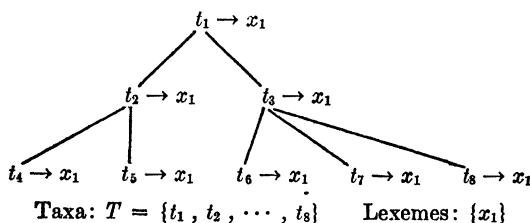
Taxa: $T = \{t_1, t_2, \cdots, t_8\}$ Lexemes: $\{x_1\}$

FIGURE 9.

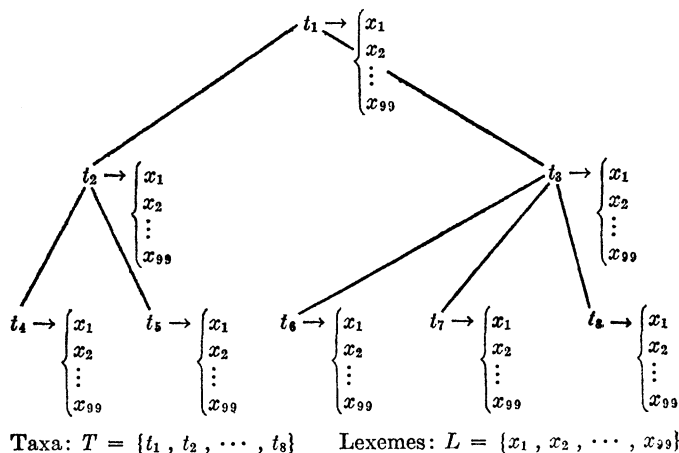Taxa: $T = \{t_1, t_2, \cdots, t_8\}$ Lexemes: $L = \{x_1, x_2, \cdots, x_{99}\}$

FIGURE 10.

problem here is to constrain the formal definition of the lexical realization mapping in just the right way—that is, in the way that results in a formal definition of 'taxonomy' which fits just those empirically observed structures that we wish to call taxonomies. In particular we would like our formal structure to be just general enough to admit as examples the natural taxonomies that have been described in detail, such as those referred to in the introduction.

The strongest and mathematically 'simplest' constraint requires that the mapping be 'one-to-one onto'. That is, each taxon is realized by a unique and distinct lexeme. (The 'onto' provision simply means that each lexeme realizes some taxon in the relevant taxonomic structure.) All possibilities of synonymy (more than one lexeme per taxon), multiple meaning (more than one taxon per lexeme), and unnamed categories (lexically unmarked taxa) are excluded.

It is very easy to find empirical counter-examples to the no-synonymy condition. For example, in one dialect of Southern Louisiana the forms *bass, black bass, trout,* and *green trout* may all be applied to the same taxon, and a similar relation holds within other such sets of lexemes as {*sac-à-lait, white perch*} and {*perch, bream, sunfish*}. We cannot accept the no-synonymy constraint.

We may relax the no-synonymy condition by defining, as the counter-domain of our mapping, a collection of SETS of synonymous lexemes. This model allows for synonymy, but still prohibits all multiple meaning. That is, it prohibits a given lexeme from being the realization of more than one taxon.

However, one of the principal achievements of empirical work on taxonomies in the last decade has been the discovery that a given lexeme may in fact realize more than one taxon. As Frake points out (1961:119), the English lexeme *man* is used as a name for a variety of taxa which might be glossed 'human', 'male human', 'adult male human', and 'virile adult male human'. Characteristically, this sort of multiple meaning involves a proper subset of T whose members can be arranged in a sequence of immediate precedence. In other words, multiple meaning in natural taxonomies appears to be restricted to a special subtype of polysemy with the following property: if a lexeme is polysemous in a taxonomy, its various senses always correspond to taxa which can be arranged in a sequence of immediate precedence.

A final condition is needed to account for lexically unmarked taxa or covert categories (Berlin, Breedlove, & Raven): the domain of the mapping is defined to be a subset T' of the set T of taxa.[7]

To review, taxonomy is defined in terms of a mapping involving a taxonomic structure and a set of lexemes. Synonymy is accounted for by partitioning the set of lexemes into cells, each of whose member lexemes realize the same taxa; the cells of this partition, rather than individual lexemes, form the converse domain of the mapping. Covert categories are accommodated by taking as the domain of the mapping, not the set T of taxa, but a subset T' of T. Finally, polysemy is incorporated in the model just to the extent that, if a single lexeme is the image of more than one taxon, then all the taxa of which it is the image can be arranged in a chain of immediate precedence [Appendix, 31, 32].

In summary, a taxonomy presupposes a taxonomic structure τ (defined on a finite set of taxa T) and a finite set of lexemes L. The taxonomy is defined by a mapping m of a subset T' of T onto a partition P of L such that for each $p \, \epsilon \, P$, (a) all the members of p are synonymous lexemes, and (b) either p is the image of a unique taxon $t_i \, \epsilon \, T'$, or the several taxa of which p is the image can be arranged in a chain of immediate precedence [Appendix, 31, 32].

APPENDIX I: AXIOMATIC TREATMENT

(1) Def. A binary relational structure $\tau = \langle T, \supset \rangle$, where T is a finite non-null set of non-null sets and $\supset$ is the relation of strict inclusion restricted to the members of T, is a TAXONOMIC STRUCTURE iff ('if and only if') the following two axioms are satisfied.

(2) Axiom. There is a unique $t_1 \, \epsilon \, T$, called the UNIQUE BEGINNER, such that, for any $t_i \, \epsilon \, T \, (t_i \neq t_1)$, $t_1 \supset t_i$.

(3) Axiom. For any $t_i \, \epsilon \, T$, the set $c(t_i) = \{t_j \mid t_j \, \epsilon \, T, \, t_i \supset t_j$ and there is no $t_k \, \epsilon \, T$ such that $t_i \supset t_k$ and $t_k \supset t_j\}$ is either null or is a partition of t_i.

(4) Def. IMMEDIATE PRECEDENCE $= \{\langle t_i, t_j \rangle \mid t_i, t_j \, \epsilon \, T, \, t_i \supset t_j$ and there is no $t_k \, \epsilon \, T$ such that $t_i \supset t_k$ and $t_k \supset t_j\}$.

If $\langle t_i, t_j \rangle$ is a member of Immediate Precedence, we write equivalently '$t_i \rightarrow t_j$', 't_i immediately precedes t_j', 't_i is the immediate predecessor of t_j', 't_j immediately succeeds t_i', etc. If $\langle t_i, t_j \rangle \not\epsilon$ Immediate Precedence, we write $t_i \nrightarrow t_j$. Remark: Thus in Axiom 3, $c(t_i) = \{t_j \mid t_i - t_j\}$.

---

[7] By defining T' as a subset rather than as a proper subset, we leave open the theoretical possibility that there may be some natural taxonomies without covert categories. However, it appears that natural taxonomies ordinarily contain covert categories. Berlin 1969b discusses the usual absence of a lexical realization for the unique beginner in the biotaxonomies of non-literate peoples. We may ultimately discover additional formal properties of taxa that are lexically unmarked.

The following properties of Immediate Precedence (5 a–d) follow directly from Definition 4, the well-known properties of proper inclusion of sets, and the fact that T is finite (Def. 1).

(5) For any distinct t_i , t_j , $t_k \in T$

    (a) $t_i \nrightarrow t_i$;

    (b) If $t_i \rightarrow t_j$, then $t_j \nrightarrow t_i$;

    (c) If $t_i \rightarrow t_k$ and $t_k \rightarrow t_j$, then $t_i \nrightarrow t_k$;

    (d) If $t_i \supset t_j$ and $t_i \nrightarrow t_j$, then there is a $t_k \in T$ such that $t_i \rightarrow t_k$ and $t_k \supset t_j$.

(6) Theorem. For any t_i , $t_j \in T$, if $t_i \supset t_j$ and $t_i \nrightarrow t_j$, then there is a UNIQUE $t_k \in T$ such that $t_i \rightarrow t_k$ and $t_k \supset t_j$. Proof: The existence of t_k is guaranteed by 5d. Hence, by Axiom 3, $c(t_i)$ is non-null. The uniqueness of t_k now follows from Axiom 3, which ensures that (i) $t_k \in c(t_i)$ and (ii) $c(t_i)$ is a partition of t_i . The proof is complete.

Note on notation: For convenience we abbreviate '$t_a \rightarrow t_b$, $t_b \rightarrow t_c$, $t_c \rightarrow t_d$, ...' as $t_a \rightarrow t_b \rightarrow t_c \rightarrow t_d$, ...' One should bear in mind that '$t_a \rightarrow t_b \rightarrow t_c$' does NOT mean $t_a \rightarrow t_c$; on the contrary, it implies $t_a \nrightarrow t_c$ (see 5c). Also $(a, b) = (b, a); \langle a, b \rangle \neq \langle b, a \rangle;$ 'Φ' denotes the null set.

(7) Theorem. For any t_i , $t_j \in T$, if $t_i \supset t_j$ and $t_i \nrightarrow t_j$ there exists a unique finite sequence of distinct elements of T of the form $t_i \rightarrow t_k \rightarrow \cdots \rightarrow t_j$. Proof: From Theore 6m, there is a unique $t_k \in T$, $t_i \rightarrow t_k$ and $t_k \supset t_j$. By hypothesis $t_i \nrightarrow t_j$; hence, $t_k \neq t_j$. Either $t_k \rightarrow t_j$ or $t_k \nrightarrow t_j$. If $t_k \rightarrow t_j$, the desired sequence is established, namely $t_i \rightarrow t_k \rightarrow t_j$. If $t_k \nrightarrow t_j$, by applying Theorem 6 again, we establish a unique $t_m \in T$ such that $t_k \rightarrow t_m$ and $t_m \supset t_j$. If $t_m \rightarrow t_j$, we have the desired sequence, namely $t_i \rightarrow t_k \rightarrow t_m \rightarrow t_j$. If $t_m \nrightarrow t_j$, the next term of the sequence is uniquely established by applying Theorem 6 to t_m and t_j . This procedure is continued until the term $t_n \in T$ $(t_n \rightarrow t_j)$ is found and the sequence is complete. Since no element t_r of T can occur more than once in the sequence $(t_r \rightarrow \cdots \rightarrow t_r$ implies $t_r \supset t_r$, which is absurd), and since the sequence is finite (because T is itself finite, cf. Def. 1), we are assured of finding the required $t_n \in T$, $t_i \supset t_n$, $t_n \rightarrow t_j$. This completes the proof.

(8) Def. A subset s_i of T is a CONTRAST SET iff there is a $t_i \in T$ such that $c(t_i) \neq \Phi$ and $s_i = c(t_i)$.

(9) Theorem. Every contrast set is a proper subset of T. Proof: For any $t_i \in T$, $t_i \notin c(t_i)$ (Axiom 3). This completes the proof.

(10) Theorem. Every contrast set has at least two members. Proof: Assume the contrary: there are t_i , $t_j \in T$ such that t_j is the unique member of $c(t_i)$. Since $c(t_i)$ is a partition of t_i (Axiom 3), $t_j = t_i$. However, also by Axiom 3, $t_i \rightarrow t_j$. Hence $t_i \neq t_j$ (5a). Thus a contradiction is established and the proof is complete.

(11) Def. For any $t_i \in T$, t_i is TERMINAL iff $c(t_i) = \Phi$.

(12) Theorem. For any $t_j \in T$ $(t_j \neq t_1)$, there is a unique $t_i \in T$, $t_i \rightarrow t_j$. Proof: By Axiom 2, $t_1 \supset t_j$. Thus, this theorem follows directly from the special case of Theorem 7, which specifies the unique sequence $t_1 \rightarrow \cdots \rightarrow t_j$.

(13) Def. For any $t_j \in T$,

    (a) t_j HAS LEVEL ZERO $[L(t_j) = 0]$ iff $t_j = t_1$;

    (b) t_j HAS LEVEL $n+1$ $[L(t_j) = n+1]$ iff $L(t_i) = n$ and $t_i \rightarrow t_j$ $(t_i \in T, n$ is a non-negative integer).

(14) Remark. The preceding definition is justified by the following fact. For any $t_j \in T$, there is a unique non-negative integer n such that $L(t_j) = n$. Proof: If $t_j = t_1$, $L(t_j) = 0$ according to 13a. If $t_j \neq t_1$, the uniqueness of $n = L(t_j)$ follows from Theorem 12.

(15) Theorem. For distinct t_i , $t_j \in T$; $t_i \supset t_j$ iff t_i occurs in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$. Proof: Suppose $t_i \supset t_j$. Then by Def. 4 and Theorem 7, there is a unique sequence $t_i \rightarrow \cdots \rightarrow t_j$. If $t_i = t_1$, then clearly t_i occurs in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$. If $t_i \neq t_1$ and if t_i were not in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$, then the sequence $t_1 \rightarrow \cdots \rightarrow t_j$ would not be unique, since there would be distinct sequences $t_1 \rightarrow \cdots \rightarrow t_i \rightarrow \cdots \rightarrow t_j$ and $t_1 \rightarrow \cdots \rightarrow t_j$. Hence if $t_i \supset t_j$, t_i must occur in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$, and the first half of the proof is complete. If t_i is in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$, and $t_i \neq t_j$, then either $t_i \rightarrow t_j$ or there is a unique $t_k \in T$, $t_i \supset t_k$ and $t_k \rightarrow t_j$ (Theorem 6). Since $t_k \rightarrow t_j$ implies $t_k \supset t_j$ (Def. 4) and since '$\supset$' is transitive, in either case $t_i \supset t_j$. The proof is now complete.

(16) Theorem. For distinct t_i , $t_j \, \epsilon \, T$, either $t_i \supset t_j$ or $t_j \supset t_i$ or t_i and t_j are disjoint. Proof: If either t_i or t_j is the unique beginner, then it properly includes the other (Axiom 2). Otherwise there are unique sequences $t_1 \rightarrow \cdots \rightarrow t_i$ and $t_1 \rightarrow \cdots \rightarrow t_j$. (a) $t_i \supset t_j$ iff t_i occurs in the sequence $t_1 \rightarrow \cdots \rightarrow t_j$ (Theorem 15). Similarly, (b) $t_j \supset t_i$ iff t_j occurs in the sequence $t_1 \rightarrow \cdots \rightarrow t_i$. It only remains to be shown that if neither condition (a) nor condition (b) holds, then t_i and t_j are disjoint. Consider t_p , $t_q \, \epsilon \, T$, $t_p \supseteq t_i$, $t_q \supseteq t_j$, and $t_1 \rightarrow t_p$, $t_1 \rightarrow t_q$. If $t_p \neq t_q$, then t_p and t_q are disjoint, since t_p , $t_q \, \epsilon \, c(t_1)$ (Axiom 3). If t_p and t_q are disjoint, then t_i and t_j are disjoint. If $t_p = t_q$, then consider t_r , $t_s \, \epsilon \, T$, $t_p = t_q \rightarrow t_r$, $t_p = t_q \rightarrow t_s$, $t_r \supseteq t_i$, $t_s \supseteq t_j$. Again if $t_r \neq t_s$, then t_r and t_s are disjoint because t_r , $t_s \, \epsilon \, c(t_p) = c(t_q)$, and consequently t_i and t_j are disjoint. This procedure is continued until t_m , $t_n \, \epsilon \, T$ $(t_m \supseteq t_i$, $t_n \supseteq t_j$, $t_m \wedge t_n = \Phi)$ are discovered. Such t_m , t_n must eventually be discovered if, as we have assumed, t_i does not occur in $t_1 \rightarrow \cdots \rightarrow t_j$ and t_j does not occur in $t_1 \rightarrow \cdots \rightarrow t_i$. Since $t_m \wedge t_n = \Phi$ and $t_m \supseteq t_i$ and $t_n \supseteq t_j$, then $t_i \wedge t_j = \Phi$. The proof is thus complete.

(17) Theorem. For any distinct t_i , $t_j \, \epsilon \, T$, if there is a $t_k \, \epsilon \, T$ such that t_i , $t_j \, \epsilon \, c(t_k)$, then $L(t_i) = L(t_j)$. Proof: By Theorem 14, $L(t_k)$ is unique; let $L(t_k) = m$. Then, by 13b, $L(t_i) = L(t_j) = m+1$. Remark: The converse of Theorem 17 does not hold in general. That is, the following statement is NOT true: For any t_i , $t_j \, \epsilon \, T$, if $L(t_i) = L(t_j)$, then t_i and t_j are in the same contrast set. We demonstrate this by counterexample. Consider $T = \{t_1 , t_2 , t_3 , t_4 , t_5 , t_6 , t_7\}$ where $t_1 \rightarrow t_2 \rightarrow t_4$, $t_1 \rightarrow t_2 \rightarrow t_5$, $t_1 \rightarrow t_3 \rightarrow t_6$, $t_1 \rightarrow t_3 \rightarrow t_7$. Clearly, Axioms 2 and 3 are satisfied. $L(t_4) = L(t_6) = 2$. However, there is no $t_i \, \epsilon \, T$ such that $t_i \rightarrow t_4$ and $t_i \rightarrow t_6$. That is, although t_4 and t_6 are at the same level, they are not in the same contrast set.

(18) Def. A taxonomic structure $\tau = \langle T, \supset \rangle$ has DEPTH n iff there is a $t_i \, \epsilon \, T$ such that $L(t_i) = n$ and there is no $t_j \, \epsilon \, T$ such that $L(t_j) > n$.

(19) Theorem. If a taxonomic structure $\tau = \langle T, \supset \rangle$ is of depth n, then for each integer m $(0 \leq m \leq n)$ there is at least one $t_i \, \epsilon \, T$ such that $L(t_i) = m$. Proof: By Def. 18, there is some $t_j \, \epsilon \, T$ such that $L(t_j) = n$. Consider the sequence $t_1 \rightarrow \cdots \rightarrow t_j$. By applying the recursive definition of level (Def. 13) to succeeding terms in this sequence, we establish that the sequence has $n+1$ terms and the ith term has level $i-1$. The proof is complete.

(20) Theorem. The subset $S_a = \{t_i \mid t_i \, \epsilon \, T$ and t_i is terminal$\}$ of T is a partition of t_1 . Proof: If t_i , $t_j \, \epsilon \, T$ are both terminal, then by Def. 11 neither $t_i \supset t_j$ nor $t_j \supset t_i$. Hence, by Theorem 16, any distinct t_i , $t_j \, \epsilon \, T$ which are terminal are disjoint. It remains to be shown that the union of all $t_i \, \epsilon \, S_a$ exhausts t_1 . For an arbitrary individual x, if $x \, \epsilon \, t_1$ then there must be a $t_i \, \epsilon \, T$ such that $L(t_i) = 1$ and $x \, \epsilon \, t_i$ (Axiom 3). Now, either $c(t_i)$ is null or it is a partition of t_i (Axiom 3). Hence either t_i is terminal and $x \, \epsilon \, t_i$ or there is exactly one $t_j \, \epsilon \, T$ such that $t_j \, \epsilon \, c(t_i)$ and $x \, \epsilon \, t_j$. Again t_j may either be terminal, in which case there is a terminal member of T that contains the arbitrary individual x, or x is a member of exactly one member t_k of $c(t_j)$. Continuing in this way, we eventually find a terminal t_n such that $x \, \epsilon \, t_n$ (since T is finite). We have thus established that for an arbitrary individual $x \, \epsilon \, t_1$, there is a terminal t_n such that $x \, \epsilon \, t_n$. That is, the $t_i \, \epsilon \, S_a$ exhaust t_1 . The proof is thus complete.

(21) Def. DIRECT CONTRAST = $\{ (t_i , t_j) \mid t_i$, $t_j \, \epsilon \, T$, $t_i \neq t_j$, t_i and t_j are in the same contrast set$\}$.

(22) Def. INDIRECT CONTRAST = $\{ (t_i , t_j) \mid t_i$, $t_j \, \epsilon \, T$ and there are t_m , $t_n \, \epsilon \, T$ such that $t_m \supseteq t_i$ and $t_n \supseteq t_j$ and $(t_m , t_n) \, \epsilon$ Direct Contrast and either $t_m \supset t_i$ or $t_n \supset t_j\}$.

If $(t_i , t_j) \, \epsilon$ Direct Contrast, we say equivalently 't_i and t_j contrast directly', 't_i and t_j are in direct contrast', 't_i and t_j are in the same contrast set'. If $(t_i , t_j) \, \epsilon$ Indirect Contrast, we say 't_i and t_j contrast indirectly via t_m and t_n respectively', 't_j and t_i contrast indirectly via t_m and t_n respectively', 't_i and t_j contrast indirectly'.

(23) Def. INCLUSION CONTRAST = $\{ (t_i , t_j) \mid t_i$, $t_j \, \epsilon \, T$ and either $t_i \supset t_j$ or $t_j \supset t_i\}$.

If $(t_i , t_j) \, \epsilon$ Inclusion Contrast, we say equivalently 't_i and t_j are in the relation of inclusion contrast', 't_i and t_j contrast inclusively'.

(24) Theorem. For any distinct t_i , $t_j \, \epsilon \, T$, one and only one of the following:

    (a) t_i and t_j contrast inclusively,

    (b) t_i and t_j contrast directly,

(c) t_i and t_j contrast indirectly.

Proof: We show first that conditions (a), (b), (c) are pairwise contradictory; that is, that the sets Direct Contrast, Indirect Contrast, and Inclusion Contrast are mutually exclusive. Conditions (b) and (c) are contradictory because of the provision in Def. 22 that 'either $t_m \supset t_i$ or $t_n \supset t_j$'. Conditions (a) and (b) are contradictory because, if t_i and t_j are in the same contrast set, then neither $t_i \supset t_j$ nor $t_j \supset t_i$, and conversely (Axiom 3, Definition 9). Conditions (a) and (c) are contradictory because if t_m and t_n contrast directly, then $t_m \wedge t_n = \Phi$. Then since $t_m \supseteq t_i$ and $t_n \supseteq t_j$, $t_i \wedge t_j = \Phi$. Consequently neither $t_i \supset t_j$ nor $t_j \supset t_i$. If $t_i \supset t_j$ or $t_j \supset t_i$, then t_i and t_j are obviously not in the same contrast set. It remains only to be shown that either (a), (b), or (c) holds for every pair of distinct t_i, $t_j \in T$. Consider the sequences $t_1 \to \cdots \to t_r \to t_i$ and $t_1 \to \cdots \to t_s \to t_j$. If t_i occurs in the latter or t_j in the former, then (a) holds for t_i, t_j. Iff $t_r = t_s$, (b) holds for t_i, t_j. If neither (a) nor (b) holds, we read down the two sequences in the way we did in the proof of Theorem 16, eventually discovering t_m and t_n such that $t_m \supseteq t_i$, $t_n \supseteq t_j$, and t_m, t_n are in direct contrast. Thus, at least one of the three conditions (a), (b), (c) must hold for any t_i, $t_j \in T$. The proof is complete.

(25) Def. TERMINAL CONTRAST $= \{ (t_i, t_j) |\ t_i, t_j \in T, t_i \neq t_j,\ t_i \text{ and } t_j \text{ are each terminal} \}$. If $(t_i, t_j) \in$ Terminal Contrast, we say equivalently 't_i and t_j are in terminal contrast', 't_i and t_j contrast terminally'.

(26) Theorem. For any t_i, $t_j \in T$, if t_i and t_j are in terminal contrast then t_i and t_j are not in inclusion contrast. Proof: By Defs. 11 and 25, there is no $t_k \in T$ such that $t_i \supset t_k$, and similarly there is no $t_k \in T$ such that $t_j \supset t_k$. Hence neither $t_i \supset t_j$ nor $t_j \supset t_i$. The proof is complete. Remark: For any t_i, $t_j \in T$, if t_i and t_j are in terminal contrast, then

    (a) t_i and t_j may or may not contrast directly,

    (b) t_i and t_j may or may not contrast indirectly.

We demonstrate this remark by examples. Assume $T = \{t_1, t_2, t_3, t_4, t_5\}$ where $t_1 \to t_2 \to t_3$, $t_1 \to t_2 \to t_4$, $t_1 \to t_5$. Axioms 2, 3 are satisfied.

    (a) t_3 and t_4 contrast terminally and also directly; t_4 and t_5 contrast terminally but do not contrast directly.

    (b) t_4 and t_5 contrast terminally and also indirectly; t_3 and t_4 contrast terminally but do not contrast indirectly. The demonstration is complete.

(27) Theorem. For any $t_j \in T$ ($t_j \neq t_1$), there is exactly one contrast set $c(t_i)$ ($t_i \in T$) such that $t_j \in c(t_i)$. Proof: This theorem follows directly from Axiom 3 and Theorem 12. Remark: In an arbitrary taxonomic structure $\tau = \langle T, \supset \rangle$, it is NOT true in general that, for any terminal t_i and t_j (t_i, $t_j \in T$), $L(t_i) = L(t_j)$. This remark is established by example, as follows: Consider $T = \{t_1, t_2, t_3, t_4, t_5\}$ where $t_1 \to t_2 \to t_3$, $t_1 \to t_2 \to t_4$, $t_1 \to t_5$. Axioms 2 and 3 are satisfied. t_4 and t_5 are terminal, but $L(t_4) = 2$ while $L(t_5) = 1$.

(28) Def. There is a unique, non-null subset G of T which partitions t_1 and whose members $g_i \neq t_1$ are called GENERIC. Remark: In each empirical case, the generic taxa must be isolated by empirical operations, subject only to the above conditions.

(29) Def. GENERIC CONTRAST $= \{ (t_i, t_j) |\ t_i, t_j \in T, t_i \neq t_j,\ t_i \text{ and } t_j \text{ are each generic} \}$.

(30) Theorem. For distinct t_i, $t_j \in T$, if $(t_i, t_j) \in$ Generic Contrast, then t_i and t_j do not contrast inclusively. Proof: The theorem follows directly from Defs. 23 and 29. Remark: For any t_i, $t_j \in T$, if t_i and t_j are in generic contrast, then

    (a) t_i and t_j may or may not contrast directly;

    (b) t_i and t_j may or may not contrast indirectly;

    (c) t_i and t_j may or may not contrast terminally.

The demonstration of the remark that follows Theorem 26 serves also as a demonstration (i) for parts (a) and (b), if we assume that the terminal taxa in that example (t_3, t_4, and t_5) are also the generic taxa, and (ii) for part (c), if we assume alternatively that t_2 and t_5 are the generic taxa.

(31) Def. A taxonomy is a ternary relational structure $\mathfrak{J} = \langle \tau, L, m \rangle$ where

    (a) $\tau = \langle T, \supset \rangle$ is a taxonomic structure;

    (b) L is a finite, non-null set of lexemes on which a partition P is defined, of which each cell p_j is non-null;

    (c) m is a mapping of a subset T' of T onto P, which satisfies the following axiom:

(32) AXIOM. If distinct t_i, t_j ϵ T' each has the image p_j under m (p_j ϵ P; p_j $\subset$ L), then
    (a) either $t_i \supset t_j$ or $t_j \supset t_i$;
    (b) for each t_k ϵ T, if $t_i \supseteq t_k \supseteq t_j$ or $t_j \supseteq t_k \supseteq t_i$, then t_k also has the image p_j under m.

APPENDIX II: GREGG'S 1967 FORMULATION

Gregg's (1967) n-rank Linnaean structures are roughly comparable to taxonomic structures of depth n. Both involve a finite set of taxa arranged in a hierarchy of immediate precedence with a unique beginner. (In terms of graph theory, either may be represented by a digraph in the form of a rooted tree. Taxonomic structures add the restriction, absent from Gregg's Linnaean structures, that no vertex has a positive degree of one. The latter condition reflects the fact that taxonomic structures do not allow monotypic taxa; i.e., if a taxon has subtaxa, it has at least two of them.) Gregg's structures may be OPEN or NOT OPEN. There is nothing in the theory of taxonomic structure corresponding to Gregg's open Linnaean structures. Gregg's theory is thus the more general one. An open Linnaean structure is one in which the unique beginner contains at least one member contained in no terminal taxon. Thus, an open structure has at least one taxon that STRICTLY includes the union of the members of the contrast set it dominates. In symbols, there is in each open Linnaean structure at least one taxon t_i such that

$$t_i \supset \bigcup_{t_j \epsilon c(t_i)} t_j \, .$$

This condition violates Axiom 3, which, it will be recalled, requires that $c(t_i)$ be a partition of t_i for all t_i in T. It is not surprising that taxonomic structures are closed systems in Gregg's sense, since taxonomic structures prohibit monotypic taxa, and since the explicit motivation for Gregg's introduction of the notion of openness is to deal with the 'PROBLEM OF MONOTYPY' (Gregg 1967:204; emphasis in original). I will return briefly to Gregg's problem of monotypy below.

The remaining major differences between taxonomic structures and Linnaean structures as defined and classified by Gregg involve the notion of absolute taxonomic category. The notion of an absolute scale of generality of taxa—constituted by a series of categories such as species, genus, family, order, class etc., and assigning each taxon to one such category—plays a crucial role in Gregg's formulation, but does not exist as such in the present development. For Gregg each category is a set of disjoint taxa. An n-rank Linnaean structure contains n categories; the unique beginner is the sole member of category n, and each category C_i ($1 \leq i \leq n$) contains at least one taxon. Each taxon belongs to exactly one category, but the category of a taxon is not necessarily exactly one less than that of its immediate predecessor; nor are all terminal taxa necessarily in category one. If a Linnaean structure contains one or more taxa whose category is not exactly one less than that of its immediate predecessor, Gregg calls such a structure IRREGULAR. If a Linnaean structure contains one or more terminal taxa that are not in category one, Gregg terms it a TRUNCATED structure. Figure 11a depicts a truncated but regular, closed, 3-rank Linnaean structure (Gregg prefers 'not irregular', and also 'not truncated', 'not open', 1967:195 ff.) Figure 11b depicts a non-truncated, irregular, closed, 3-rank Linnaean structure. A glance at the figures will reveal that these structures are identical from the point of view of taxonomic structure. A Linnaean structure must be either truncated or irregular (or both) if its terminal taxa are not all at the same level (as 'level' is defined in taxonomic structure).

For completeness, Figures 12a-b show 3-rank Linnaean structures that are (a) truncated and irregular and (b) non-truncated and regular, respectively. Figure 12b (non-truncated, regular) corresponds of course to a taxonomic structure whose terminal taxa are all at the same level. Recalling that it is possible, but by no means necessary, that all the terminal taxa in a taxonomic structure have the same level, we arrive at the following general comparison of n-rank Linnaean structures and taxonomic structures:

(1) Both a Linnaean structure and a taxonomic structure consist of a finite set of taxa, hierarchically ordered by the relation of immediate precedence. (For a useful formal definition of hierarchy, see Gregg 1967:194–5.)

(2) Open Linnaean structures violate Axiom 2 of the theory of taxonomic structure. Taxonomic structures are thus, in Gregg's sense, closed.
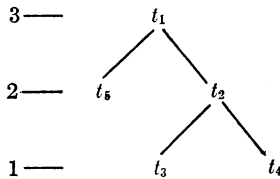
Category

FIGURE 11a. Abstract 3-rank, closed, truncated, regular Linnaean structure.
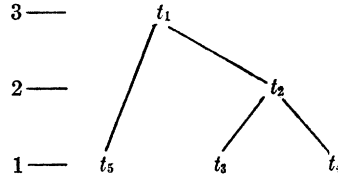
Category

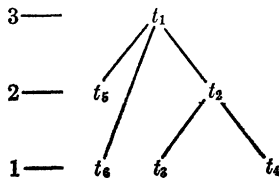FIGURE 11b. Abstract 3-rank, closed, non-truncated, irregular Linnaean structure.

Category

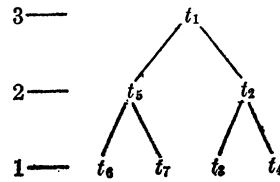FIGURE 12a. Abstract 3-rank, closed, truncated, irregular Linnaean structure.

Category

FIGURE 12b. Abstract 3-rank, closed, non-truncated, regular Linnaean structure.

(3) Gregg's notions of truncation and regularity, being based on a series of absolute, ranked categories, are not directly applicable to taxonomic structures. Any closed, n-rank Linnaean structure thus determines a unique taxonomic structure (of depth n).

(4) The converse of the last remark does not hold; that is, a given taxonomic structure does not determine a unique Linnaean structure. For example, Figures 11a–b diagram two distinct Linnaean structures but a single taxonomic structure.

Gregg's explicit motivation for the introduction of the notion of openness is to deal with what he calls the problem of monotypy (1967:201 ff., esp. 204). Gregg offers, as an example of monotypy, the fact that in Simpson's classification of mammals (1945) the subclass Protheria contains the single order Monotremata. Gregg proves that any structure containing monotypic taxa is open. (The converse does not hold: it is not the case that any open structure necessarily contains monotypic taxa. However, Gregg points out [personal communication] that even where there is no problem of monotypy, provision for undiscovered taxa is sufficient empirical motivation for openness.)

In the case of the protherians and the monotremes, Gregg's formulation commits the biologist to the empirical claim that there are some protherians that are not monotremes. Similarly, Gregg's formulation forces us to maintain that there are some members of the family Ornithorhynchidae that are not members of the genus Ornithorhynchus (1967:202–3). Thus, although Gregg (1967:205) accurately characterizes previous serious attempts to resolve the problem of monotypy—such as those of Beckner 1959, Sklar 1964, and van Valen 1964—as examples of 'technical artifice', it appears that his own attempt is not immune from that charge. Indeed, after presenting his example, Gregg says (1967:203), 'At this point, it is only fair to mention that the version of monotypy demanded by our model may not be generally acceptable to taxonomists; but for the moment we shall postpone discussion of the possible difficulties that are raised.'

The present formulation 'solves' the problem of monotypy by discarding the notion of absolute category, introducing in its place various definitions of types of contrast relation. It is unlikely that this will be considered a solution by biotaxonomists, as the intuitive notion of absolute category seems important in Western scientific biotaxonomy, although not in the folk taxonomies of natural languages, Western or otherwise. The present formu-

lation suffers the further defect that it does not allow for openness—the provision for as yet undiscovered taxa—a property that appears to be common both to Western scientific taxonomies and to natural folk taxonomies.

REFERENCES

BECKNER, M. 1959. The biological way of thought. New York: Columbia University Press.

BERLIN, BRENT. 1969a. A universalist-evolutionary approach to ethnographic semantics. (Language-Behavior Research Laboratory, working paper 27.) Berkeley: University of California.

——. 1969b. Speculations on the growth of ethnobotanical nomenclature. Unpublished.

BERLIN, BRENT; DENNIS E. BREEDLOVE; AND ROBERT M. LAUGHLIN. 1969. Lexical retention and cultural significance in Tzeltal-Tzotzil comparative ethnobotany. Paper presented at the Symposium on Mayan Ethnolinguistics, 68th Annual Meeting of the American Anthropological Association, New Orleans, La., 20–23 November. (Issued as Language-Behavior Research Laboratory, working paper 29. Berkeley: University of California, 1970.)

BERLIN, BRENT; DENNIS E. BREEDLOVE; and PETER H. RAVEN. 1966. Folk taxonomies and biological classification. Science 154.273–5. (Reprinted in Tyler, 60–6.)

——, ——, ——. 1968. Covert categories and folk taxonomies. American Anthropologist 70.290–9.

CONKLIN, HAROLD C. 1962a. Comment on Frake 1962. In Gladwin and Sturtevant, 86–91.

——. 1962b. Lexicographical treatment of folk taxonomies. Problems in lexicography, ed. by F. W. Householder and S. Saporta (Research Center in Anthropology, Folklore, and Linguistics, publication 21), 119–41. Bloomington: Indiana University. (Reprinted in Tyler, 41–59.)

——. 1964. Ethnogenealogical method. In Goodenough, 25–56. (Reprinted in Tyler, 93–122.)

FRAKE, CHARLES O. 1961. The diagnosis of disease among the Subanun of Mindanao. American Anthropologist 63.113–32.

——. 1962. The ethnographic study of cognitive systems. In Gladwin and Sturtevant, 72–85. (Reprinted in Tyler, 28–41.)

GLADWIN, THOMAS, and W. C. STURTEVANT (eds.) 1962. Anthropology and human behavior. Washington, D.C.: Anthropological Society of Washington.

GOODENOUGH, WARD H. (ed.) 1964. Explorations in cultural anthropology. New York: McGraw-Hill.

GREGG, JOHN R. 1954. The language of taxonomy: an application of symbolic logic to the study of classificatory systems. New York: Columbia University Press.

——. 1967. Finite Linnaean structures. Bulletin of Mathematical Biophysics 29.191–206.

——. 1968. Buck and Hull: a critical rejoinder. Systematic Zoology 17.342–4.

——, and F. T. C. HARRIS (eds.) 1964. Form and strategy in science. Dordrecht: Reidel.

KAY, PAUL. 1966. Comment on Ethnographic semantics: a preliminary survey, by B. N. Colby. Current Anthropology 7.20–3. (Reprinted with addendum in Tyler, 78–90.)

LOUNSBURY, FLOYD G. 1964. A formal account of the Crow- and Omaha-type kinship terminologies. In Goodenough, 351–94. (Reprinted in Tyler, 212–55.)

SIMPSON, G. G. 1945. The principles of classification and a classification of mammals. Bulletin of the American Museum of Natural History 85.i–xvi, 1–350.

SKLAR, A. 1964. On category overlapping in taxonomy. In Gregg & Harris, 395–401.

TYLER, STEPHEN A. (ed.) 1969. Cognitive anthropology. New York: Holt, Rinehart & Winston.

VAN VALEN, L. 1964. An analysis of some taxonomic concepts. In Gregg & Harris, 402–15.