Generic taxa in Hawaiian

Brett Kessler

December 25, 1995

Abstract

A computer-assisted study was performed on all native Hawai‘ian folk genera to see to what extent the Hawai‘ian system accords with Berlin's theory of ethnobiological taxonomy. It was found that the system can indeed be described as some 600 genera grouped in a shallow hierarchy, with a small percentage of them, mostly taxa of some cultural significance, being subdivided further into taxa named by contrast sets built on the name of the genus. Particular attention was applied to the question of correspondence between the Hawai‘ian system and the Linnaean. Berlin's prediction that smaller life forms tend to be named more inclusively (i.e., correspond to higher Linnaean ranks) held up well, in contrast to an alternate reading that holds evolutionary complexity to be the causal factor; however, economic importance, which may be linked to size, can not be ruled out as the underlying cause. No support was found for the idea that a culture tends to name at either the level of Linnaean genus or species, but not both. It was further found that

Hawaiʻian genera with the shortest names are more likely to correspond to higher Linnaean ranks.

1 Introduction

Berlin (1992) presents a theory of ethnobiological taxonomy that is on the whole congenial and credible. His philosophy is that taxa have objective existence, and that all humans will naturally classify and to a large extent denominate the living world in roughly the same way. The classification will be based on inherent properties of the taxon itself rather than on its use, and will be organized into a fairly shallow hierarchy. Primary names are given to taxa at or above a certain perceptually salient rank in the hierarchy (his GENUS); below that, taxa are distinguished mostly with the help of qualifiers. Such extension of the taxonomy to subgeneric entities is characteristic of agricultural societies.

Much of Berlin's theory falls out directly from adopting a universalistic view of human cognition. Specialists in child language acquisition note that children across cultures acquire vocabulary in similar patterns. They automatically pick up on the idea that nouns denote classes of distinct objects, and that these classes are discrete. At very young ages they learn to apply vocabulary at different levels. When confronted with different names for the same object, they naturally adopt the strategy of assuming that one name is more inclusive than the other. Moreover, they pick up one level as being more basic than other levels. They tend to learn words at that level more readily. When confronted with names that are more specific, they regularly use that name as a qualifier to the basic name. For example, they may perceive *car* as a basic level term, and when taught *Volvo*, insist on saying *Volvo-car*. The hierarchies are fairly shallow, and often include unnamed categories. So children perceive that mammals have something in common long before learning *mammal* or even *animal*. This is evidenced, for example, by their extension of the word *dog* to designate animals whose name they do not know, even though comprehension experiments clearly show that they normally know what *dog* really means, and that they are using the term faute de mieux. Researchers also find that children tend to classify objects in terms of inherent features such as shape, and only rarely in terms of colour or function (Eve Clark, p.c.).

Thus I find that most of Berlin's strongest claims, not to mention his fundamental anti-Whorfian stand in itself, are strongly supported by current understanding of universals of the acquisition of vocabulary and semantics. Indeed, it appears that the bulk of his claims are not claims about ethnotaxonomy in particular, but rather are extensions to biology of claims that are made of lexical systems in general. Although one could conceive of possible universes where the facts were different, I don't think most people would be shocked by Berlin's principles (p. 21-26), especially if recast in terms of general lexical acquisition. Principle 1 says that we don't have names for everything; 2 says that we classify based on form, not function; 3 says that we have levels of naming; 4 says that things grouped in classes don't have to be identical, and that classes are perceived as distinct, not continuous; 5 says that basic-level names are the most common, and that more specific names are reserved for cases where making distinctions is important; 6 says that basic and more specific classes have prototypicality (some beds are more bed-like than others). Similarly his principles of nomenclature (p. 26-31) are largely uncontroversial within the universalist tradition: 1 says that the higher level classes are generally unnamed; 2-4 say that contrast sets of names *(sofa bed, water bed)* are used if and only if the class name is more specific than the basic level, unless the class is a prototype or the class names something very important (in which case ordinary names may be used); and 5 says that names tend to refer to some inherent feature of the object named. I accept all this without question.

A few of his other claims are more questionable but are not easy to research from a distance. He claims that biological taxa are organized into exactly six levels of classification, or ranks. Certainly six is the right ballpark for a general upper limit to the number of class levels one might expect for biota: at a branching level as small as 3, this would accommodate 729 taxa, which exceeds the number of taxa he claims are found in the richest systems. But it is far from obvious that six is exactly the right number. More particularly, it is questionable that rank itself is a primary concept, beyond the idea that a particular item in a hierarchy can be designated as being at the basic level. Rather, I would expect that class inclusion is a recursive concept: one groups related objects into a class generally without ever attempting to peg that class to some class rank. Just introspecting, my naive classification of an oak is that is a deciduous tree, which is a tree, which is a woody plant, which is a vascular plant, which is a plant. I naively think of a chihuahua as a dog, which is a canine, which is a carnivorous mammal, which is a mammal, which is a legged land animal, which is a chordate, which is an animal. It doesn't bother me that some hierarchies may be longer than others, I could easily accept extra nodes in a hierarchy when I learn more about biotes (or reflect more deeply on them), and I certainly don't have any intuitions about whether chordates are equivalent to vascular plants. And clearly these hierarchies don't fit into the six ranks Berlin lays out.

Furthermore one might wish to question the very idea that classification hierarchies are necessarily proper trees with a single root. People can easily place things into different categories depending on the criteria they employ. From its ethology, a large insect that walks about on land seems in many ways similar to a small reptile; from its anatomy, it doesn't seem much like a vertebrate at all. Whether people consistently think of one classification as primary is an interesting question. Another is whether there is something special about biota that make people do so, or whether the classification tasks that have been presented to them force them to do so. If people in a given society really do consistently give one particular classification when forced to make unique groupings, that fact in itself is interesting, but it is somewhat different from saying that a society recognizes one unique proper taxonomical hierarchy. In fact I suspect that the nebulous state of “intermediate” taxa in Berlin’s theory and their low degree of cross-cultural agreement may mean precisely that people tend to agree on classifications when class members are objectively close together, such that anatomical similarities are overwhelming (genus level), but they get fuzzier when asked to group things that are farther apart, when various classificatory criteria are at hand.

Although it is easy to theorize abstractly about some of the aspects of Berlin’s approach to taxonomy, it is difficult to really test them without the benefit of field work. Few sources give much information about what native terms subsume other terms, and none discuss covert categories unless incidentally. There are some claims, however, that are subject to examination using reference sources. Berlin claims that tropical horticulturalists will have up to 500 or 600 generic (basic-level) terms for plants and animals, and that 80% of these will contain no subtaxa. Generic terms can be identified because subtaxa will be named by contrast sets incorporating the genus name, except for certain species of high cultural significance. Furthermore, since taxa have objective existence, people in different cultures should identify the same taxa, and this includes Western biologists. Berlin specifically claims that generic taxa will correspond to genus or species in Linnaean taxonomy for higher life forms, but to higher ranks for lower life forms. It has also been claimed more precisely that some cultures tend to give basic names to what corresponds to Linnaean genera, and others to Linnaean species.

I have investigated these claims with respect to the native taxonomic system of Hawai‘i. Hawai‘i is an agricultural society that technically lies within the tropics, although it should be recognized that its climate is more temperate than typical tropical areas. It remained completely isolated from most of the world until two hundred years ago, so far all practical purposes one may consider its taxonomical system to be a natural continuation of the native Polynesian, with no non-Polynesian influences until Cook. Accounts of the vocabulary and native flora and fauna were begun soon after contact, and it is easy by and large to distinguish native species from introduced ones, and quite easy to identify borrowed vocabulary. Thus Hawai‘i constitutes an ideal closed system. In several respects, however, Hawai‘i is unusual. Its extreme isolation means that the terrestrial fauna is highly restricted (mostly insects and birds), and the flora is atypical, consisting of a high percentage of ferns, for example. Hawai‘ians depended heavily on food species they brought with them from abroad, and to a very large extent on food from the sea. Among tropical horticulturalist societies, Hawai‘i may be about as far removed as is possible from the South American societies Berlin investigated. This is of course both good and bad. If Berlin's universal predictions are wrong, Hawai‘i would seem a good candidate for a counterexample. On the other hand, it should be admitted that Hawai‘i is odd enough that it could rightly be considered a special case.

2 Procedure

2.1 Data

The first step was to compile a list of the Hawai‘ian words that name taxa of biota. For this I relied primarily on Pukui and Elbert (1986), which is by far the most authoritative and comprehensive dictionary of the language. This in turn draws upon (among others) Abbott and Williamson (1974) and Reed (1907) for seaweeds; Neal (1965), Handy (1940), Krauss (1974), Rock (1913a), St. John (1975a, 1975b, 1982), *Taro* 1970, Titcomb (1948), Whitney et al. (1939), and Hillebrand (1888) for land flora; Pyle (1977) and Berger (1981) for birds; Gosline and Brock (1960), Summers (1964), Titcomb (1972), and Kay (1979) for aquatic fauna; and McKeown (1978) for land fauna. I have not consulted these works myself, although that would be a logical next step if I continued this project. I extracted the names by the simple method of reading through each entry in the Hawai‘ian-English section of the dictionary. I also spot-checked in the English-Hawai‘ian section, which was particularly useful when entries there gave lists of ferns, etc. I also cross-checked by looking up all the names mentioned in Mitchell (1992) and in Handy (1958), who drew also on Rock (1913b), Degener (1945), and Hubbard (1952). This not only verified that the dictionary was complete (it didn't lack any of the names found in those works), but also gave a good grounding in which biota were culturally significant.

When extracting names, I systematically excluded taxa that clearly belonged below Berlin's generic level. This was most clear-cut when the names formed contrast sets. For example, the *lehua* tree has several varieties known as *lehua ʻApane* 'red lehua', *lehua kū-ma-kua* 'standing-back lehua' (a variety with sessile cordate leaves), *lehua lau liʻi* 'small-leafed lehua', *lehua maka noe* 'misty-faced lehua' (a bog-dwelling variety), and *lehua mamo* '*mamo* lehua' (i.e., yellow like the *mamo* bird). A more difficult case was when there were several taxa narrowly delineating an item of great cultural importance. Berlin recognizes that domesticated biota are more likely to be given subgeneric names (p. 24) and also that those subgeneric names are more likely not to form contrast sets (p. 30). That doesn't seem to leave any linguistic means of identifying them as subgeneric. I therefore fell back on the strictly cultural test of domestication, omitting names of taxa that clearly were differentiated through human influence. In Linnaean terms, these amounted to cultivars of a single species. This was most important in the case of types of taro *(kalo)* (300 varieties), the sweet potato *(ʻuala)* (200 varieties), banana *(maiʻa)* (50 varieties), and sugar cane *(kō)* (40 varieties). Oddly enough, none of my sources have indicated such a proliferation of subgeneric names for domesticated animals (the dog, pig, and chicken). I did not apply such a filter to fish, even though they were often kept in ponds.

I also excluded entries for which the Linnaean category was not given by the dictionary. Almost all names were so identified. The exceptions were a few cases that were obvious, and didn't need identification, such as *īlio*, 'dog', in which case I supplied the Linnaean name (*Canis familiaris*). Other exceptions were some names denoting biota of uncertain identity. The fact that some names were omitted means there was a small but not negligible skew in the database toward plants of some cultural significance; such a bias is almost inevitable in any sort of study.

Finally, I omitted taxa that I had good reason to believe were introduced to Hawai'i after European contact. In addition to explicit statements that a species was introduced, I also relied on clues such as names that are adapted from a European word, or names that include elements such as *haole* 'foreign'. Doubtless many biota slipped by that are in fact non-native. This could have been controlled more carefully by researching each of the taxa in the scientific literature, but that would have been prohibitively time-consuming. The important thing is to weed out words that are obviously built on a secondary basis, which has been accomplished for the worst offenders at least.

After collecting some 642 taxa, I attempted to identify names that were suprageneric. *Kupukupu* refers to ferns, particularly those with a single stalk. *Mau'u* refers to all sedges and grasses. *Limu* is usually identified as 'seaweed', but it also includes any plant that grows in water, as well as the lower cryptogams. The word *lā‘au* is identified in the dictionary as meaning ‘plant’. The root idea seems to be ‘stick’, and may refer to woody plants especially. Unfortunately it is hard to tease out whether a pre-contact Hawaiian would have used it to refer to cryptogams or sea plants or even to grasses. I have tentatively equated it with land plants, including ferns and sedges, but in opposition to *limu*. When collecting statistics, I gathered under *lā‘au* any plant that does not belong to the other three categories, so that if someone has a broader conception of its meaning, they can simply add the statistics together. In the animal kingdom, *manu* is ‘bird’, and *i‘a*, which is usually translated ‘fish’, actually refers to any water animal. Here again, it is hard to tease out the exact meaning of the broad term *holoholona*. This is usually translated ‘animal’, and clearly includes insects. Since its root meaning is to ‘walk’, I have interpreted this as being land animals, in opposition to air and water animals, but again, further research is needed in order to understand the pre-contact usage. I grouped all the identified Hawaiian taxa under these seven categories, which Berlin might consider life-forms, and further distinguished vertebrates from invertebrates. In addition, I identified several words that are probably to be considered intermediate categories; it can be surprisingly difficult to decide subclass relationships from dictionary entries. Into this category I placed *aloalo* ‘hibiscus’; *ha‘iwale* ‘slender shrubs’; *kolokolo* ‘creeping vines’; *līpālāwai* ‘algae’; *nalo* ‘two-winged insects’; *leho* ‘cowries’; *naunau* ‘shellfish’; *‘ōlepe* ‘bivalves’; *‘opihi* ‘limpets’.

The information was placed in a flat file suitable for `awk` processing. The records were multi-line entities separated by blank spaces; the fields were single lines, untagged. A line beginning with an asterisk introduced a section of the file dedicated to a new life form, and an ampersand at the beginning of a record indicated that it (probably) referred to an intermediate taxon. All the rest of the records (622) were implicitly genus-level. The fields for each record are identified by position. The first line gives the basic Hawai'ian name for the taxon, in the modern orthography; I used a backquote to indicate glottal stop, and capital letters to indicate long vowels. The second line gives the Linnaean name for the taxon. If the Hawaiian name corresponds to more than one Linnaean taxon, all were given, separated by commas or semicolons. Normally this was a genus/species pair, but often a genus alone. Occasionally the taxon was above the genus level, in which case I preposed a designation identifying it as *class, subclass, ord., subord., fam.* or *subfam.* When the sources gave several words identified as Hawaiian subgenera which corresponded to different Linnaean taxa, I listed all the Linnaean taxa in the record for the Hawaiian genus. For some records, I added a tag to each taxon, identifying the Hawai'ian qualifier that corresponded to the taxon, but I began doing this after doing much of the work, and did not go back and consistently expand the earlier records. Consequently, no use is made of that information. The third field of the record gives an English gloss. The fourth field gives Hawai'ian synonyms. This was distinguished from the name in the first field only by considering which one the dictionary chose to give the main definition under.

The data were processed by means of an `awk` programme. Each Hawai‘ian taxon was treated as a Hawai‘ian names (field 1) paired with an array of Linnaean names (field 2). The Linnaean names were processed in order to resolve the taxonomic rank of each component, with abbreviations like *var.* and *subsp.* being ignored, and genus initials expanded. All the taxa were compared with each other to ensure that no two had the same Linnaean name. A few dozen conflicts arose, and they were studied in depth to ascertain whether the entries were synonyms (in which case they were merged) or whether they stood in a subtype relation, in which case either the broader term was marked as intermediate (and henceforth ignored) or the narrower term was reevaluated to see whether it was really subgeneric.

The most important test performed by the programme was to determine what Linnaean rank each Hawai‘ian taxon corresponded to. This may seem straightforward, since the syntax of the Linnaean names makes it clear what their rank is: a single capitalized name is a genus, unless it is preceded by one of the suprageneric rank abbreviations; two names is specific; three subspecific. But if a Linnaean genus is monotypic, a species name for all intents and purposes also denotes the genus. Even if the genus is polytypic, the same clearly holds if only one of the species is found in Hawai‘i. If more than one species is found in Hawai‘i but only one of them is named, then it would appear to be just as reasonable to take it as a genus name. Either the other species aren't known to the natives, in which case the situation is just like the preceding, or the Hawai‘ians didn't consider the difference between the species to be salient enough to require a different name. The recorder of the Linnaean name would have presumably given the name for the species he or she examined, which if we're lucky was a prototype, but which could just as well have been the first exemplar encountered. On the other hand, species that are effectively monotypic do not allow one to decide for or against the hypotheses that the culture prefers to name things at the level of the Linnaean genus, or at the level of the species. Note that the need to determine whether multiple species within the same genus have different Hawai‘ian names means that this kind of study cannot be performed with a random sample of the lexicon, but must be reasonably comprehensive.

In accordance with these considerations, the programme breaks the Hawai‘ian taxa into four categories. Taxa known to include biota from more than one Linnaean genus are classified as "above the genus level". Taxa known to include more than one species from the same genus are classified as "at the genus level". It should be noted that it is possible in some cases that the genus is a member of a family that is effectively monotypic. However, this case wasn't considered because this information wasn't available in my sources, and at any rate the number of taxa classified in this category was quite small. If a taxon names a species or subspecies in a genus which has another species or subspecies named by another taxon, both of those taxa are classified "below the genus level." The remainder—those taxa identified as a particular species in an effectively monotypic genus—are classified as "at genus or species level." These are the ones that I have argued are basically generic, but which do not constitute evidence against the hypothesis that Hawai'ians like to name things at the species level.

For each life form, the programme counts how many taxa are in that category (*lā'au*, land plants, are counted separately from *kupukupu*, ferns, and *mu'u*, sedges). It tells how many taxa are in each of the four categories described above. For those above and below the genus level, it lists the taxa, grouping the latter by the Linnaean genus. The programme then sums the information across all life forms.

The programme also computes the correlation between the length of the Hawai'ian name and the Linnaean rank the name corresponds to. This was done as a poor man's morphological analysis, since it would be very difficult to determine the etymology of all 622 of the Hawai'ian names in the database. The hypothesis is that longer names will be more likely to be compound. Berlin's theory is silent as to whether there is any correlation between the structure of primary names and the Linnaean rank. The implication is that there should not be: primary names normally denote taxa at the native genus rank and above, and the mapping between native and Linnaean ranks is based on pragmatic, not linguistic, considerations. For this study, length of name was determined in three ways: by segment count, by mora count, and by syllable count. In addition to a printed report, the programme generates data files suitable for input into David Harrison's `xgraph` programme. The correlation computed is the standard Pearson's *r* statistic.

The data file, programme, and its raw output are included as appendices to this report. Please refer to the programme for further information as to the exact algorithms that were used.

3 Results

Table 1 summarizes how the Hawaiʻian generic taxa align with the Linnaean taxa. Recall that *lāʻau* are essentially Angiosperms other than grasses; *kupukupu* are ferns (Filicineae); *mauʻu* are grasses and sedges; and *limu* are plants other than Pteropsida (algae, seaweeds, fungi, mosses). *Holoholona* are land animals—both vertebrates and invertebrates. *Manu* are birds. *Iʻa* are water-dwelling animals, here broken down into invertebrates and vertebrates.

The columns of the table show what percent of the taxa in these native categories are clearly above the level of the Linnaean genus (i.e., they comprise two or more genera); clearly at the level of the genus (i.e., they comprise two or more species in one genus); are ambiguous between the genus and the specific rank (they name a species in a monotypic genus); and are clearly below the level of the genus (they name a species in a genus in which there are other terms for other species).

Taxon	Above	Genus	Ambig	Below	Total
lāʻau	11%	13%	48%	29%	231
kupukupu	16%	3%	36%	45%	58
mauʻu	22%	15%	30%	33%	27
limu	8%	14%	63%	16%	51
holoholona v'brate	0%	0%	100%	0%	3
holoholona inv'brate	53%	6%	41%	0%	17
manu	4%	9%	47%	40%	55
iʻa inv'brate	28%	8%	55%	9%	65
iʻa v'brate	22%	10%	53%	15%	115
Total	16%	10%	49%	25%	622

Table 1: Linnaean rank of Hawaiʻian taxa.

	Above	Genus	Ambig	Below
segment	-0.59	-0.33	-0.06	0.68
mora	-0.89	0.27	0.74	0.57
syllable	-0.98	0.59	-0.08	0.38

Table 2: Correlation between word length and Linnaean rank.

Table 2 summarizes how good a correlation there is between the length of the Hawai‘ian word, computed in segments, moras, or syllables, and the Linnaean rank it corresponds to.

More detailed data can be gleaned by reading the printout. For each Hawai‘ian life form, it lists the Hawai‘ian taxa that are definitely above and below the Linnaean genus level. It also tells exactly how many taxa names are of each length in the three different measuring systems. An appendix to this report also presents the latter information in graphic form.

4 Discussion

During the course of this study, I informally verified that many of Berlin's predictions hold for Hawai‘ian. The taxa are arranged in a shallow hierarchy. Below a certain level, further distinctions are made by adding to the basic word terms that are more or less descriptive. The terms are added after the basic term, since Hawai‘ian compounds and noun phrases are left headed. For example, the word *puhi* ‘eel’ can be used alone or qualified by the words *ao* ‘(day)light’, *hāpala* ‘painted’ (*Gymnothorax pictus*), *ho‘olā‘au* ‘stick-like’, *kāpā* ‘squeezing’ (a fierce species), *kāpīpī* ‘lightly salted’, *kīna‘u* ‘stained’, *koali* ‘morning-glory’, *kumuone* ‘sandstone’ (sand-coloured), *l=a‘au* ‘stick’, *lau milo* ‘milo leaf’, *lei hala* ‘pandanus lei’, *nanaka* ‘cracked’, *nau ‘ai* ‘food-chewing’, *nonokea* ‘pink?’, *nuku ‘ula* ‘red-mouthed’, and so on. Most of these terms are no different in kind from the terms one might use in English to describe species and cultivars. In particular, they do tend to denote inherent properties or behaviours of the plant or animal, not the use people put them to. There were, however, occasional qualifiers, such as *ʻoʻa*, that have no independent meaning. I also noted many cases of prototyping. *Hau* plain and simple designates the *Hibiscus tiliaceus* tree; *hau hele ʻula,* ʻrubescent hau' and *hau kuahiwi* ʻmountain hau' name related trees (*Kokia* and *Hibiscadelphus*). Furthermore, Berlin's prediction that culturally significant biota will be classified with more granularity holds up. Taro, sweet potatoes, bananas and sugar cane all had dozens or even hundreds of words, many of which could be used independently, as primary names.

On a more quantitative tack, the total count of 622 generic Hawaiʻian taxa holds up pretty well to Berlin's prediction that tropical horticulturalists will have an upper limit of five to six hundred (p. 23). And my impression was that he was broadly correct in predicting that some 80% of the generic taxa would be monotypic, although I did not systematically collect subgeneric names so that I could quantify this precisely.

My main purpose in this project was to investigate the hypothesis that societies tend to give basic-level names at either the Linnaean genus rank or the species rank, but not both. The data in Table 1 do not suggest a clear answer, partly because 49% of all the names are for the only named species in a genus. Certainly such names cannot be used to disprove either that the

Hawai‘ians prefer to name at the genus level or that they prefer the species level. Even if we were able to show that all of the ambiguous names should be classed as genus-level, there would still be a residue of 41% that are not genus-level. Or if they were all shown to be species-level, then there would still be a residue of 26% that are not species-level. So even in the best of cases, where we could show that all of these ambiguous-level terms name species in polytypic genera and that Hawai‘ians refuse to apply the names to the other species in the same genus, we could claim only that 74% of the basic names are at the species level, which is not an overwhelming predominance. As it stands, however, we should discard those ambiguous-level taxa, in which case we see that 52% of the remaining biota are named at the genus level or above, and 48% are named below the genus level, which is about as close to random as one can get. The hypothesis that cultures name either at the genus or the species level should be rejected.

Berlin himself, however, gives a much more cautious prediction as his principle 7. On p. 25-26, he states:

> For the higher vascular plants and larger vertebrate animals, generic taxa often approximate in their content the genera and species of Western scientific biology. For the smaller vascular plants, lower cryptogams, smaller vertebrates, and many invertebrates, the correspondence of folk generic taxa more closely approximates scientific taxa of the ranks of family, order, or class.

This is not a very strong statement. No prediction at all is made to distinguish between the genus and the species level, or between the family, order, and class levels. Berlin divides the biological world into two parts, but is vague as to whether the difference is one of size or evolutionary complexity, or where the cutoff should be in either case. Where do ferns fit in, which don't seem to me to be either "higher vascular plants" or "lower cryptogams" but rather higher cryptogams or lower vascular plants. Do grasses qualify as higher vascular plants or smaller ones? Finally, the statement "often approximates" is quite a hedge. The statement would be fulfilled if several such correspondences could be found in some system in the world, but it does not even require that such correspondences predominate. Surely if we divide the Linnaean ranks into two parts and the biological world into two parts, the specified correspondence would "often" occur even if the correspondences were assigned totally at random.

Much as one might enjoy tearing the statement apart, it does contain the germ of a verifiable claim: that there is an inverse correspondence between the size and/or complexity of a plant or animal on the one hand and the height of the Linnaean rank its folk generic term will correspond to. It might be interesting to see to what extent my data support the prediction, although it should however be kept in mind that the study was not set up for this purpose. In particular, the database does not contain information about the average size of mature specimens. Neither does it contain information to determine whether a genus is a member of a monotypic family. Although this case is less common than that of monotypic genera, there is the possibility that some of the names I have classified at the genus level could actually be ambiguous between genus and family. For example, since the Hawai‘ians knew of only one rodent, the rat, there is no obvious way of deciding whether their word *‘iole* should be considered a species, genus, family, or order term. (I don't think it's any help to know that they called the non-native beaver a *‘īlio hulu pāpale*, ‘hat-fur dog’.)

Keeping this danger in mind, the breakdown in Table 1 does nevertheless provide some useful information. Among the plants, *mau‘u* ‘grasses’ have the highest percentage of taxa named above the genus level (22%), followed by *kupukupu* ‘ferns’ at 16%, *lā‘au* ‘other land plants’ at 11%, and *limu* ‘sea plants’ at 8%. The same ranking obtains if one counts the percentage of taxa unambiguously above the species level (37%, 29%, 24%, 22%). There is no obvious ranking here in terms of Western notions of evolutionary complexity, where the ranking would be, in descending order, *lā‘au*, *mau‘u*, *kupukupu*, *limu*. But it does appear that size and/or a rough idea of geometrical complexity may be involved in the ordering. Grasses tend to be rather one-dimensional, giving an overall impression of small size and morphological simplicity. Ferns are larger overall (especially in Hawai‘i), and have a two-dimensional aspect. The other land plants include everything from wildflowers to trees, all three dimensional, and may be expected to average out to intermediate size, and fairly complex apparent morphology. *Limu* too have quite a range in size, since they include some quite large seaweeds and (expanses of) algae. But before accepting the size hypothesis, the possibility of cultural importance should not be overlooked either. Hawai‘ians had no grain foods. A few ferns were used for ferns, but were of very minor importance. The other Angiosperms and many seaweeds, however, were an important part of the diet.

Among the animals, the ranking in terms of number of taxa above the genus level is *holoholona*, land invertebrates, at 53%; *i‘a*, aquatic, invertebrates, at 28%, followed by aquatic vertebrates at 22%; then *manu*, birds, at 4%. I omit *holoholona* vertebrates, since there are only three of them (dog, pig, rat), all of which would be of ambiguous ranking. The same ranking obtains if one counts all taxa unambiguously above the species level. The ranking roughly follows notions of evolutionary complexity, with the arthropods and mollusks at the bottom, followed by fish, then birds. It also roughly follows a size ranking, although of course many fish are larger than birds. Finally, it also follows a general cultural importance ranking, with aquatic invertebrates (crabs, mollusks, squid, etc.) forming a more important diet element than most land arthropods (lice, etc.), and both of them being less important than fish. Birds were not too significant as a food source, but several species were highly prized for their feathers.

It would seem therefore that Berlin's observation may well be right, that inherent properties may correlate to Linnaean rank level. If that is so, then the case of the Hawai‘ian plants indicates that it is a matter of size or gross geometric complexity than one of morphological complexity as Western scientists conceive it. However, the data here are insufficient to distinguish this hypothesis from the idea that cultural relevance (or overall familiarity) that is the correlating factor.

Finally, Table 2 and the graphs in the appendix give some information about the relationship between the length of the Hawai‘ian word and the level it names. It will be recalled that Berlin merely says that any sort of name can be applied at the levels of the genus and above, which is what this database consists of. (His only prediction is that subgeneric names will normally consist of basic name plus qualifier, which I have already informally corroborated above.) It would be interesting to know however whether there is any correlation between the complexity of a word and the Linnaean rank that the native genus it names corresponds to. Three measures of word length were introduced as a rough measure of complexity.

The reader should be cautioned that Pearson correlations rarely mean quite as much as they seem to imply. A sharp spike in an otherwise random graph can easily result in very high correlation coefficients. In an experiment like this, one should probably take seriously only correlations at 0.8 or above. One should also keep in mind that the high number of “ambiguous” entries removes a lot of confidence in the “genus” and “below genus” columns. Nevertheless there does appear to be one clear result: the shorter the word, the more likely it is to name something above the Linnaean genus rank. The best measure for word length is the number of syllables, which gives virtually a straight-line correlation. It is possible that this correlation is actually due to morphological complexity; perhaps derived forms or compounds are used less often for terms at higher ranks.

Whatever the linguistic basis, this is a strong effect not directly predicted by Berlin's theory. Nor can his theory account for the effect, since there are no grounds for distinguishing different types of native genera such that one group of them could be assigned shorter or longer words. But I believe that this finding supports his general approach, which holds that all peoples will tend to classify the natural world in quite similar ways. This correlation suggests that even when a folk genus is not at the equivalent Linnaean rank, the taxon nevertheless must be perceived to have some distinguishing characteristic that signals that it is more inclusive than other folk genera, and on which basis shorter words can be assigned. Thus the alignment between ranks in the different systems is closer than would appear from applying the single criterion of whether the taxon is the narrowest one that has a primary name.

Although this study has turned up some intriguing statistics, I should emphasize that this was a rough and ready survey based on secondary sources. There are many further steps that could be taken if more time and resources become available. One could go further in determining the complete Linnaean classification of biota, so as to see whether families are monotypic and at what particular suprageneric rank the word size effect is strongest. This would also allow one to study whether the Hawai'ian taxa actually correspond to particular Linnaean taxa at all, or rather consist of bits and pieces from different Linnaean taxa. The words could be analyzed morphologically, to see whether it is morphological or phonological complexity that accounts for the word size effect. Field work might be able to help with determining whether the "ambiguous" taxa are in fact monotypic, and if so, in what sense: that Hawai'ians only know about the one species, or that they use the name for the species as a prototypical name for all biota in the genus. Most importantly, it would be useful to find out for sure which terms can subsume other terms. More work should also be done with the synonyms. Certainly many terms are restricted to use on particular islands, but the multiplicity of synonyms leads one to fear that many of them may be referring to different taxa, perhaps at different levels. Finally, it would be interesting to replicate this experiment with other cultures, to determine which of these findings are universal and which may be attributable to the special ecology and culture of the Hawai'ian Islands.

References

Abbott, Isabella Aiona and Eleanor Horswill Williamson. 1974 *Limu: an ethnobotanical study of some edible Hawaiian seaweeds.* Lawai, Hawaii:

Pacific Tropical Botanical Garden. Non vidi.

Berger, Andrew J. 1981. *Hawaiian birdlife.* 2nd ed. Honolulu: University Press of Hawaii. Non vidi.

Berlin, Brent. 1992. *Ethnobiological classification.* Princeton, N.J.: Princeton University Press.

Degener, Otto. 1945. *Plants of the Hawai'i National Park.* Ann Arbor: Edwards Bros. Non vidi.

Gosline, William A. and Vernon E. Brock. 1960. *Handbook of Hawaiian fishes.* Honolulu: University of Hawaii Press. Non vidi.

Handy, E. S. Craighill. 1940. *The Hawaiian planter: his plants, methods and areas of cultivation.* (Bernice P. Bishop Museum Bulletin 161). Honolulu. Non vidi.

Handy, Elizabeth Green. 1958. Ka-'u, Hawai'i, in ecological and historical perspective. E.S. Craighill and Mary Kawena Pukui, *The Polynesian family system in Ka-'u, Hawai'i*, 207-252.

Hillebrand, W.F. 1888. *Flora of the Hawaiian Islands.* Heidelberg. Non vidi.

Hubbard, Douglass H. 1952. *Ferns of Hawai'i National Parl.* Non vidi.

Kay, E. Alison. 1979. *Hawaiian marine shells, reef and shore fauna of Hawaii.* Section 4, Mollusca. (Special publication 64.) Honolulu: Bishop Museum Press. Non vidi.

Krauss, Beatrice H. 1974 Ethnobotany of Hawaii. Honolulu: Dept. of Botany, University of Hawaii. Non vidi.

McKeown, Sean. 1978. *Hawaiian reptiles and amphibians.* Honolulu: Oriental Publishing Co. Non vidi.

Mitchell, Donald D. Kilolani. 1992. *Resource units in Hawaiian culture.* 4th ed. Honolulu: Kamehameha Schools Press.

Neal, Marie C. 1965. *In gardens of Hawaii.* New and rev. ed. (Special publication 50.) Honolulu: Bishop Museum Press. Non vidi.

Pukui, Mary Kawena and Samuel H. Elbert. 1986. *Hawaiian dictionary.* Rev. and enlarged ed. Honolulu: University of Hawaii Press.

Pyle, Robert L. 1977. Preliminary list of the birds of Hawaii. *Elepaio* 37:110-121. Non vidi.

Reed, Minnie. 1907. *The economic seaweeds of Hawaii.* Washington: US GPO. Non vidi.

Rock, Joseph F. 1913a. *List of Hawaiian names of plants.* (Botanical bulletin, 2). Honolulu: Territory of Hawaii Board of Agriculture and Forestry. Non vidi.

Rock, Joseph F. 1913b. *The indigenous trees of the Hawaiian Islands.* Honolulu. Non vidi.

St. John, Harold. 1975a. The variability of the Hawaiian *maile. Phytologia* 32:377-386. Non vidi.

St. John, Harold. 1975b. *List and summary of the flowering plants in the Hawaiian islands.* (Memoir, 1). Lawai, Hawaii: Pacific Tropical Botanical Garden. Non vidi.

St. John, Harold. 1982. *Vernacular plant names used on Niʻihau Island.* (Hawaiian plant studies, 69. Occasional papers, 25, 3.) Honolulu: Bishop Museum. Non vidi.

Summers, Catherine C. 1964. *Hawaiian fishponds.* (Special publication, 52). Honolulu: Bishop Museum. Non vidi.

The Taro collection. 1970. Honolulu: Lyon Arboretum, University of Hawaii. Non vidi.

Titcomb, Margaret. 1948. Kava in Hawaii. *Journal of the Polynesian Society* 57:105-171. Non vidi.

Titcomb, Margaret. 1972. *Native use of fish in Hawaii.* Honolulu: University Press of Hawaii. Non vidi.

Whitney, Leo D., et al. 1939. *Taro varieties in Hawaii.* (Bulletin, 84). Honolulu: University of Hawaii Agricultural Experiment Station. Non vidi.