

PROBLEMS IN LEXICOGRAPHY

Edited by
FRED W. HOUSEHOLDER
and SOL SAPORTA

Published by
INDIANA UNIVERSITY, BLOOMINGTON
MOUTON & Co., THE HAGUE, THE NETHERLANDS

LEXICOGRAPHICAL TREATMENT OF FOLK TAXONOMIES¹

By Harold C. Conklin
Yale University

0. Introduction

Many lexical problems are of considerable importance to linguists and ethnographers. With the interests of both groups in mind, I would like to discuss certain aspects of folk classification which I feel deserve more rigorous lexicographic attention than they have typically received.

An adequate ethnographic description of the culture [Goodenough 1957] of a particular society presupposes a detailed analysis of the communications system and of the culturally defined situations in which all relevant distinctions in that system occur. In this regard, accurate knowledge of both the grammar and lexicon of the local spoken language constitutes a minimum requirement. When the ethnographer works in an area for which adequate statements about the local language are unavailable in published sources, his first and often continuing task is the construction of a set of valid rules for the interpretation of the local language. In his phonological and grammatical analysis of new speech forms, he may find many helpful models in the descriptive linguistic literature. In attempting, however, to account for the obligatory semantic relations inherent in his lexical corpus, he may not be so fortunate. While extant dictionaries and vocabularies do provide glosses and definitional information, many of the nontrivial, and often essential, semantic and contextual relationships obtaining among lexical items are often either neglected or handled in an imprecise and unsystematic manner [cf. Newman 1954:86].

For formal linguistic analysis it is necessary that utterances be acceptable and interpretable grammatically. For ethnographic (including lexicographic) analysis utterances must also be acceptable and interpretable semantically. While an "appeal" to meaning does not improve grammatical analysis, neither does an

intuitive appeal to morphosyntactic form yield the most appropriate analysis of meaning and reference (see 1.5. below). In fact, an adequate grammar may generate semantically unacceptable propositions [Chomsky 1955:149, 1957:103-4; cf. Landar 1960:352; Frake 1961:113]. Results of some recent attempts to develop non-intuitive procedures for the evaluation of the grammaticalness and meaningfulness of sentences [e.g., Maclay and Sleator 1960; cf. Joos 1958] indicate that this difference is of considerable importance. The distinction between these two aspects of the analysis of speech is apparent even in the treatment of isolated forms.

In the course of several years of linguistic and ethnographic field work among the Hanunóo in the Philippines, it became abundantly evident that providing such segments as sah, tabākuq, samparansisku-qalistun, and lāda.balaynun.tagnānam.qiruṅ-pādiq each with the same gloss '(distinct) kind of plant' was--while adequate for certain syntactic purposes--most unsatisfactory for the task of semantic analysis. Had I not modified this procedure, I would have ended up with more than 2000 lexical items (including several hundred referential synonyms) each labeled identically. While employing glosses like 'tea' and 'tobacco' (in the first two cases above) proved useful in labeling familiar objects, the majority of these culturally significant Hanunóo designations referred to entities which to me were quite unfamiliar. In this type of ethnographic context one finds many instances where the problems faced traditionally by the compilers of bilingual dictionaries are considerably magnified [Nida 1958]. For the ethnographer, the semantic structure of such folk classification is of paramount significance. Upon his analysis of it depends the accuracy of many crucial statements about the culture being described. Problems of analyzing and presenting such structures in a succinct fashion may be of interest even to lexicographers who work only in relatively familiar cultural surroundings.

1. Folk classification

In the lexicographic treatment of folk classification, we are concerned primarily with (1) the identification of relevant syntactic segments, (2) the identification of fundamental semantic units in specific contexts, (3) the delineation of significant sets of semantic units in particular domains, and (4) the translation (and marking) of these units so that important semantic relationships will not be obscured. In discussing different systems of classifying segments of the natural and social environment, the neutral

term segregate [Conklin 1954] serves as a label for any terminologically-distinguished (i. e., conventionally-named) grouping of objects.

1.1. Linguistic structure. The shape and combinatorial structure of the linguistic forms which designate folk segregates are irrelevant, in a strict sense, to the analysis of the system of classification itself; i. e., to the semantic structure [Conklin 1957]. Labels and categories can change independently and therefore must be analyzed separately. On the other hand, a knowledge of the linguistic structure involved is essential for understanding the principles of folk nomenclature; and in working out this structure, clues for isolating folk segregate labels and for eliciting information about such segregates may be found.

1.2. Lexical units and contexts. A full lexical statement (i. e., an adequate dictionary) should provide semantic explanation, as well as phonological and grammatical identification, for every meaningful form whose signification cannot be inferred from a knowledge of anything else in the language. It is convenient to refer to these elementary lexical units as lexemes [cf. Swadesh 1946; Newman 1954; Jordan 1955; Goodenough 1956], although other terms have been suggested (e. g., idiom [Hockett 1956; cf. Householder 1959:508-24; Weinreich 1960:337]). So far as lexemic status is concerned, the morphosyntactic or assumed etymological relations of a particular linguistic form are incidental; what is essential is that its meaning cannot be deduced from its grammatical structure. Single morphemes are necessarily lexemes, but for polymorphemic constructions the decision depends on meaning and use (implying an analysis of the constraints imposed by the semantic structure, and the specification of relevant immediate contexts).

Formal segments such as black bird (vs. blackbird) or in the old house (vs. in the doghouse) can be excluded from the lexical statement because they are predictable, meaningfully, in that they can be considered semantically endocentric [Nida 1951:12-3, 1958:286; cf. Chao 1953:385]. Put another way, those constructions which are never semantically exocentric may be classed as nonlexemic forms (e. g., sunburned face, long pink strand). Problems do arise, however, in degrees of lexemic exocentricity [Nida 1958:286] and, again, if caution is not exercised in distinguishing clearly between grammatical and semantic criteria. The compounds firewater and silverfish, for example, are endocentric morphosyntactically (either on an attribute-plus-head basis or on

the perhaps stronger grounds of formal selection rules [Lees 1960:128, 158]), but semantically they are as exocentric as vodka and moth.

In the study of segregate labels in folk classification, and despite some of the difficulties of technical definition noted, I find it useful to distinguish by explicit semantic criteria two kinds of lexemic units: unitary lexemes (no segments of which may designate categories which are identical with, or superordinate to, those designated by the forms in question) and composite lexemes (one or more segments of which, under specified conditions, may [a] designate the same categories as those designated by the forms in question [abbreviation], or [b] designate categories superordinate to those designated by the forms in question [generalization], see 2.-2.2.) Unitary lexemes may be either simple (unsegmentable) or complex (segmentable). These distinctions are exemplified below:

<u>Lexemes</u>		
<u>Unitary simple</u>	<u>Unitary complex</u>	<u>Composite</u>
oak	poison oak	white oak
pine	pineapple	pitch pine
son	grandson	son-in-law
dart (an artifact)	darts (a game)	Baldwin apple
Jack	jack-in-the-pulpit	Port-orford cedar
dandelion	black-eyed Susan	black-crowned night heron
caterpillar (larva)	cat's-eye	caterpillar tractor

For contrast, consider a few similar but nonlexemic forms: cheap pine, pine and oak, black-eyed Joe, darts (plural of dart [Hockett 1956:229]). For a native speaker, such distinctions cause little concern, but in new linguistic and cultural environs difficulties may arise.

For example, on first inspection, the following partially-identical Hanunóo forms [Conklin 1954] might appear to belong to a simple paradigm (they could all be recorded during a conversation about rice cultivation and weeding problems):

1 <u>paray:paray</u>	'cattail'
2 <u>pāray·māyah</u>	'immature wild <u>pādaŋ</u> (plant)'
3 <u>pāray·qiŋkantuh</u>	'kind of wild sedge'

<u>4</u> <u>pāray</u> · <u>bīhud</u>	'kind of rice'
<u>5</u> <u>pāray</u> · <u>tāwuh</u>	'some one (else)'s rice'
<u>6</u> <u>pāray</u> · <u>tīdah</u>	'that rice'

The glosses, however, indicate that several types of lexical units may be involved. Are there any formal linguistic clues?

Each of the six forms is easily segmented into two morphs, as I have indicated by the use of dots. Loose-joining, phonemically, is represented by a single raised period. Except for the closely-joined doubling in item number 1, the forms in this set provide no obligatory intonational or junctural contrasts. Furthermore, each form occurs in many identical frames such as tūhay ṅāni tī _____. '_____ is (are) certainly different.' Thus, for most of the semantically distinct types of joining suggested by the glosses, there are no phonological clues and few, if any, immediate, formal indications. (A full syntactic statement covering the structure of compounds would separate out some of these forms on grammatical grounds [cf. Lees 1960].) Given the necessary semantic information, however, these distinctions can be noted easily for lexicographical purposes by rewriting the forms as follows:

<u>1</u> <u>parayparay</u>
<u>2</u> <u>pāray-māyah</u>
<u>3</u> <u>pāray-qīṅkantuh</u>
<u>4</u> <u>pāray</u> · <u>bīhud</u>
<u>5</u> <u>pāray</u> <u>tāwuh</u>
(<u>5a</u>) (<u>5b</u>)
<u>6</u> <u>pāray</u> <u>tīdah</u>
(<u>6a</u>) (<u>6b</u>)

This procedure clearly marks 1, 2, and 3 semantically exocentric, unitary lexemes; 4 as a composite lexeme; and 5 and 6 as non-lexemic, semantically endocentric constructions the initial lexeme of which is superordinately related to 4. Minimally, forms 1, 2, 3, 4, 5a, and 6a could be labeled 'kind of plant,' but by not attending to essential semantic distinctions this type of short cut would obscure such important contrastive relations as the mutual exclusion of coordinate categories (1 : [pādaṅ], implied--but not covered--by the specific growth stage term number 2) : 3 : 5a or 6a), and the possible total inclusion of subordinate categories (4 by 5a/6a; but not 1, 2, or 3 by 5a/6a). Statements about such relations, hinted at in some glosses and definitions, may be demonstrated only by systematic pairing in minimal, and relatively

controlled, linguistic and semantic contexts.

1.3. Lexical sets and domains. In many ways it can be said that the more discrete the phenomena referred to, the simpler the task of treating the associated terminology in a lexicographically adequate manner [cf. Wallace and Atkins 1960]. If this is true for particular lexical items it is equally true for the semantically structured sets which such items may comprise [Frake 1961]. Minimally, a lexical set consists of all semantically contrastive lexemes which in a given, culturally relevant context share exclusively at least one defining feature [Lounsbury 1956:61-2]. The semantic range of all such lexemes defines the domain of the lexical set. The initial establishment of domain boundaries, while widely recognized as an ideal goal, is often a very difficult task [cf. Voegelin and Voegelin 1957]. Effective eliciting frames and procedural tests used to determine such boundaries, and convincing demonstrations of their intracultural reality, are subjects not often discussed in the linguistic or ethnological literature. Some of the essential factors involved in this type of analysis are treated briefly below under "levels" (2.) and "dimensions" (3.) of contrast. In general, the number and complexity of boundary problems increases as one moves from the investigation of lexical domains within a particular language to an attempt to "match" the domains of different languages [Ohman 1953; cf. Quine 1960:26-79]. This does not, however, preclude rigorous contrastive analysis.

1.4. Translation and semantic structure. With few exceptions, the lexical items employed in systems of folk classification always comprise a segment of the everyday vocabulary of the particular language [Conklin 1957]. The rules governing the obligatory semantic relations among the categories in such lexical sets are thus to be determined, evaluated, and described for each language. Such rules cannot be prescribed merely on the basis of familiarity in another system with the "concrete" denotata of the sets involved. In the case of folk botany, for example, this means that a local system of plant classification cannot be described accurately by attempting to obtain only vernacular "equivalents" for botanically recognized species. Translation labels (glosses) are frequently necessary, but they should be considered neither as definitions nor as exact equivalents [Lounsbury 1956:163; for an attempt to use acronyms as a partial mnemonic solution to such translation problems, see Landar et al. 1960: 371]. This well-established and perhaps obvious semantic principle is sometimes forgotten where the assumed absolute nature

(in a cross-linguistic sense) of "scientific" names or of other long-established traditional distinctions in certain Western languages is involved [Öhman 1953; cf. Simpson 1961:11].

1.5. Syntactic vs. semantic structure. Implicit in the preceding remarks is the assumption that the relation between formal linguistic (syntactic, in the general, semiotic sense [Morris 1946]) structure and semantic structure need not be isomorphic [Lounsbury 1956:189]. If this assumption is taken seriously, a full dictionary should state explicitly the necessary and sufficient conditions for the unambiguous structural interpretation of each included lexeme in the context of the total lexicon as well as in that of the grammar. While such coverage has rarely been achieved, even for relatively small lexical domains, I feel that recognition of this goal has considerable relevance for this discussion. A brief illustration may help to indicate the kind of crucial lexical data that are often ignored, especially where meanings are either assumed on the basis of semantic patterning in a more familiar language, or where they are treated only partially (as in the derivation of definitional statements from translational labels).

Consider the following situation (which, with minor differences, I have encountered on a number of occasions): a woman, whose brother (x) and husband (y) are both named Juan, has a son, also named Juan (z) and a daughter who in turn has a son named Pedro (P). The genealogical situation is diagrammed in Fig. 1 (we can ignore the broken lines for a moment). Two fluent speakers of English, F, a Filipino whose first language was Tagalog, and A, a native speaker of a dialect of American English, both know Pedro and the specified members of his family. The fact that one of the Juans (x, y, or z) has died is known only to A (or F) who in turn wishes to relate this circumstance to his friend F (or A). A straightforward statement completing the sentence P's _____ Juan died would seem to do the trick; and, depending on the circumstances, one of two unitary lexemes (grandfather, uncle) might be used to fill the blank:

- 1 Pedro's Grandfather Juan died.
- 2 Pedro's Uncle Juan died.

However, if A uses Grandfather, F may ask Which grandfather?; if F uses Uncle, A may ask Which uncle? indicating a kind of two-way ambiguity which can only be resolved by recognizing that despite their unquestionable grammaticality and morphosyntactic identity,

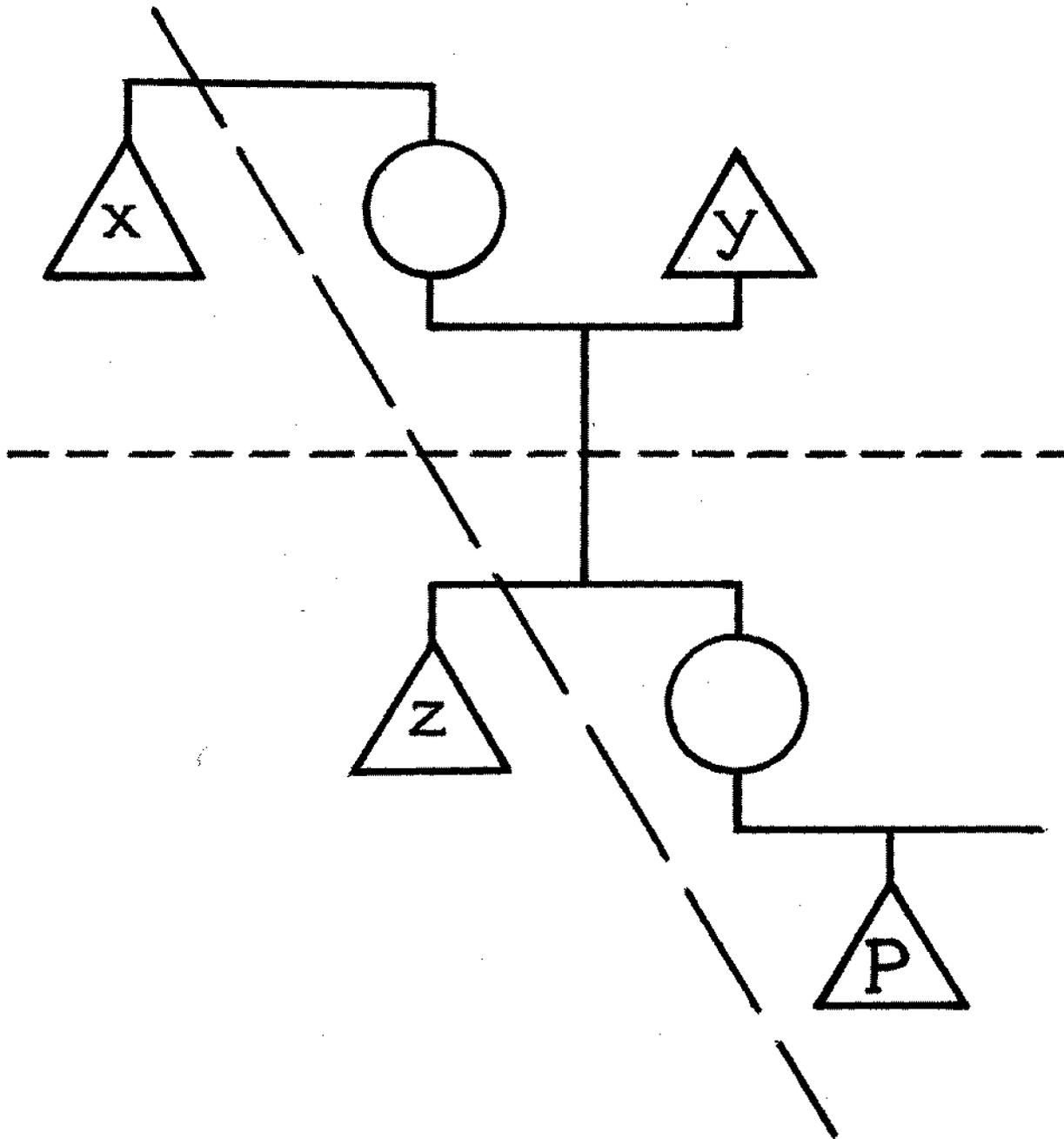


Figure 1. A genealogical illustration of contrasting systems of kinship classification.

A's sentences 1 and 2 and F's sentences 1 and 2 differ semantically:

<u>Sentence</u>	<u>Kin term used</u>	<u>Kin type(s) included (Pr = parent's)</u>	
A1	<u>Grandfather</u>	y	(PrFa)
F1	"	x and y	(PrFa, PrPrBr)
A2	<u>Uncle</u>	x and z	(PrBr, PrPrBr)
F2	"	z	(PrBr)

This, of course, reflects only a small part of a very fundamental structural difference in Central Philippine and North American systems of kinship classification: universal terminological recognition of generation in the former vs. universal terminological recognition of degree of collaterality in the latter (these two "limits" to the lexical extension of kin class membership are indicated on the kinship diagram in Fig. 1 by the horizontal and diagonal broken lines, respectively). Although any careful investigator might learn this systematic distinction after a few days of field work, the principle goes unaccounted for in the relevant and extant bilingual dictionaries. The restrictions involved in this illustration are just as obligatory and inescapable within the respective semantic systems represented as is the distinction of singular vs. plural in English grammar.

2. Levels of contrast

Folk categories within the same domain may be related in two fundamentally different ways: by inclusion, which implies separate levels of contrast, and by exclusion, which here applies only within single-level contrastive sets. There may also be subcategoric, or componential, intersection (see 3. below). In studying semantic relationships, as among folk categories, it has often been demonstrated that likeness logically and significantly implies difference [Kelly 1955:303-5]. It is also pertinent, however, to note that total contrast (complete complementary exclusion)--which logically relates such segregates as ant and ship or cough and pebble--is less important than restricted contrast within the range of a particular semantic subset (compare the relations within and between the partial sets robin - wren - sparrow; spaniel - terrier - poodle; and bird - dog). When we speak of the category dime being included in the category coin we imply that every dime is also a (kind of) coin--but not necessarily the reverse. Furthermore, when we state (a) that the category dime contrasts with that of quarter and (b) that the category coin contrasts with that of

bill we are speaking of two instances of relevant mutual exclusion at two different levels of contrast [Conklin 1955, 1957; Frake 1961]. Such alignments of folk categories are common to all languages, though systematic indications of these relationships are rare even in the more detailed monolingual dictionaries.

2.1. Hierarchic structure. Where the articulation between successive levels, each consisting of a set of contrastive lexical units, is ordered vertically by inclusion such that each monolexemic category at one level is totally included in only one category at the next higher level, we can speak of a lexical hierarchy. The two axes of such a structure involve the horizontal differentiation of contrastive but coordinate categories and the vertical increase of generalization or specification resulting from ascent to superordinate (including) or descent to subordinate (included) levels, respectively [Gregg 1954; Conklin 1957; Beckner 1959; 55-80; Frake 1961:117]. These axes are fixed and cannot be merged or interchanged, nor can the succession of levels be modified. Dime is not contrasted with coin, but at the same level with nickel, quarter, penny, etc. Subhierarchies of varying "depths" are often discernible within larger hierarchic structures. The depth (in levels) of the subhierarchy including the categories hawk, pigeon, and starling is less than that of the subhierarchy including hawk, horse, and crocodile; i. e., the first three segregates are included in a superordinate category at a lower level than that of the segregate ultimately including hawk, horse, and crocodile. The embedding of shallow subhierarchies within increasingly deeper ones is characteristic of many systems of folk classification.

2.2. Folk taxonomy. A system of monolexemically-labeled folk segregates related by hierarchic inclusion is a folk taxonomy; segregates included in such a classification are known as folk taxa [Conklin 1957; cf. Lawrence 1951:53; Simpson 1961:19]. Some of the additional requirements of "model" or "regular" taxonomic systems [Woodger 1952:201ff.; Gregg 1954; Beckner 1959:55-8; Simpson 1961] are: (1) at the highest level, there is only one maximal (largest, unique) taxon which includes all other taxa in the system; (2) the number of levels is finite and uniform throughout the system; (3) each taxon belongs to only one level; (4) there is no overlap (i. e., taxa at the same level are always mutually exclusive). Folk systems vary widely with respect to these more specific "requirements," but the presence of hierarchically arranged though less "regular" folk taxonomies is probably universal. Most of the examples given here are taken from folk botany, but similar illustrations could be taken from other domains [Thomas

1957; Frake 1961].

Several important differences distinguish folk taxa from the taxonomic groups of biological systematics [Conklin 1957; Simpson 1961]. The former usually relate only to locally relevant or directly observable phenomena. They are defined by criteria which may differ greatly from culture to culture. The number and position of levels of contrast may change from one sector of a folk system to another. There are no formal rules for the nomenclatural recognition or rejection of taxa [cf. Lawrence 1951: 213-5], though new groupings may be added productively with considerable ease. In respect to any particular local biota, there is no reason to expect the folk taxa to match those of systematic biology--either in number or in range. The Hanunóo classify their local plant world, at the lowest (terminal) level of contrast, into more than 1800 mutually exclusive folk taxa, while botanists divide the same flora--in terms of species--into less than 1300 scientific taxa.

2.3. Special problems. Although they cannot be discussed here at length, a number of lexicographically important problems encountered in the analysis of folk taxonomies include:

(1) Multiple and interlocking hierarchies. Unlike scientific taxa, folk segregates may belong simultaneously to several distinct hierarchic structures. The same segregates may be classed as terminal categories in a taxonomy based on form and appearance and also as terminal or nonterminal categories in another taxonomy based on cultural treatment (e.g., morphologically distinguished kinds of floral segregates vs. functional categories of plants as food cultigens, medicines, ornamentals, etc.) [Conklin 1954]. Subhierarchies may be interarticulated in numerous ways [e.g., Joos 1956:296-7] and there is always the potentiality of partial inclusion or domain overlap.

(2) Extrahierarchic relations. Not all folk categories are directly related by class inclusion or contrast within the range of a particular superordinate category. For example, numerous difficulties may arise if lexemes designating separate ontogenetic stages or parts of members of particular segregates (see 1.2. above) are not distinguished from hierarchically arranged folk taxa [Chao 1953:387-9; Conklin 1954, 1957; Frake 1961]. Part-of (part-whole) relations are often complicated by ambiguities [Nagel 1961:381-3] not encountered in the analysis of kind-of (class inclusion) relations (e.g., the segregates plant, stem, sap are not related taxonomically like plant, tree, elm).

... homonymy and nomonymy. When, within the context of a particular folk taxonomy, a single taxon may be labeled by phonemically distinct forms, as in the case of minor dialect variants or abbreviated terms [see 1.2.], we may speak of referential synonyms (or synonymous lexemes); e. g., fin, finnif, five, fiver, five-spot, five-dollar bill. In many such cases, it may be difficult to demonstrate taxonomic identity and the absence of categoric overlap. Alternative substructuring of the subhierarchy may be involved. Phonemically identical (homonymous) lexemes may designate separate taxa of different ranges of generalization at successive levels. Such situations (e. g., animal and man in the following partial contrastive sets: animal¹ vs. plant, animal² vs. man¹, man² vs. woman [cf. Frake 1961:117-9]) are not uncommon but they require careful contrastive pairing and testing for inclusion at each level involved. Similar steps must also be taken in working out problems concerned with distinguishing polysemy from homonymy [Wells 1958:662-3; cf. Chomsky 1957:95; Garvin 1960:147].

(4) Types of contrast. Paired folk taxa of some lexical subsets are related by simple, binary, segregate opposition. Many larger sets and some dyadic ones involve important types of semantic contrast other than antonymy [cf. Lyons 1960:622]. Structurally, for example, taxa may be contrasted in serial, complementary, or discontinuous arrays. (For subcategoric attribute relations, see 3. below.)

2.4. Folk vs. botanical taxonomy. Ideally, in the study of interrelated lexical sets in folk taxonomies, priority and preference should be given to unanimously-agreed-upon, obligatory distinctions in specified contexts. When tested by means of what are essentially crucial experiments--by pairing and contrasting negatively and positively--one should be able to construct a model (i. e., a theoretical statement) of the hierarchic structure such that assertions of membership and inclusion in any of the implied taxa are unanimously and unambiguously denied whenever such assertions are incongruent (i. e., meaningless within the system) [cf. Joos 1958:65]. The assertion 'Poodles, dogs, and animals are kinds of snails' would thus be rejected by speakers of my dialect of English--and on very easily specified semantic grounds. Within a particular universe of discourse (a taxonomic domain) how can one construct a nontrivial model by means of which only semantically acceptable, congruent propositions may be generated? An example from Hanunóo folk botany may serve as a partial answer.

In a situation where one Hanunoo farmer wishes to draw another's attention to a particular individual pepper bush Q, he may, of course, attempt to describe some of Q's unique attributes without naming the plant. Much more often, however, even in the course of a "unique" description, he will resort to the use of one or more of at least eight lexical units each of which might complete the frame māluq, qinda pag _____, 'Hey, take a look at this _____,' but at different levels of contrast (allowing for different degrees of desired or required specificity):

I	<u>kuwag</u>	'entity' (i. e., something that can be named)
II	<u>bāgay</u>	'thing' (not a person, animal, etc.)
III	<u>kāyuh</u>	'plant' (not a rock, etc.)
IV	<u>qilamnun</u>	'herbaceous plant' (not a woody plant, etc.)
V	<u>lādaq</u>	'pepper (plant)' (not a rice plant, etc.)
VI	<u>lāda. balaynun</u>	'houseyard pepper (plant)' (not a wild pepper plant)
VII	<u>lāda. balaynun. mahārat</u>	'houseyard chili pepper (plant)' (not a houseyard green pepper plant)
VIII	<u>lāda. balaynun. mahārat. qūtin-kutiq</u>	'"cat-penis" houseyard chili pepper (plant)' (not a member of any of five other terminal houseyard chili pepper taxa such as <u>lāda. balaynun. mahārat. tāhud-manuk</u> , the "cock's-spur" variety).

Within the domain of Hanunoo plant taxonomy, from level III down, and specifically within the range of lādaq, from level V down, conversations recorded during many similar situations would

6-11-7

ultimately provide the lexicographer with fifteen unitary and composite lexemes (including a terminal set of eleven 'pepper plant' names) arranged at four levels in the form of a discrete subhierarchy (Fig. 2). Specification below the level of the terminal taxa noted in the diagram (Fig. 2:1-11), and hence outside this system of classification, may be provided only by semantically endocentric constructions describing individual plant variations, on which unanimous accord is rare and unpredictable. In this particular case, folk taxa 15, 14, and 11 happen to correspond rather closely with the scientific taxa Capsicum, C. annum L., and C. frutescens L., respectively; but the twelve remaining folk taxa involve distinctions not recognized as significant botanical subspecies by taxonomic botanists who have classified the same flora. Structurally speaking, however, some of the most important patterns of semantic contrast involve not only the hierarchic separation of these varied, lower-level, folk taxa, but also a large number of nonhierarchic relations governed by sublexemic class intersection (see 3.2). Although such relations cannot be diagrammed with the taxonomic implications of Fig. 2, nor can they be treated effectively at all in terms of our hierarchic model, they should nevertheless be of considerable interest to linguists and others concerned with systems of folk classification.

3. Dimensions of contrast

At any given level within a well-defined folk-taxonomic subhierarchy, the relations obtaining among three or more coordinate taxa may involve varying dimensions, or kinds of subcategory contrast. The conjunction of these dimensions, or more precisely, of the values (or specific attributes [cf. Bruner *et al.* 1956:26-30]) along the several dimensions, define the categories involved within an essentially paradigmatic (i. e., nonhierarchic) subsystem [Lounsbury 1956, 1960:27-8; for a discussion of the structurally similar though more typologically-oriented procedures of attribute space substruction and reduction see Greenberg 1957 and Lazarsfeld 1961].

3.1. Nonhierarchic structure. Such multidimensional contrasts do not imply, and indeed do not allow, the ordering of the resultant categories by hierarchic inclusion. These features of nonhierarchic semantic structures, while not always sharply distinguished from the principles inherent in hierarchic systems, have been recognized and carefully analyzed in a number of domains, notably in kinship [Goodenough 1951:92-110, 1956; Lounsbury 1956;

<u>kuwaq</u>										
<u>bāgay</u>										
<u>kāyuh</u>										
<u>qilamnun</u>										
15 <u>lādaq</u>										
14 <u>lāda. balaynun</u>										<u>lada. tirindukun-</u> <u>tigbayaq</u>
12 <u>lāda. balaynun. mahārat</u>						13 <u>lāda. balaynun. tagnānam</u>				
<u>l. b. m.</u> <u>batūnis</u>	<u>l. b. m.</u> <u>hapun</u>	<u>l. b. m.</u> <u>pasītih</u>	<u>l. b. m.</u> <u>pinasyak</u>	<u>l. b. m.</u> <u>qūtin-</u> <u>kutiq</u>	<u>l. b. m.</u> <u>tāhud-</u> <u>manuk</u>	<u>l. b. t.</u> <u>mali-</u> <u>puṅkuk</u>	<u>l. b. t.</u> <u>pasītih</u>	<u>l. b. t.</u> <u>pātuktuk</u>	<u>l. b. t.</u> <u>qarābaq</u>	
1	2	3	4	5	6	7	8	9	10	11

Figure 2. A segment of Hanunóo plant taxonomy. All folk taxa included in the taxon lādaq, are indicated.

Frake 1960; Wallace and Atkins 1960], color [Conklin 1955; cf. Lenneberg and Roberts 1956; Landar et al. 1960], orientation [Haugen 1957], disease [Frake 1961], and, beginning with Jakobson's pioneering efforts, in such partly modulational [Joos 1958:70] paradigms as case and pronoun systems [Jakobson 1936; Sebeok 1946; Harris 1948; Lotz 1949; Wonderly 1952; Austerlitz 1959]. The following example of multidimensional contrast in a regular paradigmatic structure will illustrate some of these points.

3.2. Significant classification vs. cataloguing. If, omitting the high-level, wide-ranging kuwaq (see 2.4), we list all the Hanunóo personal name substitutes occurring in various frames such as māluq, qinda pag binwat ni _____, 'Hey, take a look at what _____ did (here),' we will invariably end up with an exhaustive and mutually exclusive lexical set consisting of just eight units (in each case representing a single morpheme). Arranged in the least meaningful type of catalogue, an alphabetical index (as in a dictionary), these lexical units are:

<u>dah</u>	'they'
<u>kuh</u>	'I'
<u>mih</u>	'we'
<u>muh</u>	'you'
<u>tah</u>	'we two'
<u>tam</u>	'we all'
<u>yah</u>	'he, she'
<u>yuh</u>	'you all'

The shapes provide little that is structurally suggestive, but the glosses do indicate that an ordering in terms of eight "traditional" distinctions along three quasi-semantic dimensions

- (1) first person : second person : third person
- (2) singular : dual : plural
- (3) exclusive : inclusive

might be attempted. But the resulting applied structure is hardly elegant, economical, or convincing:

<u>kuh</u> 1s	<u>tah</u> 1d	<u>mih</u> 1pe
---	---	<u>tam</u> 1pi
<u>muh</u> 2s	---	<u>yuh</u> 2p
<u>yah</u> 3s	---	<u>dah</u> 3p

If a close examination is made of the distinctive contrasts involved, not in terms of labels but in terms of actual, minimal, obligatory differences, a more satisfactory, economical, and

semantically verifiable solution is reached. The necessary and sufficient conditions for defining each of the eight categories depend on the regular intersection of six components which comprise three simple oppositions:

minimal membership : nonminimal membership (M: \bar{M})
 inclusion of speaker : exclusion of speaker (S : \bar{S})
 inclusion of hearer : exclusion of hearer (H : \bar{H})

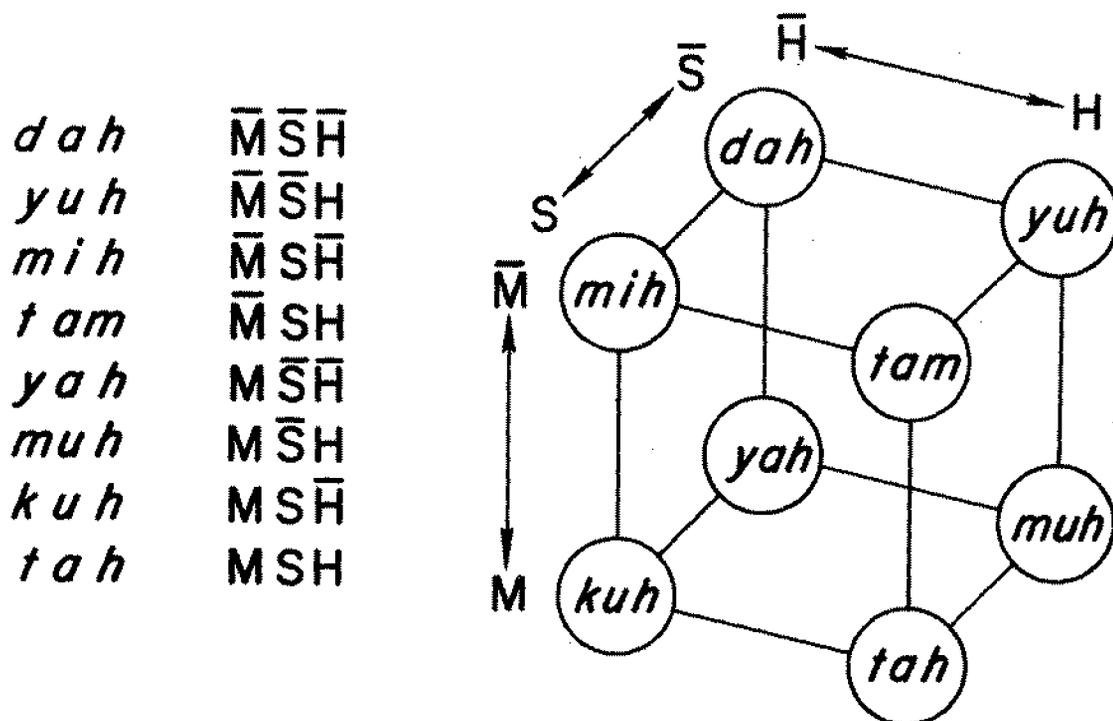


Figure 3. Paradigmatic structure of a Hanunóo pronominal set.

These relations can be represented in list or diagrammatic form (Fig. 3). Even without further elaboration, the basic semantic structure of this lexical set should now be clear. (In passing, it may be noted that pronoun systems in Tagalog, Ilocano [Thomas 1955], Maranao [McKaughan 1959], and some other Philippine languages exhibit very similar, if not identical, obligatory semantic relationships.)

This example also illustrates a very important, though perhaps less obvious, characteristic of paradigmatic relations at one level in a taxonomic subhierarchy in contrast to the noncommutative relations of class inclusion governing the larger taxonomic system. Within such a contrastive lexical set (as in Fig. 3), ordered by

class intersection, the constituent categories cannot be arranged in a taxonomic hierarchy. Any arrangement (e. g., a circular, block, or branching diagram) superficially appearing to contradict this statement will prove on closer inspection either (a) to constitute what the biologists call a key [Mayr et al. 1953:162-8; Simpson 1961:13-6], essentially another kind of catalogue or finding list ordered by successive--but not necessarily taxonomically significant--dichotomous exclusion, or (b) to be based on some other artificially imposed, and hence semantically nonsignificant, classification.

4. Lexicographic treatment

The ways in which the problems mentioned in this paper may be treated in bilingual dictionaries, especially ethnographic dictionaries, are practically unlimited. That very few of the possibilities have been explored to date is disappointing, but not discouraging. There have been a number of new attempts at expanding the analytic procedures of descriptive linguistics to include a more rigorous, thorough, and theoretically rewarding analysis of semantic structure [e. g., Goodenough 1956; Lounsbury 1956; Nida 1958; Frake 1961]. Despite these more encouraging signs, I realize that most dictionaries will continue to be organized primarily as alphabetical indices. Suggestions regarding the ways in which structural semantic information (especially with reference to folk taxonomies) might be more adequately covered in such dictionaries would include, wherever possible: (1) consistent marking of each entry as to its status as a lexical unit and taxon, its immediately subordinate taxa and superordinate taxon, and all coordinate taxa included with it in this next higher taxon (simple diacritics and abbreviations can be devised for systematic use in compilation and checking); (2) differential marking of translation labels and of definitions; (3) concise indication of distinctive attributes which define categories belonging to analyzed lexical sets; (4) systematic cross-referencing to maximal taxa in all major subhierarchies, to referential synonyms, and to all units involved in categoric overlap; and (5) frequent use of structural charts and diagrams. Where only limited opportunities are available for accomplishing such tasks, priority might be given to those parts of the lexicon which, on the basis of nonintuitive and intracultural criteria, appear to involve semantic relations of an everyday, obligatory nature. In number of segregates, paradigmatic complexity, and hierarchic depth, certain lexical domains are likely

to be more highly structured than others [Brown 1956:307; Nida 1958:283-4; Worth 1960:277; Frake 1961:121-2]. For the student of folk taxonomy, focusing attention on these domains should lead not only to more interesting analytic problems but also to results of greater lexicographical and general cultural relevance.

5. References cited

- Austerlitz, Robert. 1959. "Semantic components of pronoun systems: Gilyak." Word 15.1:102-109. New York.
- Beckner, Morton. 1959. The Biological Way of Thought. viii, 200 pp. New York: Columbia University Press.
- Brown, Roger W. 1956. "Language and categories." pp. 247-312 (Appendix) in A Study of Thinking, by Jerome S. Bruner et al. New York: John Wiley and Sons.
- Bruner, Jerome S., Jacqueline J. Goodnow, and George A. Austin. 1956. A Study of Thinking (with an Appendix on Language by Roger W. Brown). xii, 330 pp. New York: John Wiley and Sons.
- Chao, Yuen Ren. 1953. "Popular Chinese plant words, a descriptive lexico-grammatical study." Language 29.3:379-414. Baltimore.
- Chomsky, A. Noam. 1955. "Semantic considerations in grammar." Georgetown University Monograph Series on Languages and Linguistics No. 8, pp. 141-150. Washington, D. C.
- Chomsky, Noam. 1957. Syntactic Structures. 116 pp. 's-Gravenhage: Mouton.
- Conklin, Harold C. 1954. "The relation of Hanunóo culture to the plant world." Unpublished Ph. D. dissertation in anthropology, Yale University. New Haven.
- _____. 1955. "Hanunóo color categories." Southwestern Journal of Anthropology 11:339-344. Albuquerque.
- _____. 1957. "Ethnobotanical problems in the comparative study of folk taxonomy." Paper read at the Ninth Pacific Science Congress, 1957. Bangkok. [Published in Proceedings, Ninth Pacific Science Congress, 1957, 4:299-301, Bangkok, 1962.]
- Frake, Charles O. 1960. "The Eastern Subanun of Mindanao." pp. 51-64 in Social Structure in Southeast Asia, ed. George Peter Murdock. Chicago: Quadrangle Books.

- _____. 1961. "The diagnosis of disease among the Subanun of Mindanao." American Anthropologist 63.1:113-132. Menasha.
- Garvin, Paul L. 1960. "On structuralist method." Georgetown University Monograph Series on Languages and Linguistics No. 11, pp. 145-148. Washington, D.C.
- Goodenough, Ward H. 1951. Property, Kin, and Community on Truk. Yale University Publications in Anthropology. No. 46, 190 pp. New Haven.
- _____. 1956. "Componential analysis and the study of meaning." Language 32.2:195-216. Baltimore.
- _____. 1957. "Cultural anthropology and linguistics." Georgetown University Monograph Series on Languages and Linguistics No. 9, pp. 167-173. Washington, D.C.
- Greenberg, Joseph H. 1957. "The nature and uses of linguistic typologies." International Journal of American Linguistics 23.2:68-77. Baltimore.
- Gregg, John R. 1954. The Language of Taxonomy: an application of symbolic logic to the study of classificatory systems. xii, 70 pp. New York: Columbia University Press.
- Harris, Zellig. 1948. "Componential analysis of a Hebrew paradigm." Language 24.1:87-91. Baltimore.
- Haugen, Einar. 1957. "The semantics of Icelandic orientation." Word 13.3:447-459. New York.
- Hockett, Charles F. 1956. "Idiom formation." pp. 222-229 in For Roman Jakobson. The Hague: Mouton.
- Householder, Fred W., Jr. 1959. Review of: A course in modern linguistics, by Charles F. Hockett. Language 35.3:503-527. Baltimore.
- Jakobson, Roman. 1936. "Beitrag zur allgemeinen Kasuslehre." Travaux du Cercle Linguistique de Prague 6:240-288. Prague.
- Joos, Martin. 1956. Review of: Machine translation of languages: Fourteen essays. Ed. William N. Locke and A. Donald Booth. Language 32.2:293-298. Baltimore.
- _____. 1958. "Semology: a linguistic theory of meaning." Studies in Linguistics 13.3-4:53-70. Buffalo.

- Jorden, Eleanor Harz. 1955. "The syntax of Modern Colloquial Japanese." Language 31.1 (part 3): i-vi, 1-135. Baltimore.
- Kelly, George A. 1955. The Psychology of Personal Constructs. Vol. 1., A Theory of Personality. xviii, 566 pp. New York: W. W. Norton.
- Landar, Herbert J. 1960. "A note on accepted and rejected arrangements of Navaho words." International Journal of American Linguistics 26.4:351-354. Baltimore.
- Landar, Herbert J., Susan M. Ervin, and Arnold E. Horowitz. 1960. "Navaho color categories." Language 36.3:368-382. Baltimore.
- Lawrence, George H. M. 1951. Taxonomy of Vascular Plants. xiv, 823 pp. New York: Macmillan.
- Lazarsfeld, Paul F. 1961. "The algebra of dichotomous systems." pp. 111-157 in Studies in Item Analysis and Prediction, ed. Herbert Solomon. Stanford: Stanford University Press.
- Lees, Robert B. 1960. "The grammar of English nominalizations." International Journal of American Linguistics 26.3 (part II): i-xxvi, 1-205, (Publication 12, Indiana University Research Center in Anthropology, Folklore, and Linguistics). Baltimore.
- Lenneberg, Eric H., and John M. Roberts. 1956. "The language of experience, a study in methodology." International Journal of American Linguistics Memoir No. 13. xi, 33 pp. Baltimore.
- Lotz, John. 1949. "The semantic analysis of the nominal bases in Hungarian." Travaux du Cercle Linguistique de Copenhague 5:185-197. Copenhague.
- Lounsbury, Floyd G. 1956. "A semantic analysis of the Pawnee kinship usage." Language 32:158-194. Baltimore.
- _____. 1960. "Similarity and contiguity relations in language and culture." Georgetown University Monograph Series on Languages and Linguistics No. 12, pp. 123-128. Washington, D. C.
- Lyons, J. 1960. Review of: Language change and linguistic reconstruction, by Henry M. Hoenigswald. Bulletin of the School of Oriental and African Studies 27. 3:621-622. London.

- Maclay, Howard, and Mary D. Sleator. 1960. "Responses to language: judgments of grammaticalness." International Journal of American Linguistics 26.4:275-282. Baltimore.
- Mayr, Ernst, E. Gorton Linsley, and Robert L. Usinger. 1953. Methods and Principles of Systematic Zoology. x, 328 pp. New York: McGraw-Hill.
- McKaughan, Howard. 1959. "Semantic components of pronoun systems: Maranao." Word 15.1:101-102. New York.
- Morris, Charles W. 1946. Signs, Language, and Behavior. 365 pp. New York: Prentice-Hall.
- Nagel, Ernest. 1961. The Structure of Science: problems in the logic of scientific explanation. xvi, 618 pp. New York and Burlingame: Harcourt, Brace and World.
- Newman, Stanley. 1954. "Semantic problems in grammatical systems and lexemes: a search for method." pp. 82-91 in Language in culture, ed. Harry Hoijer. Chicago: University of Chicago Press.
- Nida, Eugene A. 1951. "A system for the description of semantic elements." Word 7.1:1-14. New York.
- _____. 1958. "Analysis of meaning and dictionary making." International Journal of American Linguistics 24.4:279-292. Baltimore.
- Öhman, Suzanne. 1953. "Theories of the 'Linguistic Field.'" Word 9.2:123-134. New York.
- Quine, Willard Van Orman. 1960. Word and Object. xvi, 294 pp. Cambridge Mass., and New York and London: The Technology Press of M. I. T., and John Wiley & Sons, Inc.
- Sebeok, Thomas A. 1946. "Finnish and Hungarian case systems: their form and function." Acta Instituti Hungarici Universitatis Holmiensis, Series B, Linguistica 3. Stockholm.
- Simpson, George Gaylord. 1961. Principles of Animal Taxonomy. xiv, 247 pp. New York: Columbia University Press.
- Swadesh, Morris. 1946. "Chitimacha." pp. 312-336 in Linguistic Structures of Native North America, by Harry Hoijer, et al. New York: Viking Fund.

- Thomas, David. 1955. "Three analyses of the Ilocano pronoun system." Word 11.2:204-208. New York.
- _____. 1957. "An introduction to Mansaka lexicography." 6 typescript pp. Nasuli, Philippines: SIL.
- Voegelin, Charles F. and Florence M. 1957. "Hopi domains, a lexical approach to the problem of selection." International Journal of American Linguistics Memoir No. 14, 82 pp. Baltimore.
- Wallace, Anthony F. C., and John Atkins. 1960. "The meaning of kinship terms." American Anthropologist 62.1:58-80. Menasha.
- Weinreich, Uriel. 1960. "Mid-century linguistics: attainments and frustrations." Romance Philology 8.3:320-341. Berkeley.
- Wells, Rulon. 1958. "Is a structural treatment of meaning possible?" Proceedings of the Eighth International Congress of Linguists, Oslo, 1957, pp. 654-666. Oslo.
- Wonderly, William L. 1952. "Semantic components in Kechua person morphemes." Language 28.3:366-376. Baltimore.
- Woodger, J.H. 1952. Biology and Language: an introduction to the methodology of the biological sciences including medicine. xiv, 364 pp. Cambridge: The University Press.
- Worth, D.S. 1960. Review of: Leksikologija anglijskogo jazyka, by A. I. Smirnickij. Word 16.2:277-284. New York.

NOTES

1. The work on which most of this paper is based has been supported by the National Science Foundation. A number of students and other friends have offered constructive criticism of an earlier draft of this statement. For especially helpful and more detailed comments I am particularly indebted to Y. R. Chao, David Crabb, Arthur Danto, C. O. Frake, Paul Friedrich, W. H. Goodenough, J. H. Greenberg, Einar Haugen, P. F. Lazarsfeld, F. G. Lounsbury, and Volney Steffire.