How does one know that two languages are related? Or that two language families are related? Every linguist purports to know the answers to these questions, but the answers vary surprisingly from one linguist to another. And the divergence of views concerning what is actually known is even greater than that exhibited on the question of how one arrives at this body of information. This is not a particularly satisfactory state of affairs. In what follows we will explore these questions in a global context. We conclude that, despite the generally antipathetic or agnostic stance of most linguists, the case for monogenesis of extant (and attested extinct) languages is quite strong. We will present evidence that we feel can only be explained genetically (i.e. as
the result of common origin), but we will also attempt to answer some of the criticism that has been leveled at work such as ours for over a century.

14. Global Etymologies

The result of common origin), but we will also attempt to answer some of the criticism that has been leveled at work such as ours for over a century.

THE BASIS OF LINGUISTIC TAXONOMY

That ordinary words form the basis of linguistic taxonomy is a direct consequence of the fundamental property of human language, the arbitrary relationship between sound and meaning. Since all sequences of sounds are equally well suited to represent any meaning, there is no tendency or predisposition for certain sounds or sound sequences to be associated with certain meanings (leaving aside onomatopoeia, which in any event is irrelevant for classification). In classifying languages genetically we seek, among the available lexical and grammatical formatives, similarities that involve both sound and meaning. Typological similarities, involving sound alone or meaning alone, do not yield reliable results.

The fundamental principles of taxonomy are not specific to linguistics, but are, rather, as applicable in fields as disparate as molecular biology, botany, ethnology, and astronomy. When one identifies similarities among molecular structures, plants, human societies, or stars, the origin of such similarities can be explained only by one of three mechanisms: (1) common origin, (2) borrowing, or (3) convergence. To demonstrate that two languages (or language families) are related, it is thus sufficient to show that their shared similarities are not the result of either borrowing or convergence. As regards convergence—the manifestation of motivated or accidental resemblances—linguists are in a more favorable situation than are biologists. In biology, convergence may be accidental, but is more often motivated by the environment; it is not by accident that bats resemble birds, or that dolphins resemble fish. In linguistics, by contrast, where the sound/meaning association is arbitrary, convergence is always accidental.

It is seldom emphasized that similarities between language families are themselves susceptible to the same three explanations. That we so seldom see mention of this corollary principle is largely because twentieth-century historical linguistics has been laboring under the delusion that language families like Indo-European share no cognates with other families, thus offering nothing to compare. At this level, it is alleged, similarities simply do not exist.

What is striking is that this position—for which considerable evidence to the contrary existed already at the start of this century (Trombetti 1905) and which on a priori grounds seems most unlikely (Ruhlen 1988a)—came to be almost universally accepted by linguists, most of whom have never investigated the question themselves. Those few scholars who have actually investigated the question, such as Trombetti (1905), Swadesh (1960), and Greenberg
14. Global Etymologies

(1987), have tended to favor monogenesis of extant languages. Even Edward Sapir, often considered an exemplar of linguistic sobriety (despite his alleged excesses in the Americas), looked favorably upon the work of Trombetti, as seen in a letter to Kroeber in 1924: “There is much excellent material and good sense in Trombetti in spite of his being a frenzied monogenist. I am not so sure that his standpoint is less sound than the usual ‘conservative’ one” (quoted in Golla 1984: 420). We maintain that a comparison of the world’s language families without preconception reveals numerous widespread elements that can only be reasonably explained as the result of common origin.

BORROWING

Linguists employ a number of well-known techniques to distinguish borrowed words from inherited items. Most important, clearly, is the fact that basic vocabulary, as defined by Dolgopolsky (1964) and others, is highly resistant to borrowing. Though it is no doubt true that any word may on occasion be borrowed by one language from another, it is equally true that such basic items as pronouns and body parts are rarely borrowed. Furthermore, borrowing takes place between two languages, at a particular time and place, not between language families, across broad expanses of time and place. Thus to attribute the global similarities we document here to borrowing would be ludicrous. And as regards the alleged cases of mass borrowing in the Americas (the so-called “Pan-Americanisms”), Greenberg (1990: 11) quite rightly protests “that basic words and pronouns could be borrowed from Tierra del Fuego to British Columbia . . . is so utterly improbable that it hardly needs discussion.” It seems to us even less likely that basic vocabulary—the grist for most of the etymologies we offer herein—could have been borrowed from one language to another all the way from Africa across Eurasia to South America.

CONVERGENCE

A common criticism of work like ours is that, with around 5,000 languages to choose from, it cannot be too hard to find a word in some African language that is semantically and phonologically similar to, or even identical with, some word in an American Indian language.¹ There are so many possibilities, runs this argument, that one can hardly fail to find accidental “look-alikes” everywhere (Goddard 1979, Campbell 1988). But this sort of mindless search is exactly the reverse of how the comparative method proceeds. The units we are comparing are language families, not individual languages (a language isolate like Basque has traditionally been considered, taxonomically,

¹ For a more fundamental discussion of convergence, see Chapter 2.
a family consisting of a single language). Specifically, we will be comparing items in the following 32 taxa, each of which we believe is a genetically valid group at some level of the classification: Khoisan, Niger-Congo, Kordofanian, Nilo-Saharan, Afro-Asiatic, Kartvelian, Indo-European, Uralic, Dravidian, Turkic, Mongolian, Tungus, Korean, Japanese-Ryukyuan, Ainu, Gilyak, Chukchi-Kamchatkan, Eskimo-Aleut, Caucasian, Basque, Burushaski, Yeniseian, Sino-Tibetan, Na-Dene, Indo-Pacific, Australian, Nahali, Austroasiatic, Miao-Yao, Daic (= Kadai), Austronesian, and Amerind.

One may legitimately wonder why, for the most part, we are comparing relatively low-level families like Indo-European and Sino-Tibetan rather than higher-level taxa like Eurasiatic/Nostratic and Dene-Caucasian, especially since both of us support the validity of these higher-level families (Bengtson 1991a,b, Ruhlen 1990a). We do this to emphasize that higher-level groupings do not require the prior working out of all the intermediate nodes, contrary to the opinion of most Amerindian specialists (the field is all but bereft of generalists!). As is well known, both Indo-European and Austronesian were recognized as families from the early years of their investigation, long before specialists had reconstructed all their intermediate levels (a task that is, of course, still incomplete). In taxonomy it is a commonplace that higher-level groupings are often more obvious—and easier to demonstrate—than are lower-level nodes. We maintain that this is particularly so when one considers the entire world. Current contrary opinion notwithstanding, it is really fairly simple to show that all the world’s language families are related, as we shall see in the etymologies that follow. Discovering the correct intermediate groupings of the tree—the subgrouping of the entire human family—is a much more difficult task, and one that has only begun. Exactly the same is true of Amerind, which itself is a well-defined taxon (Greenberg 1987, Ruhlen 1991a); the subgrouping within Amerind involves far more difficult analyses and taxonomic decisions (Ruhlen 1991c).

Each of our 32 genetic groups is defined by a set of etymologies that connects grammatical and lexical items presumed to be cognate within that group; the postulated membership and putative subgrouping within each of these groups is given in Ruhlen (1987a). The precise number of etymologies defining each of the 32 groups ranges from several thousand (for close-knit and/or well-documented groups like Dravidian or Indo-European) to several dozen (for ancient and/or poorly studied groups like Indo-Pacific or Australian). For the most part the many etymologies defining each group have been discovered independently, by different scholars. (In this regard Greenberg’s work—in Africa, New Guinea, and the Americas—represents an exception to the rule.) So instead of drawing our etymologies from thousands of languages, each containing thousands of words, we are, rather, limited to less
than three-dozen families, some of which have no more than a few hundred identifiable cognates. The pool of possibilities is thus greatly reduced, and accidental look-alikes will be few.

We believe that the failure of our critics to appreciate the truly minuscule probability of accidental similarities is the chief impediment to their understanding why all the world’s languages must derive from a common origin. Accordingly, let us consider this question in some detail. Each of the etymologies we cite involves at least a half-dozen of the 32 supposedly independent families, precisely because the probability of finding the same accidental resemblance in six different families is close to zero. The multiplication of the (im)probabilities of accidental resemblance, as more and more families are considered, quickly assures the attentive taxonomist that similarities shared by numerous families, often separated by vast distances, cannot be due to chance. This crucial point has been emphasized by Collinder (1949), Greenberg (1957, 1963, 1987), and Dolgopolsky (1964), among others, but even Trombetti (1905) was well aware of the statistical importance of attestation in multiple families, rather than in just two. The biologist Richard Dawkins (1987: 274) makes the same point: “Convergent evolution is really a special kind of coincidence. The thing about coincidences is that, even if they happen once, they are far less likely to happen twice. And even less likely to happen three times. By taking more and more separate protein molecules, we can all but eliminate coincidence.”

To see just how unlikely accidental look-alikes really are, let us consider two languages that each have just seven consonants and three vowels:

```
  p  t  k  i  u
  s  a
  m  n  l
```

With a few notable exceptions the vast majority of the world’s languages show at least these phonological distinctions. Yet even this minimal inventory is capable of producing 147 CVC roots, as shown in Table 5. The probability of accidental phonological identity is only 1/147, though the probability of accidental phonological resemblance might be 2/147, 3/147, etc., depending on how many other phonological shapes in Table 5 are deemed sufficiently similar. A perusal of Table 5 suggests, however, that most of these putative roots are quite distinct phonologically and are not readily connected by common phonological processes.
TABLE 5 Possible CVC Roots for a Language with Seven Consonants and Three Vowels

<table>
<thead>
<tr>
<th></th>
<th>KAK</th>
<th>LAK</th>
<th>MAK</th>
<th>NAK</th>
<th>PAK</th>
<th>SAK</th>
<th>TAK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KAL</td>
<td>LAL</td>
<td>MAL</td>
<td>NAL</td>
<td>PAL</td>
<td>SAL</td>
<td>TAL</td>
</tr>
<tr>
<td></td>
<td>KAM</td>
<td>LAM</td>
<td>MAM</td>
<td>NAM</td>
<td>PAM</td>
<td>SAM</td>
<td>TAM</td>
</tr>
<tr>
<td></td>
<td>KAN</td>
<td>LAN</td>
<td>MAN</td>
<td>NAN</td>
<td>PAN</td>
<td>SAN</td>
<td>TAN</td>
</tr>
<tr>
<td></td>
<td>KAP</td>
<td>LAP</td>
<td>MAP</td>
<td>NAP</td>
<td>PAP</td>
<td>SAP</td>
<td>TAP</td>
</tr>
<tr>
<td></td>
<td>KAS</td>
<td>LAS</td>
<td>MAS</td>
<td>NAS</td>
<td>PAS</td>
<td>SAS</td>
<td>TAS</td>
</tr>
<tr>
<td></td>
<td>KAT</td>
<td>LAT</td>
<td>MAT</td>
<td>NAT</td>
<td>PAT</td>
<td>SAT</td>
<td>TAT</td>
</tr>
<tr>
<td></td>
<td>KIK</td>
<td>LIK</td>
<td>MIK</td>
<td>NIK</td>
<td>PIK</td>
<td>SIK</td>
<td>TIK</td>
</tr>
<tr>
<td></td>
<td>KIL</td>
<td>LIL</td>
<td>MIL</td>
<td>NIL</td>
<td>PIL</td>
<td>SIL</td>
<td>TIL</td>
</tr>
<tr>
<td></td>
<td>KIM</td>
<td>LIM</td>
<td>MIM</td>
<td>NIM</td>
<td>PIM</td>
<td>SIM</td>
<td>TIM</td>
</tr>
<tr>
<td></td>
<td>KIN</td>
<td>LIN</td>
<td>MIN</td>
<td>NIN</td>
<td>PIN</td>
<td>SIN</td>
<td>TIN</td>
</tr>
<tr>
<td></td>
<td>KIP</td>
<td>LIP</td>
<td>MIP</td>
<td>NIP</td>
<td>PIP</td>
<td>SIP</td>
<td>TIP</td>
</tr>
<tr>
<td></td>
<td>KIS</td>
<td>LIS</td>
<td>MIS</td>
<td>NIS</td>
<td>PIS</td>
<td>SIS</td>
<td>TIS</td>
</tr>
<tr>
<td></td>
<td>KIT</td>
<td>LIT</td>
<td>MIT</td>
<td>NIT</td>
<td>PIT</td>
<td>SIT</td>
<td>TIT</td>
</tr>
<tr>
<td></td>
<td>KUK</td>
<td>LUK</td>
<td>MUK</td>
<td>NUK</td>
<td>PUK</td>
<td>SUK</td>
<td>TUK</td>
</tr>
<tr>
<td></td>
<td>KUL</td>
<td>LUL</td>
<td>MUL</td>
<td>NUL</td>
<td>PUL</td>
<td>SUL</td>
<td>TUL</td>
</tr>
<tr>
<td></td>
<td>KUM</td>
<td>LUM</td>
<td>MUM</td>
<td>NUM</td>
<td>PUM</td>
<td>SUM</td>
<td>TUM</td>
</tr>
<tr>
<td></td>
<td>KUN</td>
<td>LUN</td>
<td>MUN</td>
<td>NUN</td>
<td>PUN</td>
<td>SUN</td>
<td>TUN</td>
</tr>
<tr>
<td></td>
<td>KUP</td>
<td>LUP</td>
<td>MUP</td>
<td>NUP</td>
<td>PUP</td>
<td>SUP</td>
<td>TUP</td>
</tr>
<tr>
<td></td>
<td>KUS</td>
<td>LUS</td>
<td>MUS</td>
<td>NUS</td>
<td>PUS</td>
<td>SUS</td>
<td>TUS</td>
</tr>
<tr>
<td></td>
<td>KUT</td>
<td>LUT</td>
<td>MUT</td>
<td>NUT</td>
<td>PUT</td>
<td>SUT</td>
<td>TUT</td>
</tr>
</tbody>
</table>

Now were we to compare two languages with a more typical phonemic inventory, say, fourteen consonants and five vowels,

\[
\begin{array}{ccccccc}
p & t & k & j & w & i & u \\
b & d & g & e & o \\
\ddot{c} & a \\
m & n \\
l & r \\
\end{array}
\]

we would find that the number of possible CVC roots in each language jumps to 980. Again, of course, the probability of chance resemblance will depend on certain phonological assumptions, but precious few accidental identities or resemblances, vis-à-vis the stock of some other language or group of languages, could be expected.

One may appreciate just how unlikely an explanation of chance resemblance—indeed development in each family—really is by considering the prob-
ability that the resemblances noted in etymology 21 (below) arose by convergence. We have chosen this etymology for our argument because the meaning involved is rarely borrowed and has no onomatopoeic connections. It thus offers a clear case, where the similarities must be due either to common origin or to accidental convergence. Let us try to calculate the probability that these similarities arose independently. To do this we must make certain assumptions, and at each such stage we shall adopt a minimalist approach that in fact underestimates the true probability. Let us assume, as we did above, that each language family uses only seven consonants and three vowels, yielding the 147 syllable types shown in Table 5. What, then, is the probability that two languages will accidentally match for a particular semantic/phonological domain, in the present case ‘female genitalia’? Clearly it is 1/147 or .007. Whatever the form that appears in the first language family, the second family has only one chance in 147 of matching it. And the probability that a third family will offer a match will be (1/147)² or .000049; that of a fourth family, (1/147)³ or .0000003; and so forth. In the etymology we give, 14 of the 32 taxa show apparent cognates, though the evidence is for the moment slim in Australian and the vowel in Austronesian (and many Amerind forms) is e rather than the expected u. But if we ignore these details, then the probability that the particular sound/meaning correlation “PUT/female genitalia” arose independently fourteen times will be (1/147)¹³, or about one chance in ten octillion, by our rough calculations. We feel that this qualifies as a long shot; certainly descent from a common source is the more likely explanation.

The foregoing constitutes what we consider to be the basis of genetic classification in linguistics. The application of these basic principles to the world’s language families leads inevitably, in our opinion, to the conclusion that they all derive from a single source, as suggested by the 27 etymologies presented below. We have not yet dealt, however, with a number of other topics that in the minds of many linguists are inextricably tied up with taxonomy, questions like reconstruction, sound correspondences, and the like. We believe that these topics are not in fact of crucial importance in linguistic taxonomy, and that mixing the basic taxonomic principles with these other factors has led to much of the current confusion that we see concerning the classification of the world’s languages. So that these ancillary topics not be invoked yet again, by those opposed to global comparisons, we will take them up one by one and explain why they are not relevant to our enterprise. Let us begin with a topic that is at the heart of many current disputes, the alleged incompatibility between Greenberg’s method of multilateral comparison and the traditional methods of comparative linguistics.
MULTILATERAL COMPARISON VS. THE TRADITIONAL METHOD

Many linguists feel that Greenberg’s use of what he calls multilateral comparison to classify languages in various parts of the world is incompatible with—or even antagonistic to—the methods of traditional historical linguistics, which emphasize reconstruction and sound correspondences (about which, see below). Thus, Bynon (1977: 271) claims that “the use of basic vocabulary comparison not simply as a preliminary to reconstruction but as a substitute for it is more controversial. . . . Traditional historical linguists . . . have not been slow in pointing out the inaccuracies which are bound to result from a reliance on mere similarity of form assessed intuitively and unsubstantiated by reconstruction.” In a similar vein, Anna Morpurgo Davies (1989: 167) objects that “we do not yet know whether superfamilies outlined in this way have the same properties as families established with the standard comparative method. If they do not, there is a serious risk that the whole concept of superfamily is vacuous.” And Derbyshire and Pullum (1991: 13) find Greenberg’s Amerind hypothesis “startling, to say the least, when judged in terms of the standard methodology . . . .”

The confusion displayed in the previous three quotes (and one could give many others) results from a failure to realize that the comparative method consists essentially of two stages. The first stage is classification, which is really no different from what Greenberg calls multilateral comparison. The second stage, which might be called historical linguistics, involves family-internal questions such as sound correspondences and reconstruction. In practice, there is no name for this second stage simply because the two stages are seldom distinguished in the basic handbooks on historical linguistics, in which, almost without exception, the initial stage, classification, is overlooked (Bynon 1977, Hock 1986, Anttila 1989). Also overlooked in these basic texts are language families other than Indo-European. The origin of this anomaly—which knows no parallel in the biological world—is a consequence of the primogeniture of Indo-European in the pantheon of identified families, and the subsequent elaboration of the family by Europeans in the nineteenth century.

That the initial stage of comparative linguistics, classification, is so systematically overlooked today lies in the origin of the Indo-European concept itself. When Sir William Jones announced in 1786 that Sanskrit, Greek, and Latin—and probably Gothic and Celtic as well—had all “sprung from some common source,” he essentially resolved the first stage of comparative linguistics at the outset: he identified five branches of Indo-European and hypothesized that all five were altered later forms of a single language that no longer existed. What was left unstated in Jones’s historic formulation was the fact that languages such as Arabic, Hebrew, and Turkish—languages that Jones knew well—were
excluded from his Indo-European family.\footnote{The term Indo-European was not introduced until the nineteenth century.} For Indo-European, and for the Indo-Europeanists who came to dominate historical linguistics, the problem of classification was essentially resolved by Jones, and the later additions of a few more obvious branches, such as Tocharian and Anatolian, did not alter this state of affairs.

The problems that Greenberg confronted, however, when he set out to classify the languages of Africa, were quite different from those facing a historical linguist investigating an already-defined family. Greenberg was confronted by over 1,000 languages, only some of which fit into well-defined families (e.g. Semitic, Bantu), and among which there was little understanding of the relationships. Under these circumstances, where does one start? Obviously the only way to begin is by the comparison of basic lexical items and grammatical formatives in all the languages, which inevitably leads to a classification of the languages into a certain number of groups defined by recurring similarities. This is exactly what Jones had done when he identified Indo-European, stressing, as he did, “a stronger affinity, both in the roots of verbs and in the forms of grammar, than could possibly have been produced by accident.” He said nothing of sound correspondences or reconstruction, for in fact these concepts came to prominence (despite the earlier work of Rask, Grimm, and Bopp) only in the second half of the nineteenth century.

We believe, in short, that there is really no conflict between Greenberg’s method of classifying languages and what is often referred to rather inexpressively as “the standard methodology.” The standard methodology is used to investigate family-internal problems; it does not—at least as it is explained in the basic textbooks referred to above—tell one how to identify language families. Accordingly, it does not tell one how to classify the world’s languages. This, rather, is what Greenberg’s work does, and it is, furthermore, how Greenberg views what he does. It has recently been alleged that he himself subscribes to the view that his methods differ from the standard methodology: “Greenberg (1987) makes clear that he believes such groupings [as Altaic, Hakan, and Amerind] cannot be reached by the standard comparative method; a wholly different method, MASS COMPARISON, is required” (Nichols 1990: 477). That this is, in fact, exactly the opposite of Greenberg’s views is shown in the following:

Statements from certain American Indianists that I have rejected comparative linguistics and have invented a new unorthodox method called mass or multilateral comparison are repeated again and again in the press. However, as I clearly stated in Greenberg (1987: 3), once we have a well-established stock I go about comparing and reconstructing just like anyone else, as can be seen in my various contributions to historical linguistics. However, as I pointed out long ago in regard to my generally accepted African classification, the first step has to be to look very broadly,
on at least a continent-wide scale, to see what the obvious groupings are. How can one start to apply the comparative method until one knows what to compare? (Greenberg 1990: 8)

**RECONSTRUCTION**

It is remarkable how frequently reconstruction is confounded with taxonomy. For a moment’s reflection should make it clear that one can only begin reconstructing a proto-language *after* one has decided which languages belong to the putative family. Until one has delineated a set of seemingly related languages, collectively distinct from all others, by the methods outlined at the outset of this chapter, there is simply nothing to reconstruct. (After the fact, of course, reconstruction and (re)classification may enjoy a fruitful feedback.) And as for the supposed validating effect of reconstruction, would anybody claim that a bad reconstruction invalidates a well-defined family such as Indo-European? Or that a brilliant reconstruction could show that Slavic, Ob-Ugric, and Basque form a valid family? As a process, reconstruction is entirely different from taxonomy, and the two should not be confused. It is for this reason that Bynon’s claim that Greenberg uses multilateral comparison as a “substitute” for reconstruction really makes no sense, and it is certainly not anything that Greenberg has ever written or said or even suggested.

**SOUND CORRESPONDENCES**

Perhaps the greatest source of confusion in recent taxonomic debates has been the role that sound correspondences, for example Grimm’s Law, play in classification. It is clear that many historical linguists see regular sound correspondences as playing some crucial role in identifying valid linguistic taxa. In reality, sound correspondences are discovered only *after* a linguistic family has been identified, for the simple reason that sound correspondences are properties of particular linguistic families. They are not—and could not be—a technique for discovering families. When the Indo-European sound correspondences were worked out in the nineteenth century, not for a minute did any of the Indo-Europeanists imagine that they were “proving” Indo-European, the validity of which had not been in doubt for decades.

There are several reasons why sound correspondences have become enmeshed with taxonomic questions. First, it is sometimes alleged that it is only by means of regular sound correspondences that borrowings can be discriminated from true cognates. It has long been recognized, however, that loanwords often obey regular sound correspondences as strictly as do true cognates, a point emphasized on several occasions by Greenberg (1957, 1987). Campbell (1986: 224) makes the same point: “It ought to be noted that such
agreements among sounds frequently recur in a number of borrowed forms, mimicking recurrent sound correspondences of true cognates."

Another alleged use of sound correspondences is to discriminate superficial look-alikes from true cognates (see the quote by Bynon above), and cognates, it is claimed, do not look alike and can only be recognized by means of sound correspondences. Thus, the commonly accepted Indo-European sound correspondences show that Armenian erku ‘2’ and Latin duo ‘2’ are cognate, despite their different form, whereas English bad and Farsi bad are not cognate, despite their identity of form. Campbell has aptly criticized such views:

Identical or very similar sound matchings do not necessarily imply loans or weak evidence of genetic connection. . . . With a time depth approaching that of the Indo-European languages of Europe, the Mayan correspondences are on the whole identical or are the result of single natural and recurrent changes. Proto-Mayan *p, *m, *n, and *y are reflected unchanged, with identical correspondences, in all of the over thirty Mayan languages. All other correspondences are very similar. Even English, after its many changes, reflects Proto-Indo-European *r, *l, *m, *n, *s, *w, and *y unchanged, on the whole.

A quick survey of once-disputed but now established remote genetic relationships reveals that identical (or very similar) sound correspondences are not that unusual....

Therefore, identical correspondences should not be shunned nor too speedily attributed to borrowing. While longer separation may provide greater opportunity for unusual and exotic correspondences to develop in cases of distant genetic relationship, it is in no way necessary for such developments to have taken place nor for correspondences to be non-identical” (1986: 221–23).

Indeed, when one looks at the reconstructions that have been proposed for almost any family, one is able to find modern languages that preserve the proposed ancestral forms virtually unchanged. To cite just a few examples, Proto-Indo-European *népot- ‘nephew, son-in-law’ is strikingly similar to modern Rumanian nepot, and Proto-Indo-European *müs ‘mouse’ was preserved without change in Latin, Old English, and Sanskrit. Proto-Austronesian *sepat ‘2’ is almost identical with Rukai sepat, and Proto-Austronesian *mat’a ‘eye’ is identical with Rukai mat’a. Proto-Uralic *tule ‘fire’ is preserved in Finnish tule-, and Proto-Uralic *morska ‘to wash’ differs little from Estonian mõske. At an even greater time depth, we find that Proto-Nostratic *nato ‘female relation by marriage’ has survived, in Uralic, as Finnish nato ‘husband’s or wife’s sister’ and, in Dravidian, as Malayalam nattūn ‘husband’s sister, brother’s wife,’ while Proto-Nostratic *pʰalV ‘tooth’ survives in Dravidian as Telugu palu and in Altaic as Ulch palu. At a time depth perhaps even greater than that of Nostratic, we find Proto-Australian *bugku ‘knee’ preserved in Dyirbal buŋku.

In the etymologies we present below, connecting all of the world’s language families, the situation is not all that different from that within the families
just discussed. There are, in fact, many examples of sound correspondences of the transparent variety discussed by Campbell. This initial stage of the analysis is necessarily characterized by the identification of easily recognizable similarities, just as was the discovery of Indo-European or any other family. The refinement represented by exotic sound correspondences of the *erku–duo* variety inevitably awaits a later stage in the analysis—the second stage, which we have called “historical linguistics.” And it is important to recognize that the work of this stage leads almost invariably to a refinement of the etymologies, rather than a refinement of the classification.

Among the world’s language families, there are no doubt exotic sound correspondences as well that we have not detected. It should be noted, nevertheless, that as early as 1986 one of us (Bengtson) proposed some global sound correspondences, and the Russian scholar Sergei Starostin (1991) has recently published the most explicit statement of interphyletic sound correspondences to date. His brief table of Nostratic–Dene-Caucasian correspondences, though not quite global in scope, accounts for a vast expanse of the linguistic world. Nostratic, for Starostin, includes ten of our 32 taxa (Kartvelian, Indo-European, Uralic, Dravidian, Turkic, Mongolian, Tungus, Korean, Japanese-Ryukyuan, and Eskimo-Aleut), and Dene-Caucasian, for Starostin, includes Caucasian, Sino-Tibetan, Yeniseian, and Na-Dene—to which one may confidently add both Basque and Burushaski (Bengtson 1991a,b). Thus, Starostin’s equations account for roughly half of our 32 taxa, as well as the vast majority of the Eurasian land mass. We find nothing in Starostin’s correspondences that is inconsistent with the etymologies proposed below.

**ON THE LIMITS OF THE COMPARATIVE METHOD**

It has recently been widely asserted that the comparative method in linguistics produces reliable results only for the past 5,000–10,000 years. According to Kaufman (1990: 23), “A temporal ceiling of 7,000 to 8,000 years is inherent in the methods of comparative linguistic reconstruction. We can recover genetic relationships that are that old, but probably no earlier than that. The methods possibly will be expanded, but for the moment we have to operate within that limit in drawing inferences.” Similar statements from a host of other scholars are given in Chapter 11, where such beliefs are identified as the central myth of historical linguistics (Chapter 13 further analyzes such myths). The origin of this myth, we believe, is an attempt by Indo-Europeanists to “explain” why Indo-European has no known genetic connections—in our view yet another myth. The fact that Indo-European is intimately connected with numerous other families has been demonstrated beyond a reasonable doubt by the Russian Nostraticists (Illich-Svitych 1971–84), a demonstration that is complemented and extended by Greenberg (to appear).
We have shown that in numerous cases sounds (particularly stable ones like nasal consonants and liquids)—and even entire words—have persisted over time spans greater than 8,000 years virtually unchanged. This raises the question why these evidently quite stable sounds must suddenly change beyond recognition, or disappear entirely, beyond the supposedly insuperable threshold of 10,000 years. If we can use modern languages to reconstruct proto-languages that existed at least 6,000–8,000 years ago (e.g. Proto-Indo-European, Proto-Uralic, Proto-Dravidian, Proto-Austronesian), why cannot such earlier languages themselves be compared (as in fact we will do) in order to discern still earlier groupings? Would it not be one of the more remarkable coincidences in the history of science if Indo-European, the family in terms of which comparative linguistics was discovered, turned out to define the temporal limit of comparative linguistics as well? That there is no such coincidence is amply demonstrated in the etymologies we give below. We feel it is time for linguists to stop selling the comparative method short and to apply it consistently to the world’s linguistic taxa, without preconception. The present chapter represents a step in this direction, an initial step that shows that all of the world’s populations are linguistically connected. The culmination of these efforts will be a comprehensive subgrouping of this single linguistic family.

BAD SEMANTICS

Another criticism of global etymologies in particular, and of long-range comparison in general, is that such liberties are taken with semantic change that literally anything can be connected with anything else, and it is certainly true that many global etymologies proposed over the years have been semantically unconvincing. But for just that reason we have constrained the semantic variation of each etymology very tightly, and few of the semantic connections we propose would raise an eyebrow if encountered in any of the standard etymological dictionaries. They are in fact semantically more conservative than many proposed connections in Pokorny (1959), the standard Indo-European etymological dictionary. Whatever damage this often alleged defect may have done to earlier programs of long-range comparison, we believe that it does not affect the etymologies presented below.

ERRORS IN THE DATA

Another often-cited criticism of long-range comparison is the presence of errors in the data, errors that invalidate the overall hypothesis. This is a specious argument, for it ignores both common sense and the standard measures of statistical significance. Genetic classification is not analogous to a mathematical proof, wherein one false step undermines the complete demon-
14. Global Etymologies

rather, the cumulative weight of all the evidence completely swamps the effects of whatever random errors may be scattered through the work. As Greenberg has often stressed—and has in fact shown in his work—multilateral comparison yields valid genetic classifications even from decidedly degenerate data. An example was Greenberg's classification of Australian languages in 1953, using little more than the vocabularies published by E. M. Curr in 1886–87. The notion that data must be pristine and copious flies in the face of commonly accepted historical method. It is all well and good for Kaufman (1990: 18) to demand at least 500 items of basic vocabulary and 100 points of grammar before “serious comparative work” can be carried out, but the fact remains that Indo-Europeanists have classified Lydian as Indo-European, without dissent, on the basis of a handful of words, as noted by Greenberg (1990: 10). Similarly, David Payne (1991: 362) reports that “all that remains of the [Shebayo] language is a vocabulary list of fifteen words collected at the end of the 17th century. . . . Despite the paucity of data from this language, it is quite clear that it is Arawakan.” Historians and historical linguists—not to mention paleontologists working from handfuls of bashed fossils—use whatever material is available; they do not demand that the evidence be complete or immaculate.

DISTRIBUTIONAL DIFFERENCES

It is often alleged that one can find anything in linguistic data if one looks for it hard enough. Thus the global etymologies we present below are a tribute more to our industry and enterprise than to real genetic connections. Such a view is widespread among linguists who have never actually compared large numbers of languages (or language families), but those of us who have done this kind of work know the reverse to be true. “Wanting” to find something is of very little help if it is not there. Greenberg (1987) points out that the Amerind family has two general words for females, TUNA ‘girl’ and KUNA ‘woman.’ Both roots are abundantly attested throughout North and South America, and both are found in all eleven branches of the Amerind family. What is interesting about their distribution, however, is that whereas KUNA is widely attested in the Old World, as we show in etymology 11 below, we have found no trace of TUNA in the Old World. If it were really so easy to find anything one looks for, why did we fail to find TUNA in the roughly 4,500 Old World languages, when it is so readily observed in the approximately 500 New World languages? The evolutionary analysis provides a simple and natural explanation: when the Amerind forebears first entered the New World they brought with them the word KUNA ‘woman,’ and only later did they invent the word TUNA ‘girl.’ That there is no trace of TUNA ‘girl’ in the Old World is because it never existed there.
14. Global Etymologies

GLOBAL ETYMOLOGIES

For each etymology, in what follows, we present a phonetic and semantic gloss, followed by examples from different language families. Though we have not attempted to present a unified phonetic transcription for all sources, we have adjusted certain transcriptions from time to time to avoid potential ambiguity. In the first etymology (but not elsewhere) yod has been normalized to j in all citations. Ejectives have been normalized to p’, t’, k’, etc.; V represents a vowel of indeterminate timbre; ˘ı is used for the Old Church Slavic soft sign and ˘u for the hard sign; and ~ separates alternative forms. In the two interrogative etymologies (10, 17), interrogative and relative uses are not distinguished (‘who?’ as in “Who is that man?” vs. ‘who’ in “The man who came to dinner.”). The intimate connection between the two is well known and uncontroversial. Most of the cited forms are, however, true interrogatives.

The source of the information for each family represented in a given entry is indicated by an abbreviation in brackets at the end of the entry. The number following the abbreviation is either the etymology number in the original source (if there is one) or the page number there. Since the existence of these roots as characteristic features of the language families cited has already been established by other scholars, and is not for the most part in question, we do not give the complete documentation for each family, limiting ourselves in most instances to an indication of the range of semantic and phonological variation within the family. The reader who wishes to see every relevant form for a given family should consult the sources cited. For Amerind, however, we give extensive citations, in order to counterbalance the fallacious criticism that has been directed at Greenberg’s work. Parts of etymologies that are problematic, by dint of either phonetic or semantic divergence, or by restricted distribution, are preceded by a question mark. The lack of a semantic gloss following a form means that that form has the same meaning as the preceding form.

We make no claim to being the first to discover any of the etymologies listed below. The pioneering work of Trombetti, Swadesh, Greenberg, Illich-Svitych, Dolgopolsky, and Starostin has identified numerous widespread roots. What we have tried to do is to make each etymology more complete and more soundly documented in this incarnation than it may have been in previous ones. With this goal in mind we have weeded out certain families from pre-

---

3 We do not deal here with reconstruction, and these glosses are intended merely to characterize the most general meaning and phonological shape of each root. Future work on reconstruction will no doubt discover cases where the most widespread meaning or shape was not original.
vious proposals, where the root was phonologically or semantically too divergent, or too weakly attested, to be convincing. But we believe we have also uncovered some additional etymological connections that had previously gone unnoticed. To a very great extent the recognition of these similarities has been made possible by the lower-level classificatory work of Greenberg in Africa, the New Guinea area, and the Americas, and by that of Russian scholars on Kartvelian, Caucasian, and other families of the former Soviet Union. Before all this work appeared, in recent decades, it was difficult, if not impossible, for a taxonomist to be sure that a root was truly diagnostic of some family, simply because there was no understanding of what the valid genetic families were, much less what cognates defined them. Trombetti, for example, dealt in terms of languages only where he was forced to by a lack of any general overall classification. Wherever possible, he worked with established language families (e.g. Indo-European, Uralic, Bantu), since he was well aware of the unavoidable methodological quandary presented by poorly documented families.

We harbor no illusions, of course, that every etymological connection we propose will be found, ultimately, to be correct, but we do believe that the removal of such errors as may exist in these etymologies will not seriously affect the basic hypothesis, which does not depend on any specific link for its validity. Furthermore, the number of widespread etymologies can be vastly increased over the fragment we present here. In the long run we expect the evidence for monogenesis of extant languages to become so compelling that the question will be not whether all the world’s languages are related, but why it took the linguistic community so long to recognize this obvious fact.

1 AJA ‘mother, older female relative’

Niger-Congo: Tenne -ja ‘mother,’ Bulom ja, Yoruba ija; Bantu: Proto-Bantu *jijà ~ *jijò. [BA IV: 190]
Nilo-Saharan: Saharan: Daza aja ‘mother,’ Kanembu jia ~ ja, Kanuri ja; Fur ija; Maban: Runga ja; Koman: Gumuz ijo; Central Sudanic: Mangbetu aja, Madi ia, Lombe jaija; East Sudanic: Gulfan aja, Midob ija, Suk iju, Nyangiya joijao ‘thy mother.’ [NS 95, CN 67, ES 77, NSD 43]
Afro-Asiatic: Omotic: Wolamo ajë ‘mother’; Cushitic: Oromo ajo, Somali hoojo; Chadic: Kotoko ija ~ ija ~ ja, Mubi jà. [WM 64]

292 14. Global Etymologies
aja ~ ija, Manda ajà, Kui aia ~ aja ~ ija, Kuwi ija ‘mother,’ aja ‘woman,’
Kurux ajan ‘mother,’ Malto aja ‘my mother.’ [D 364, NSD 43]

BURUSHASKI -āi ‘daughter, girl.’ [B 455]

?INDO-PACIFIC: Isahi aijo ‘mother,’ Korafe aja. [FS 99]

NAHALI aji ‘husband’s younger sister.’ [NA 59]

AUSTROASIATIC: Munda: Sora ajàntsà ‘bitch’ (= female-dog, cf. kin-tsàr
‘male dog’); Mon-Khmer: Proto-Mon-Khmer *ja?’grandmother.’ [PB 482,
SB 34]

MIAO-YAO: Proto-Yao *ja ‘father’s sister.’ [PB 339]

DAIC: Tai: Proto-Tai *ja ‘father’s mother’; Sek ja; Kam-Sui: Proto-Kam-
Sui *ja ‘grandmother,’ Sui ja ‘grandmother, old woman’; Li: Proto-Li *ja
‘mother, grandmother,’ Small Cloth Loi ja ‘mother’; Lakkia ja ‘grand-
mother.’ [PB 339]

AUSTRONESIAN: Proto-Austronesian *ajah ‘father,’ Atayal jaja? ‘mother,’
Pazeh jah ‘older sister,’ Malay ajah ‘father,’ Javanese (j)ajah ‘father.’
[AN 13, WW 74, PB 339]

AMERIND: Penutian: North Sahaptin àjaD ‘woman,’ Nez Perce ?ajat, Tzotzil
jaja ‘grandmother’; Hokan: Washo -ja ‘paternal aunt,’ Quinigua ñjaak,
Tonkawa ñejan ‘woman’s sister’; Central Amerind: Tewa jia ‘mother,’
Proto-Oto-Manguean *ja ‘female,’ Proto-Uto-Aztecan *je ‘mother,’ Tara-
humara ije, Yaqui ñaije, Nahua -je? Chibchan-Paezan: Xinca ajà ‘woman,’
Matagalpa joaja, Cuna jaa-kwa ‘young woman,’ Colorado ajà ‘mother’;
Andean: Ono ñoj ‘grandmother,’ Auca -ñeje: Macro-Tucanoan: Amaguaje
ajo ‘old woman,’ Masaka jaja ‘older sister,’ Ticuna jake ‘old woman’;
Equatorial: Mapidiana ajà ‘aunt,’ Tora ije ‘paternal grandmother,’ Arikem
hoja ‘aunt’; Macro-Panoan: Mayoruna jaja, Shipibo jaja ‘paternal aunt,’
Moseten eje ‘grandmother,’ jaja ‘mother-in-law’; Macro-Carib: Accawai
aja ‘mother’; Macro-Ge: Coropo ajàn, Cororado ajan, Palmas já. [CA 55,
AMN]

2 BU(N)KA ‘knee, to bend’

NIGER-CONGO: Baga -buŋ ‘knee,’ Pajade -paŋ, ?Lefana -ŋko; Bantu: Proto-
Bantu *ɓʊŋgɔ, Swahili bong’oa ‘to stoop, bend down.’ [BA III: 57]

KORDOFANIAN: Tegele mbo ‘knee’ (pl. abɔan ~ abuŋ). [VB]

AFRO-ASIATIC: Omotic: Dime boq ‘knee,’ Bako boŋa, Basketo buŋa, Oyde
bunke; Chadic: Fyer fuŋ ‘knee,’ Bura hunji ‘knee.’ [VB]

INDO-EUROPEAN: Proto-Indo-European *bhugen(h) ‘to bend’; Indic: Sanskrit
bhugna ‘bent’; Germanic: Gothic biugan ‘to bend,’ Old Icelandic bogenn
‘bent,’ English bow, elbow; Celtic: Proto-Celtic *buggo ‘flexible, mal-
leable,’ Irish bog ‘soft’; Albanian butë (< bhugen(h)-to) ‘soft’; Baltic: Lat-
vian baugurs ‘hill, rising ground.’ [IE 152, N 25]
14. Global Etymologies

**Turkic:** Proto-Turkic *bük(ā)* ~ *bök(ā)* ‘to bend,’ Chuvash pēk ~ pōk ‘to bend,’ Yakut bük, Khakas bükri ‘bent,’ Old Uighur bük ~ bök ‘to twist,’ Uighur bük ~ bök ‘to kneel.’ [N 25]

**Mongolian:** Proto-Mongolian *böke* ‘to bend,’ Written Mongolian bökeji ∼ bökeji ‘to cave in, sag’ böken ‘hump of a camel,’ bökötür ‘bent,’ Khalkha bóx(ōn) ‘hump of a camel,’ Kalmyk bökñ ‘hump, humped.’ [N 25]

**Tungus:** Proto-Tungus *bök(ā)* ‘to bend,’ Nanai bukun ‘hump,’ Evenki bükä ‘to bow,’ buku ‘bent, crooked,’ bękä ‘hump.’ [N 25]

**Ainu** he-poki-ki ‘bow down,’ he-poki-poki ‘to nod the head.’

**Indo-Pacific:** Halmahera: Tobelo buku ‘knee,’ Modole bubuqu, Loda wuwu-ku; Bougainville: Koianu poku; South New Guinea: Teri Kawalsch bugu; Northeast New Guinea: Saker habbakan. [IP 43]

**Australian:** Proto-Australian *puŋku* ‘knee,’ Tyeraity bōngōl, Maranunggu bingar, Guugu Yimidhir bugu, Kok-Nar poŋk ∼ puŋkuwāl, Gugu-Badhun buguyal, Kukatj poŋkipal, Dyirbal bugu ‘knee, bend in the river, wave,’ Yidiny bugu ‘knee, hump in a snake’s body.’ [NP 232, RD 110, 123, 223]

14. Global Etymologies

3 BUR ‘ashes, dust’

NILO-SAHARAN: Songhai: Gao bomni ‘ashes,’ Djerma boron; Berta bub(u)da;
Central Sudanic: Bongo buru-ku, Keliko ḫaro; East Sudanic: Kenzi,
Birgid u-burti, Murle bûr, Mursi bûr, Balé bûr, Shilluk, Bor bûr, Lango
bûr, Auhr bûr. [NS 9, CN 7, ES 5, NSD 6, NSB]

AFRO-ASIATIC: Proto-Afro-Asiatic *b(w)H ‘loose soil, sand, dust’; Semitic:
Proto-Semitic *br ‘dust,’ Arabic baraj ‘dust, soil,’ Mehri berô ‘sandy
seashore,’ Classical Hebrew bar ‘field, open space,’ Proto-Semitic *bwr
‘(fallow) ground,’ Arabic bawr ‘fallow ground,’ Syriac bûrô, Akkadian
bûr ‘open space’; ?Berber: Shilha tamurt ‘soil’; Cushitic: Beja bûr ‘soil,’
Bilin bûr ‘soil,’ Saho bâro, Afar bâr, Somali berri; Chadic: Proto-Chadic
*bwr ‘sand,’ Angas bûr ‘sand, dust,’ Logone bûrû, Gider burduku ‘soil.’
[CS 398, N 22]

KARTVELIAN: Svan burγ ‘to raise dust,’ birγ (< *burγ-i) ‘dust, ashes,’
?Middle Georgian bre. [N 22]

URALIC: Proto-Uralic (Illich-Svitych) *porV ‘dust, sand, dirt’; Samoyed: Kamassian pûre ‘sand’; Ugric: Ostyak per ‘ashes’; Finnic: Finnish poro ‘hot
ashes, course dust,’ Estonian pori ‘mud,’ Mansi pors ‘sweepings.’ [U 68,
N 22]

DRAVIDIAN: Proto-Dravidian *pûrV ~ *porV ‘loose soil, sand, dust,’ Malto
parsi ‘sweepings,’ Naïkri bûrdi ‘ash,’ Telugu bûdida ‘ashes,’ Tulu poyyê
‘sand,’ Malayalam purûti ‘dust, earth,’ pûyî ‘sand,’ Tamil purûti ‘dust,
dry earth,’ pûř ‘powder, dust.’ [D 4316, N 22, NSD 6]

TURKIC: Proto-Turkic *bûr, Chuvash pur(a) ‘chalk,’ Tuva pur ‘clay,’ Jakut
buor ‘soil, clay, dust,’ Altai pur ‘ashes,’ Uighur bor, Kazakh bor ‘chalk.’
[N 22]

MONGOLIAN: Khalkha bur ‘dirty, muddy, dark,’ Buriat bur ‘silt, swamp,
clay.’ [N 22]

TUNGUS: Manchu buraki ‘dust, sand,’ Nanai burâxín ‘dust,’ Oroch burâxi.
[N 22]

?ESKIMO-ALEUT: Proto-Eskimo-Aleut *pujV ~ *apju ‘dust, mud, soot.’ [EA]

BIRUSHASKI bur-di ‘the ground.’

?INDO-PACIFIC: Tasmanian bûrana ‘smoke.’

AUSTRALIAN: Proto-Australian *burin ~ *burinj ‘smoke.’ [AC 75]

AMERIND: Chibchan-Paezan: Cuna piru ‘ashes,’ Uncasica bura, Manare okâ-
obra, Move ūio-bru, Guatuso purun, Catio pora ‘dust’; Andean: Lupaca purika ‘ashes’; Equatorial: Shuarapupuur ‘dust,’ Bare bûridi ‘ashes,’
Wapishana parati, Goajiro purpura ‘dust’; Macro-Panoan: Taruma gula-
paru ‘powder’; Macro-Carib: Yagua papândru ‘ashes’; Macro-Ge: Proto-
Ge *pra ‘ashes,’ Krenje pro, Cayapo pra ‘embers,’ Guato (ma-)fora(-ta)
‘ashes,’ Caraja brûbi. [AM 11, AMN]
14. Global Etymologies

4 ČUN(G)A ‘nose; to smell’

KHOISAN: #Au.//ei ěù ‘nose,’ !Kung tʰû ~ sū; !O-!Kung tûn ~ čû; G//abake čui, Naron sô ‘to snuff,’ Nama suni ‘sniff, smell from’; /Xam sū ‘snore,’ //N-!e sùwa ‘blow the nose,’ Kakia /mûha čûni ‘nostrils.’ [K 89, SAK 488, 489, HF 1:10]

Nilo-Saharan: Saharan: Zagawa sîna ‘nose,’ Berti sâno; East Sudanic: Mei-dob i-siŋi ‘nose,’ Ongamo (a-ta-)sîŋa ‘to sneeze,’ !O-!Kung tûsû ‘sniff, smell from;’ /Xam sû ‘snore,’ /Ng-i suna ‘blow the nose,’ Kakia /nuha čuni ‘nostrils.’ [NSB, KER, HF 12]

AFRO-ASIATIC: Proto-Afro-Asiatic *t(˘w)n ~ *t(˘j)n ‘smell;’ Ancient Egyptian sn ‘to smell,’ sn sn ‘to breathe;’ Omotic: Proto-Omotic *sîn-t ‘nose,’ Bas-keto sînc.a, Chara sind. ¯a, Gimira sînt, Mao sînt; Cushitic: Burgi suna, Konso sôna, Somali san, Kaffa čînno ‘odor,’ Saho sôn ‘to smell;’ Chadic: Hausa sûnsûnû ‘to sneeze,’ Bachama sîne, Bata čînne, Klesem siŋ, Bana čîn, Dari sn. [AA 54, N 51, HF 1:10]

KARTVELIAN: Georgian sun ‘odor, to smell.’ [N 51]


BASQUE su-ður ‘nose,’ sun-da ‘smell.’

BURUSHASKI šû (etas) ~ sû (etas) ‘to smell.’ [B 335]

SINO-TIBETAN: Karen suŋ ‘odor;’ Tibeto-Burman: Proto-Tibeto-Burman *Sîna ‘nose;’ *suŋ ‘smell,’ Tibetan sna, bsuŋ ‘smell (sweet),’ Nung sôna, Janggali sîna, Digaro hona(-gam), Burmese hna, sàng ‘emit a pleasant odor.’ [ST 101, 405, HF 1:10]

14. Global Etymologies

INDO-PACIFIC: Baruya sína ‘nose,’ Musak sinami ‘to smell.’ [FS 105]
NAHAI 獐 ‘nose.’ [NA]
MIAO-YAO: Miao tšinyu ‘nose,’ Yao (pu-)&sN. [HF 1:10]
DAIC: Ong Be zoŋ ‘nose,’ suŋ (uu) ‘to blow the nose’ (= to-blow [nose-mucus]). [PB 345]
AUSTRONESIAN: Proto-Austronesian *įgųŋ *ųgųŋ ‘nose,’ Kuvalan uguŋ, Proto-Philippine *suŋD, Proto-Oceanic *isu(ŋ) ~ *untsu(ŋ), Fijian utsu, Proto-Polynesian *isu; Proto-Austronesian *t’uŋa] ~ *suNar ‘to sneer, turn up the nose.’ [AN 67, 158; WW 139, 227; PB 345]

5 KAMA ‘hold (in the hand)’

KHOISAN: Zhu xìn(–xí) ‘rub back and forth’; Kxoe xôm ‘crush with the hands.’ [SAK 852]
NIGER-CONGO: Dagomba kam ‘squeeze,’ Nupe kā, Proto-Bantu *kama, Swahili kama ‘to squeeze, to milk,’ kama-ta ‘to take hold, seize, grasp.’ [KS 59, BA III: 263]
?NILO-SAHARAN: Songhai kaŋkam ‘squeeze’ (< *kankam ?). [KS 59]
AFRO-ASIATIC: Proto-Afro-Asiatic *km ‘seize, take, squeeze’; Semitic: Akkadian kamû ‘to capture,’ Arabic kamaša ‘seize, grasp’; Cushitic: Dhalo kam ‘to hold,’ Kwadza komos ‘to grip,’ Iraqw kom ‘to have’; Berber: Tuareg ekmen ‘squeeze’; Chadic: Hausa kàm ‘to catch,’ Musgukaw ‘seize,’ Gidar gëm ‘to take,’ Masa čun. [AA 63, N 157, AB 160]

INDO-EUROPEAN: Proto-Indo-European *gem ~ *gëmo ‘to grasp with both hands, seize;’ Armenian չմել ‘I squeeze’; Greek geuto ‘he took’ (< *gëm-to); Celtic: Old Irish gemel ‘fetters’; Germanic: Old Swedish kimla ‘to crumple’; Baltic: Latvian gūmstu ‘to seize, grasp’; Slavic: Old Church Slavic žëmoš ‘I press, squeeze,’ [IE 368, N 157, AB 171]
URALIC: Proto-Uralic (Illich-Svitych) *kama-HV ~ *koma-rV ‘handful,’ (Rédei) *kona(ř) ‘palm of the hand,’ (Rédei) *káme(-ne); Samoyed: Yenisei Samoyed hammara ‘hand’; Finnic: Finnish kamahlo ~ kahmalo

DRAVIDIAN: Proto-Dravidian *kamV ‘to seize, take, hold,’ Koraga kamdi ‘to steal,’ Telugu kamucu ‘to hold, seize,’ Malto kam ‘to gather (by oneself).’ [D 1326, N 157]

TURKIC: Proto-Turkic *kam-a ~ *qam-a ‘to take, seize,’ Old Uighur qama ‘to take prisoner, surround,’ Kirghiz kama ‘to surround, arrest,’ Tatar kama ‘to herd cattle into a pen,’ Nogai kam-ty ‘to seize.’ [N 157]

MONGOLIAN: Written Mongolian qamu ‘to gather, pick up,’ Khalkha xam˘a ‘to gather, pick up.’ [N 157]

CAUCASIAN: Proto-Caucasian *k’emV ‘armful, handful,’ Proto-Lak *k’ama ‘handful,’ Proto-Lezghian *k’em(a) ‘armful.’ [C 124]


MIAO-YAO: Proto-Miao-Yao *ŋgam ‘to crush, squeeze’ (< *kamgam), Haininh Yao gam ‘to crush or squeeze with the hand.’ [PB 315]

DAIC: Tai: Proto-Tai *ŋpam ‘to lay hold of, grasp’ (< *kamgam), *kum ‘hold with the hand’; Kam-Sui: Sui ȵiam ‘to hold,’ Mak ȵam ‘to clench the fist, take hold of’ (< *ŋpam < *kamgam); Li: Proto-Li *kłam ‘to press with the hand, squeeze,’ Southern Li kom (luoi) ‘to squeeze’ (= press [down]); Ong-Be kom ‘to press down.’ [PB 315]

AUSTRO-NESE: Proto-Austronesian *kem ‘enclose, cover, grasp,’ *gengem ‘hold, grasp in the fist, make a fist,’ Rukai (wa-)ŋangəm ‘to squeeze in the hand.’ [AN 54, 74, 78; WW 103; PB 315-16]

6 KANO ‘arm’

KHOISAN: /Xam //kũ ‘arm,’ /Nu~//en //kan ‘branch;’ Naron //kõa ‘arm,’ Nama //oa(-b); !Kung //kãu ‘branch,’ ?#hã ~ #hã ‘arm.’ [SAK 130, 186]

NIGER-CONGO: Yingulum kɔnĩ ‘arm,’ Fali kan; Bantu: Proto-Bantu *kónd ‘(fore)arm,’ Nyali (i-)kón(-do) ‘hand,’ Swahili (m-)kono ‘arm, forearm, hand, front paw.’ [BA 297, AT 11]

NILO-SAHARAN: Kunama kõná ‘hand,’ u-kun-kula ‘armpit, elbow’ (= armpit), Lili kon ‘hand,’ Berta kɔŋ-k’ọloŋ ~ k”ɔŋ-k”ọlọŋ ‘elbow,’ Teso (á-)káñi ‘hand,’ Masai (enp-)káñá ‘hand.’ [CN 5, AT 79, NSB]

14. Global Etymologies


URALIC: Proto-Uralic (Rédéi) *kon̂o~ *kon̂ala ‘armpit’; Yukaghir kun(-el) ‘10,’ xanba ‘hand’; Ugric: Hungarian hón ~ hón(-alj) ‘armpit’ (-alj ‘that which is beneath’), Vogul kan(-l) ‘armpit’; Finnic: Finnish kain (-alo), Votyak kun(-ul), Zyrian kon(-uvt), (-uvt ‘that which is beneath’). [U 101, KR 178]

DRAVIDIAN: Tulu kaŋkuḷa ‘armpit’ (= arm-hole), Kannada kaŋkur̥- gaŋkolu ‘armpit,’ Kota ganjguly. [D 1234, N 220]

CAUCASIAN: Proto-Caucasian *GHwì~ı́ ‘arm, shoulder,’ Proto-Lezghian *q¨un ‘shoulder,’ Krytz qun¨a, Archi qhun. [C 156]

YENISEIAN: Proto-Yeniseian *ken ‘shoulder.’ [Y 28]

SINO-TIBETAN: Ancient Chinese *k¯en ‘shoulder’; Proto-Tibeto-Burman *kan ‘arm.’ [Y 28, TB 438]

NA-DENE: Proto-Eyak-Athabaskan *gã̂n ‘arm,’ Tanana gan, Tlatskanai oka-ne, Ingalk gàn, Galice gaane? ‘arm, branch,’ Navajo gaan ‘arm.’ [SN, ND]

INDO-PACIFIC: Tasmanian: Southeastern gouna ~ guna ~ gouana ~ wana ‘arm, hand,’ Parawen konene ‘arm,’ Jilim kanán ‘left (arm),’ Dumpu kíiη ‘shoulder,’ Gapun akan ‘arm.’ [T 83, FS 107, IP 820]

?NAHALI khanda ‘shoulder,’ akhandi ‘finger.’ [NA 59, 85; Kuiper believes khanda is probably a borrowing from Kurku.]

?AUSTROASIATIC: Vietnamese cânh ~ càn̩h ‘arm, branch, wing.’

DAIC: Tai: Proto-Tai *xeen ‘arm,’ Dioi kien ‘arm, sleeve,’ Sek keen; Kam-Sui: Sui chin, Then khyin ‘sleeve’; Ong-Be kan (mo) ‘forearm,’ gen (mo) ‘upper arm.’ [PB 379]

7 KATI ‘bone’

Khoisan: /Xam /kuttən ‘bone behind the ear,’ //Ng.‘e ketn ‘bone.’ [BD 87, 326]

Nilo-Saharan: ?Berta k’ara ‘bone,’ East Sudanic: Nera katii(n), Kenuzi kiddled, Kundugr koiđu, Turkana ako-, Lotuko -γ’îty, Bari kuyu, Debri kwedu, Proto-Southeast Surmic *qigec. [ES 21, NSB, SES]

Afro-Asiatic: Proto-Afro-Asiatic *qš ‘bone’; Ancient Egyptian qš ‘bone’; Semitic: Arabic qas.s. ∼ gas.s. ‘sternum’; Cushitic: Proto-Cushitic *(m-)qš ‘bone,’ Kambata miqqa-ta (qq < *qš); Omotic: Jeba ḫūis-u (< *k′us), Berber: Proto-Berber *qš(j) ‘bone,’ Shilha iks, Kabyle iges, Tuareg egès; Chadic: Proto-Chadic k′š(j) ‘bone,’ Proto-West Chadic ẖa-k′asi, Hausa k′ashii, Musgu keš’ke, Gerka ḡas, Somraï guse, Karbo k′asii. [CS 225, AA 11, N 219, OS 620, HF 12]


Uralic: Samoyed: Kamassian kot ‘rib;’ Finnic: Mordvin kaskā ‘sacrum.’ [N 219, SUL 492]

Draavidian: Kurux xo(-ol) ‘bone,’ Malto qoc(-lu). [D 1288]

Chukchi-Kamchatkan: Proto-Chukchi-Kamchatkan * XmlDocument not available

Basque gar-khotx(e) ‘nape (gara = ‘skull’).

Burushaski kanja ‘back of neck, neck joint.’

Sino-Tibetan: Proto-Sino-Tibetan *kut ‘bone,’ Old Chinese *kwat. [SC 57]

Na-Dene: Haida (s)kuts ∼ (s)kuği ‘bone,’ Eyak q′als. [ND]

Indo-Pacific: South New Guinea: Dabu kut ‘bone,’ Dibolog kute, Ngamai kuta, Kawam kutra, Parb kwod, Tunjumau guat, Tokwasu kurt, Bangu kuar, Keladdar kadrowa. [SNG 9]

Kapishana nya-kotsi; Equatorial: Caranga kaiču ‘bone’; Macro-Panoan: Cavineña (epere-)katse ‘rib,’ Komlek kadekotti ‘bone,’ Caduveo koda-uek’o ‘rib.’ [AM 141, P 102, CP 23, AIW, AMN]

8 K’OLO ‘hole’

Khoisan: ≠Au.//ei !kuru ‘quiver’ (n.), !Kung !køro ‘hole,’ !kuru ‘quiver’ (n.), ‘koro ‘hole, grave,’ !O-!Kung kɔlo ‘hollow;’ G//abake koro ‘hole in tree,’ (čui) kxolo ‘nostrils’ (= nose hole); /Xam !køro ‘to be hollow,’ /huru ‘hole,’ /ürú ‘anus.’ /Auni !kuru ‘quiver’ (n.). [SAK 371]

Nilo-Saharan: Songhai nkoro ‘buttocks;’ Saharan: Kanuri kuli ‘anus,’ Teda kulo; Berta k’oj-k’ọọŋ ~ k’eŋ-k’a‘ọŋ ‘elbow’ (= arm-hole, cf. the Kuna name form below); Koman: Buldii kul(ma) ‘buttocks;’ Kuna name kura ‘anus,’ ukunkula ‘armpit, elbow’ (< *kan-kul ‘arm-hole’); East Sudanic: Temein kukuruk(it) ‘buttocks,’ Nandi kulkul ‘armpit,’ So ušlkšl ‘armpit,’ Gaam kura-n ‘hollow (in ground).’ [NS 4, CN 2, 5, ES 3, NSD 3, KER 432]


Korean kul ‘cave.’ [EU]

Japanese-Ryukyuan: Japanese kur ‘hollow, scoop out.’ [EU]

Dravidian: Tulu kulliɡe ‘buttocks,’ kaŋkulə ‘armpit’ (= arm-hole), Kolami kula ‘buttock,’ ganjiguly ‘armpit,’ Gondi kula ‘buttock,’ kàkri ‘armpit,’ Kannada kaŋkur ~ gaŋkhole ‘armpit,’ Telugu kaŋgil ‘breast,’ tsaiŋkili ~ tsakkili ‘armpit’ (= arm-hole), Malayalam akkumam ‘armpit, tickling,’ Tamil akkul ‘armpit,’ akkulu ‘to tickle.’ [D 1234, 2274, Supplement 30; N 220, NSD 3]


Sino-Tibetan: Tibeto-Burman: Proto-Tibeto-Burman *kor ~ *kwar ‘hole,’ *kali ‘armpit, tickle,’ Tibetan (West) kor ‘hollow in the ground, pit,’ Lushei khaŋ ~ khr ‘hole, pit,’ Kor ‘ravine,’ Dimasa ha-khor ‘cave’ (= earth-hole), sisi-khor ‘armpit’ (= tickle-hole), Bodo ha-khor ‘hole, valley,’ Bur-
302 14. Global Etymologies

mese kāli ‘tickle,’ tsakh-kāli ∼ lak-kāli ‘armpit’ (= arm-hole, cf. the Telugu form above), Lakher kili ‘tickle,’ ba-koli ‘armpit.’ [ST 265, 349, 350]

NA-DENE: Haida kunts-qlul ∼ kwun-zool ‘nostril’ (= nose-hole). [ND]


KHOISAN: /Xam Ʇgwā ‘hyena,’ /Ng:!e Ʇxā, /’Auni Ʇkāĩ, Kakia Ʇxāĩ, /Nu-//en Ʇün, !Kung Ʇgwí, !O-Kung Ʇgwí. [SAK 380, DB 48]

AFRO-ASIATIC: Proto-Afro-Asiatic *k(j)n ‘dog, wolf’; Omotic: Haruro kānō, Basketo kana, Kullo kana, Gimira kjān, Kaffa kunānō, Mao kano; Chadic: Gamergu kenē, Jegu kān. [CS 189, N 238, UOL 175]

INDO-EUROPEAN: Proto-Indo-European *kwon ∼ *kun ‘dog;’ Phrygian kan; Greek kuōn; Italic: Latin can(-is); Armenian šun ∼ šan; Sanskrit cvana; Iranian: Avestan span; Tocharian ku ∼ kon; Germanic: Old English hund, English hound. [IE 632, UOL 175, N 238, EU]

URALIC: Proto-Uralic (Illich-Svitych) *kū⁺mä ‘wolf;’ Finno-Ugric: Northern Saami gādne ‘wolf,’ Udmurt k’yjon ‘wolf,’ Komi kōn ‘wolf;’ Samoyed: Ostyak Samoyed kana(-k) ‘dog’ (probably a borrowing) [N 238, UOL 175, EU]

TURKIC: Old Turkish qançiq ‘bitch.’ [EU]

MONGOLIAN: Mongol qani ‘a wild masterless dog.’ [EU]

TUNGUS: Proto-Tungus *xiña ‘dog,’ Manchu (inda-)xun, Udej in’ai, Oroch inaxki, Evenki ina, inakin, Lamut gen, Orok ninda. [N 238, EU]

KOREAN ka ‘dog’ (≠ kani). [N 238, EU]

GILYAK qan ∼ kan ‘dog.’ [EU]

ESKIMO-ALEUT: Sirenik qama ‘wolf.’ [EU]


BASQUE hauz-koin ‘badger’ (lit. ‘bear-dog’).

YENISEIAN: Proto-Yeniseian *kūi ∼ *güi ‘wolverine.’ [Y]


?INDO-PACIFIC: Pila kawun ‘dog,’ Saki kawug, Wodani kwıno. [FS 14]
14. Global Etymologies

?Austronesian: Proto-Oceanic *nkaun ‘dog.’ [WW 60]


10 KU(N) ‘who?’


Niger-Congo: Pam kágé ‘which,’ Dama káñi ‘which,’ Jukun ákè ‘what,’ Proto-Bantu *ki~ ká ‘which,’ Swahili ga-ni ‘what, why, what kind.’ [BA]

Nilo-Saharan: Fur kí ‘who,’ ka ‘what,’ Daza ka ‘which,’ Masai ka ‘which,’ Didinga ñání ‘who’ (< *kani ?), Liguri keneen ‘who,’ Na ya k-rem ‘how many,’ Shatt k-reï ‘how many,’ Shabo kukne ‘who.’ [NS 149, CN 126, HF 12]


Indo-European: Proto-Indo-European *kw~ ‘who,’ *k=e (coordinating conjunction); Indic: Sanskrit kas ‘who’; Iranian: Avestan kō; Armenian o (< *k‘o); Anatolian: Hittite kuš ‘who,’ kuit ‘what,’ Luwian kui ‘who,’ Lydian qis ‘who,’ qid ‘what’; Albanian kë ‘whoso’s; Italic: Latin quis ‘who,’ quis-que ‘whoever,’ quod ‘what,’ quam ‘how, as,’ quom ‘when,’ (arma virum)-que ‘(arms) and (the man)’; Celtic: Old Irish cia ‘who,’ cid ‘what’; Germanic: Gothic hwas ‘who,’ English who, what, when, where, why, how; Baltic: Old Prussian kas ‘who,’ ka ‘what’; Slavic: Old Church Slavic ksto ‘who’; Tocharian: Tocharian A kus ‘who, what.’ [IE 644, N 232, EU, UOL 70]

14. Global Etymologies


[U 44, 46, N 223, 232, EU, KR 140, 191]


Korean ka (interrogative particle). [N 232, EU]

Japanese-Ryukyuan: Japanese ka (interrogative particle, indefinitizer), ka . . . ka (alternating conjunction), Ryukyuan ča ‘what.’ [EU]

 Ainu ka (interrogative particle, indefinitizer), ka . . . ka (coordinating conjunction). [EU]

Gilyak ka (interrogative particle), ko . . . ko (coordinating conjunction). [EU]

Chukchi-Kamchatkan: Kamchadal k’e ‘who’ (genitive k’en), Chukchi mik ~ mek ‘who,’ req ~ raq ‘what,’ Koryak qej . . . qej ‘either . . . or,’ kur ‘to be who?, to be what?,’ Kerek jaq ‘who, what.’ [EU]


Caucasian: Proto-Caucasian *kʰi ‘who, which,’ Kurin ku- ‘what,’ Archi kʰi- ‘who,’ Avar kʰi-n ‘how.’ [UOL 70, SC 149]

Burushaski ke ‘if, when,’ ke ‘and,’ ke . . . ke ‘both . . . and,’ (men . . .) ke ‘(who)ever.’ [B 231, 265]

Sino-Tibetan: Old Chinese *kjei ‘how much,’ Hruso kʰi-nia ‘how many,’ kʰi-nia ‘how far.’ [SC 149]
NA-DENE: Haida gyis-to ~ kīs-to ‘who,’ gösu ~ guu(s) ‘what,’ gyinu ‘where,’
giisant ‘when’; Tingit kusu ~ gisù ~ gi ‘where,’ künsa ‘how much.’ [ND]

AUSTRALIAN: Maung gunuga ~ gigi ‘what,’ Tiwi kuwa ‘who,’ kamu ‘what.’
[RD 373, 376]

NAHALI (nani) ka ‘anyone’ (nani ‘who’), (nan) ka ‘anything’ (nan ‘what’).
[NA 92]

AUSTRASIAN: Munda o-ko-e ‘who,’ o-la ‘what,’ če-le ‘which’; Mon-Khmer:
Vietnamese gi ‘what,’ Nicobarese či ‘who,’ či-n ‘who, what,’ hahā ‘what,’
čan ~ ču ‘where.’ [UOL 70]

AUSTRONESIAN: Proto-Austronesian *ku’a/ʃ ‘how.’ [AN]

AMERIND: Almosan-Keresian: Kutenai ka ‘where,’ Wiyot gu- ‘when, where,’
Yurok kus ‘when, where,’ Passamaquoddy kek̓ ‘what,’ Chemakum ač’is
‘what,’ Quileute ak’is ‘what,’ qo- ‘where,’ Nootka ʔaquis- ‘what,’ Bella
Bella akoqiʔan ‘who,’ Pentlatitch kwana ‘where,’ kw İlk ‘when,’ Upper
Chehalis ka-n ‘do what?’ Keres hēko ‘what ever,’ Quapaw ka ‘what,’ Ofo
kaka ‘what,’ Wichita ʔekiyō ‘who,’ Caddo kwit ‘where,’ Cherokee gago
‘who,’ Onondaga kani ‘where,’ Seneca kwana ‘who,’ Mohawk ka ‘where’;
Penutian: Tsimshian gu ‘who,’ Alsea gau, Kalapuya ūk, Coos qanė ‘where,’
Siouan qani, Klamath kani ‘who,’ ka ‘which,’ Bodega Miwok ʔêke ‘what,’
ʔêketto ‘where,’ Zuni ʔak’i-pi ‘when,’ Tunica kaku ‘who,’ kanahku ‘what,’
kaʔan ‘when,’ Natchez kanne ‘someone,’ gösh ‘what,’ Huave xaŋ ‘who,’ key
‘what,’ Quiche xan ‘when’; Hokan: Achomawi ki ‘who,’ Washo kudija
‘who,’ kugate ‘what,’ kupa ‘where,’ East Pomo kia ‘who,’ k’owa ‘what,’
Chumash kune ‘who,’ keń ‘why,’ Esselen kini ‘who,’ ke ‘where,’ Walapai
ka ‘who,’ Seri ki?, Coahuilteco ka ‘what,’ Chontal kana ‘what,’ ‘when,’
Tlapacuay gana, Jicaque kai ‘where,’ Yurimangui kana ‘what,’ kuna ‘where’;
Central Amerind: Proto-Aztecan *kaan ‘where,’ *keem ‘how,’ *kee-skii
‘how much, how many,’ Nahua aʔkon ‘who,’ Zacapoaxtla akoni, Yaqui
hakuni ‘where,’ Istmus Zapotec guna?, Mazatec kʔia ‘when’; Chibchan-
Paezan: Cuna kana ‘when,’ Miskito ajkia, Paya agini, Terraba kene
‘where,’ Tirib koñe, Totoro kín ‘who,’ Paez kin ‘who,’ kíh ‘what,’ Catio
kai ‘who,’ Moguex kina ‘who, what,’ Tucuru karea ‘why’; Andean: Yah-
gan kunna ‘who,’ kain(a) ‘to whom,’ kana ‘where,’ Tehuelche keme ‘who,’
ken ‘which,’ kenaś ‘when,’ kienia ‘where,’ Araucanian kam ‘how,’ Aymara
kuna ‘what,’ kanisa ‘how,’ Quechua kana ‘who,’ Aymara kuna ‘what,’
kauki ‘where’; Macro-Tucanoan: Ticuna karu ‘where,’ kejaito ‘when’;
Equatorial: Ayore gösí ‘who,’ Tuyoneri kate ‘what,’ ‘Yaruro kanemo ‘when,’
Uru kanču, Wapishana kanun ‘what,’ Puquina kin; Macro-Carib: Yaba-
rana ekkwařiwa ‘when,’ ak̓eťto ‘where,’ Witoto akó ‘what,’ Miranya kia
‘where,’ Faak kiati, Andoke kōde ‘who’; Macro-Panoan: Lule kine-kinema,
Macca karik ‘who,’ kona ‘when’ (rel.), Taruma goga ‘what,’ Tacana ke-
tsunu ‘when,’ kepia ‘where.’ [AM: G102; UOL 70, AMN]
306  14. Global Etymologies

11 KUNA ‘woman’


TURKIC: Proto-Turkic *küni ‘one of the wives in polygamy,’ Old Turric küni ‘wife,’ Kirghiz kùnì, Azerbaijani gûnú. [N 178]


CAUCASIAN: Proto-Caucasian *q(w)änV ‘woman,’ Proto-Dagestan *qonV(βV). [EC, NSC 59]


AUSTRALIAN: Warrgamay gajin ‘female of human or animal species,’ Gamilaraay gunjarr ‘mother,’ Ngaanyatjara ngunytju, Jalnguy guyggun ‘spirit of a dead woman.’ [RD 119, UOL 180]

?AUSTROASIATIC: Mon-Khmer: Nancowry kàn ∼ kãne ‘woman.’ [UOL 179]


12 MAKO ‘child’


DRAVIDIAN: Tamil maku ‘child, young of an animal, son or daughter,’ Malayalam maku ‘son,’ makkal ‘children (esp. sons),’ Kota mog ‘child,’ Toda mok ‘child, son, male, daughter,’ Kannada maga ‘son, male person,’ makan ‘son,’ magu ‘infant, child of either sex,’ Kodagu makkka ‘children,’ Tulu mage ‘son,’ magulu ‘daughters,’ Telugu maga ‘male,’ Konda moga koro ‘boy child,’ gālu ‘daughters’ (< *mgālu), Pendo gār ‘daughter,’ Kuwi maka (vocative used to daughters and sisters in affection), Malto maeqe ‘boy,’ maqi ‘girl,’ maqo ‘small, little, young,’ maqu ‘young of an animal.’ [D 4616, AB 371]

CAUCASIAN: Proto-Caucasian *mik’wV ‘small, young one,’ Proto-Avar-Andi *mok’i ‘small, child,’ Proto-Dido *mik’V ‘small, little,’ Proto-Lezghian *mik’w’V ‘young.’ [C 151]


14. Global Etymologies


13 MALIQ’A ‘to suck(le), nurse; breast’

AFRO-ASIATIC: Proto-Afro-Asiatic: *mlg ‘breast, udder, suck,’ Arabic mlg ‘to suck the breast,’ Old Egyptian mnd (⋯< *mlg) ‘woman’s breast, udder’; Cushitic: Somali maal ‘to milk,’ Rendille maal. [N 291, LN 291]


URALIC: Proto-Uralic (Illich-Svitych) *mälö ‘breast,’ Proto-Finno-Ugric (Rédel) *mälke; Saami mielgä ‘breast, chest,’ Mordvin miilkhä ‘breast,’ Vogul mágl (with metathesis), Hungarian mell, Yukaghir mehu-t. [N 291, R 267]
14. Global Etymologies

**DRAVIDIAN:** Kurux *melkhā* ‘throat, neck’ and Malto *melqe* ‘throat,’ Tamil *melku* ‘to chew, masticate,’ Malayalam *melluka* ‘to chew, champ,’ Toda *melk* ‘mouthful,’ Kannada *mellu* ‘to chew, masticate, eat with a muttering sound,’ *melaku* ‘bringing up again for rumination,’ Telugu *mekku* ‘to eat, gobble,’ Gadaba *mekkap* ‘to eat like a glutton.’ [D 5077, 5080]

**Eskimo-Aleut:** Aleut *umlixa* ‘chest,’ Kuskokwim *muligu* ‘sucks it out,’ *mulik* ‘nipple,’ *miligárga* ‘licks (or sucks) it; kisses it (a child),’ [EU]

**Caucasian:** Proto-Caucasian *mVq’supresslV* ‘throat, larynx,’ Proto-Avar-Andi *maq’ala* ‘throat,’ Proto-Dido *muq’, Proto-Dargi *muq’luq’ ‘chute, gutter.’ [C 142]

**Amerind:** Almosan: Lower Fraser *malq*’w*’throat,*’ Nootka *m’ulg* ‘swallow,’ Kwakwala *mli’t* ‘chew food for the baby,’ *mlik’a* ‘moisten the fingers with the tongue,’ Heiltsuk *melqva* ‘chew food for baby,’ *melv-buit* ‘lick the end of something,’ Yurok *mil’olum* ‘swallow,’ Kutenai *nuq’qol*; Penutian: Chinook *pokk’u* ‘throat,’ *mlik’tan* ‘cheek,’ Wishram *ô-mêq’ l* ‘lick’; Oregon: Takelma *mîlk* ‘swallow,’ Tfalatik *milq, Kalapuya malq-mat* ‘lick’; Yokuts *môk’i* ‘swallow,’ *mik’-is* ‘throat,’ Mixe *amu’ul* ‘suck,’ Zoque *mu’k*; HOKAN: Yuma *ma’aqé* ‘neck,’ Walapai *maq’* ‘throat, neck’ Havasupai *milqé* ‘throat,’ Yavapai *melq’i* ‘neck,’ Mohave *ma’aqé* ‘throat,’ Akwa’ala *milq’i* ‘neck,’ Paipai *milq*; Chibchan: Cuna *muri’-makka* ‘swallow,’ *murgi murgi sae* ‘swallow food’; Andean: Quechua (Cochabamba) *maq’a* ‘throat,’ Quechua (Huaraz) *mallaqa* ‘be hungry’; Aymara *maq’a* ‘swallow, throat’ (a borrowing from Quechua?) Equatorial: Guamo *mirko* ‘drink.’ [P 239, AMN; this etymology is explored in greater detail in Chapter 11.]

14 MANA ‘to stay (in a place)’

?Nilo-Saharan: Tatoga *miin* ‘to stand,’ Shabo *maŋ-ka* ‘to sit.’ [NSB, HF 12]

**Afro-Asiatic:** Proto-Afro-Asiatic *mn* ‘to remain, be firm’; Ancient Egyptian *mn* ‘to remain,’ Coptic *mun*; Semitic: Proto-Semitic *’mn* ‘to be firm, safe,’ Arabic *’mu‘n* ‘to be loyal to someone,’ *’manu* ‘to be safe,’ Geez *’nn* ‘to be faithful,’ Syriac *’amín* ‘firm,’ Classical Hebrew (n-)’nn ‘to be permanent, safe’; Omotic: Gofa *mín* ‘to be firm, strong’; Cushitic: Oromo *manā* ‘house, home,’ Somali *mín, Chadic: Musgu *mín* ‘to be.’ [CS 38, N 287, UOL 192]

?Kartvelian: Georgian *mena* ‘dwelling’ (possibly a borrowing from Iranian languages). [N 287]

**Indo-European:** Proto-Indo-European *men* ‘to remain’; Indic: Sanskrit *man* ‘to linger, not budge from a place’; Iranian: Old Persian *man* ‘to remain, wait for’; Armenian *mnam* ‘I remain, wait for’; Italic: Latin *man*(-ere) ‘to remain’; Tocharian: Tocharian *A mûne* ‘waiting,’ *mûsk* (< *men-sk*) ‘to be.’ [IE 729, N 287, UOL 192]
**DRAVIDIAN:** Proto-Dravidian *man* ‘to remain in a place,’ Brahui *manning* ‘to become, be,’ Malto *mena,* Kurux *mannā,* Kuwi *man* ‘to be, remain, stay,’ Konda *man* ‘to be, stay, dwell,’ Parji *men* ‘to be, stay,’ Telugu *manu* ‘to live, exist,’ *mannu* ‘to last, be durable,’ Malayalam *mamnuka* ‘to stand fast,’ Tamil *maṇṇu* ‘to be permanent, remain long, stay.’ [D 4778, N 287]

**TUNGUS:** Evenki *māna* ‘to live settled, stay in camp for a long time in one place,’ Negidal *maņņa* ‘to remain.’ [N 287]

**CAUCASIAN:** Proto-Caucasian *÷i-ma(n)-* ‘to stay, be,’ Hurrian *mann-* ‘to be.’ [NSC 111]

**BASQUE** *min* ‘to place, set up, settle.’

**BURUSHASKI** *ma-* ‘to be.’ [B 257]

**INDO-PACIFIC:** South New Guinea: Makleu *man* ‘to sit,’ Nabua *mān* ‘to remain,’ Gende *mīna* ‘stay,’ Moge *mana(-munt)* ‘to sit,’ *Kuno amen(-nyint);* Northeast New Guinea: Langtub *min* ‘to stay’; Unclassified New Guinea: Waruna *mana* ‘to dwell,’ *Gogodala mana* ‘to sit, stay.’ [IP 65]

**AMERIND:** Almosan-Keresiouan: Nootka *ma*- ‘dwell’; Penutian: Tsimshian *mān* ‘remain,’ Kalapuy *māni-* ‘wait,’ Maidu *ma* ‘be,’ Zuni *tīma* ‘sit’; Hokan: Subtiaba -ama; Chibchan-Paezan: Cacaopera *ima* ‘wait,’ Puruha *ma* ‘be,’ Timicua -ma ‘inside’; Andean: Cholona *man-* ‘in,’ Aymara *mankxa* ‘inside,’ Araucanian *miyu,* Quechua *ma-* ‘be,’ Yahgan *manī* ‘be,’ *jumanana* ‘live,’ *mōni* ‘remain,’ *kamani* ‘stand’; Equatorial: Dzubucua *mañe* ‘remain,’ Otomi *yamania* ‘live,’ Paumari *gamanani* ‘stand,’ Coche *xamnan* ‘be,’ Macro-Carib: Yameo *mune* ‘sit down,’ Ocaina *mūn-xo* ‘remain,’ *Apiaca umano* ‘wait’; Macro-Panoan: Cashinawa *mana,* Shipibo *manei* ‘remain,’ Chacobo *man-’* ‘wait,’ *Panobo manai,* Lule -ma ‘in’; Macro-Ge: Botocudo *mēn* ‘remain,’ Creengez *moinj* ‘to sit,’ Capoxo *moinjam,* Bororo *aši.setData:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAACAAAAAgCAYAAABhdWlO AAAQElwSU1QLEQAAAABlBMVEX///8AAAABnRFWHRMSEQUMQAAAAAABJRU5ErkJggg==’ ‘to rest,’ Cayapo *kaimaniun* ‘stand,’ *kaman* ‘inside,’ *Tibagi ema* ‘dwell,’ [AM: G46, A 59, MG 99, AMN]

**AFRO-ASIATIC:** Proto-Afro-Asiatic *mn* ‘male, man, person’; Ancient Egyptian *mnn* ‘Min, a phallic deity,’ Old Egyptian *mnyw* ‘herdsman’; Omotic: Wolamo *minō* ‘warrior,’ Janjero *momo* ‘people,’ Cushitic: Proto-Cushitic *mn* ‘man,’ Burjī *mēn-a* ‘people,’ Somali *mun* ‘male,’ Hadiyya *mañana* ‘people,’ *man-čo* ‘person,’ Tembaro *mana,* Iraqw *ameni* ‘woman’; Berber:

**INDO-EUROPEAN:*** Proto-Indo-European *manu(-s) ~ *monu(-s)* ‘man;’ Indic: Sanskrit *manu* ~ *mänus* ‘man, person;’ Iranian: Avestan *manu* ‘man;’ Germanic: Gothic *manna,* Old High German *man,* English man (pl. men), woman (< *wife + man;* Slavic: Old Church Slavic *mënî* ‘man, person,’ Russian *muz* ‘husband.’ [IE 700, N 292]

**URALIC:*** Proto-Uralic (Illich-Svitych) *mänč* ‘man, person;’ Ugric: Vogul *meúči ~ maúši* (self-name), Ostyak *maúč* ~ *maúš* ~ *maš* (self-name of one Ostyak clan), Hungarian *magyar* (self-name); Finnic: Finnish *mies,* Estonian *mees.* [U 114, N 292]

**DRAVIDIAN:*** Kolami *mäš* ‘man,’ *mäč* ‘husband,’ *mäča* ‘wife,’ Naikri *mäs* ‘man,’ *mäsäl* ‘woman,’ Naiki *mäš* ‘husband,’ *mäša* ‘wife,’ Parji *mañja ~ mañña* ‘man,’ Gondi *manja* ‘man, person,’ Konda *mäši* ‘husband,’ Kurux *mët ~ mét* ‘adult man, husband,’ Tamil *mân†ar* ‘people, men.’ [D 4791; Illich-Svitych’s comparison (N 292) is with D 4774: Tamil *mân†ara* ‘king, chief, husband,’ etc. The two are probably related.]

**JAPANESE-RYUKYUAN:*** Old Japanese *(wo-)*mina ‘woman’ (mod. onna). [SY]

**AINU** *meno(-ko) ~ mene(-ko)* ‘woman.’

**CAUCASIAN:*** Proto-Caucasian *mVnxV* ‘man, male.’ [NSC 116]

**YENISEIAN:*** Proto-Yeniseian *pix- ‘man.* [NSC 116]

**INDO-PACIFIC:*** Bilakura *munan* ‘man,’ Warenbori *mendo,* Osum *aminika* ‘woman,’ Ikundun *munu‘ man.’ [FS 92, 93, 106]

**NAHALI** manchio ~ *manco* ‘man,’ man-ṭa ‘men.’ [NA 89]

**MIAO-YAO:*** Proto-Miao-Yao *hmnan* ‘person,’ Miao *hmlog ~ hmun* (self-name of the Miao), Yao *man ~ myen ~ mn* (self-name of the Yao). [PB 336]

Global Etymologies


16 MENA ‘to think (about)’

Khoisan: Sandawe mēna ‘to like.’

Niger-Congo: Fulup -maman ‘know,’ Mambila mini ‘think,’ Malinke men ‘understand,’ Bambara me, Proto-Bantu *màni ~ *mèni ~ *mèny ~ *màn ‘know,’ Namshi mei, Ibo ma, Mandyak me. [NC 28, KS 45, BA IV: 8, 12]

Kordofanian: Tumalé aiman ‘think.’ [NK 41]

Nilo-Saharan: Songhai ma ‘understand,’ Daza monar ‘know,’ Dinik mài, Lotuko mij, Proto-Daju *minage ‘to dream,’ Shatt mimig, Ik mi-n-es ‘to love,’ Teso a-min. [KS 45, NSB, KER]


Dravidian: Tamil maṇu ‘prayer, request, word,’ Kannada maṇuve ‘request,’ Telugu manavī ‘prayer, humble request,’ Irula maṇi ‘talk, speak,’ Kota mayn- ‘talk, scold, abuse.’ [D 4671, 4775, N 281]

?Turkic: Turkish manı ‘folk song,’ Crimean Turkish manā ‘folk song, melody.’ [LN 281]
Basque mun ‘medulla,’ munak (pl.) ‘brains.’ [LC 916]

?Burmushki minas ‘story, tale.’ [B 506]


17 MI(N) ‘what?’

Khoisan: #Au.//ei kama ‘if, when,’ G//abake /kam ‘when,’ Naron kama ‘if, when,’ Nama hamo ‘when,’ maba ‘where,’ Kxoe ma ‘who, which,’ /Nu-//en maba ‘where.’ [SAK 384, 757, 758, UOL 71]

314

14. Global Etymologies


INDO-EUROPEAN: Proto-Indo-European *mo- (base of interrogative adverbs); Anatolian: Hittite ma-ˇsi-ˇs ‘how much,’ ma˘h˘han ‘when,’ Luwian măn, Hieroglyphic Hittite mana ‘if, when’; Celtic: Old Irish má ‘if,’ Middle Breton ma ‘what’; Tocharian: Tocharian A mánt ‘how.’ [KA 124, 135, N 300, UOL 71]


?DRAVIDIAN: Kajkadi midă ‘what,’ Burgendí mi, Tamil (even-)um ‘(who)ever.’ [N 300, UOL 71]

TURKIC: Proto-Turkic *mi ‘what,’ Chuvash mên ‘what,’ mïše ‘how much,’ mênle ‘what kind of,’ Old Uighur mu ~ mii (sentence question enclitic), Turkish mi (sentence question enclitic). [N 300, EU]

MONGOLIAN: Mongolian -ă (< *wu < *mu) (sentence interrogative), Monguor amu ~ ama ‘what.’ [N 300, EU]

JAPANESE-RYUKYUAN: Ryukyuan mi ‘what,’ -mi (sentence interrogative enclitic). [EU]

AINU mak ~ makenak ‘what,’ makan ‘what kind.’ [EU]

CHUKCHI-KAMCHATKAN: Proto-Chukchi-Kamchatkan *m-qaV ‘what,’ *m-ke ‘who,’ *ma/or when,’ *miŋ ‘which,’ Chukchi mìkin ‘who,’ mi-k ‘where,’ Kamchadal min ‘which, what sort.’ [EU, CK]

CAUCASIAN: Proto-Caucasian *ma (interrogative particle), Chechen mìl’a ‘who,’ Bats me. [KA 135]

BURUSHASKI mën ‘who,’ amín ‘which,’ mën ( . . . kë) ‘who(ever).’ [I. 265]

YENISEIAN: Proto-Yeniseian *wi- ~ *we- ‘what.’ [Y]


AUSTRALIAN: Proto-Australian *mînhá ~ *mînya ‘what,’ Dyirbal minya, Pitta-Pitta minha, Gumbaynggir minya, Malyangapa minhaga, Yota-Yota minhe, Diyari minha. [RD 373, 376]

NAHALI mingay ‘where,’ miyan ‘how much.’ [NA 91]
14. Global Etymologies


Amerind: Almosan-Keresiouan: Kwakwala m’as ‘what,’ Tutelo mana ‘who,’ matswe ‘what,’ Tutelo m̃a, m̃ai ‘who,’ Central Sakait mō, mā ‘what.’ [NA 91, UOL 71]

Amerind: Almosan-Keresiouan: Kwakwala m’as ‘what,’ Tutelo mana ‘who,’ m̃a ‘what,’ Tutelo m̃a, m̃ai ‘who,’ Central Sakait mō, mā ‘what.’ [NA 91, UOL 71]
18 PAL ‘2’


AFRO-ASIATIC: Omotic: Kafa bæra ‘other,’ Mocha bæro, Dime bal; Cushitic: Saho bæray ‘2nd,’ Oromo bíra; Chadic: Proto-Central Chadic *(kV-)bwVR ‘2.’ [IE 802, 986, IS 356]


DRAVIDIAN: Proto-Draavidian *päl ‘part, portion,’ Tamil pā ‘part, portion, share,’ Malayalam pāl ‘division, part,’ Telugu pālu ‘share, portion, part,’ Parji pēla ‘portion.’ [D 4097, IS 356]

INDO-PACIFIC: Andaman Islands: Biada (ik-)pūur(-da) ‘2,’ Kede (ér-)pōl, Chariar (nér-)pōl, Juwoi (ró-)pāuēr; New Guinea: Ndani bere, Sauweri pere; Tasmanian: Southeastern boula ~ bura, Southern pooalih. [T 331, VB]

AUSTRALIAN: Proto-Australian *bula ‘2,’ Proto-Pama-Nyungan *(nu)NpālV ‘(you) two,’ *(nu)Npula ‘they two,’ Ngiyambi (a)bulā ‘one of a pair.’ [RD 356, BB 7, 31]

AUSTROASIATIC: Proto-Austroasiatic *(m)bær ‘2;’ Munda: Santali bar, Kharia (u-)bar, (am-)bar ‘you two,’ Juang ambar, Remo ṭłańbar ‘2;’ Mon-Khmer: Khmu’ bār, Bahnar ʔbar, Jeh bar, Old Mon ʔbar, Old Khmer ber, Sakai hma bi, Khasi ri, Riang (k-)ār, Palaung ār ~ a, par ‘you two,’ Temiar ʔb(ː)ar(ː) ‘2,’ Central Nicobarese ʔ. [PB 135, UOL 94]

MIAO-YAO: Proto-Miao-Yao *(a)war ~ *(s)war ‘2,’ Proto-Miao *way (< *war), Proto-Yao *(w)j. [PB 415]

DAIC: Mak wa ‘twin,’ Ong Be von ‘2.’ [PB 415]

AUSTRONESIAN: Proto-Austronesian *kɔ(m)bal ~ *(ŋ)kɔ(m)bar ‘twin,’ Japanese kēbar ‘doubled,’ kēmbar ‘twin,’ Motu he-kapa ‘twins,’ Roro akabaní ‘8’ (= 4-pair). [AN 76, WW 227, PB 415]
14. Global Etymologies


19 PAR ‘to fly’

Niger-Congo: Proto-West Sudanic *pil ‘to fly,’ Serer *fol, Same *pere, Ewe *fô ‘to jump,’ Yoruba *fô ‘fly,’ Grebo *fri, Igbo *fé, Ijo *fin. [KS 32]

Nilo-Saharan: Dinka *par ‘to fly,’ Nubian *fire ‘to flutter,’ Teso *a-poror ‘to fly,’ Teda *bur-ci ‘to jump,’ Songhai *firi ‘to fly,’ Ik *por-çan, Maasai *pieri, Majang *pi.


Kartvelian: Proto-Kartvelian *p’er ‘to fly,’ Georgian *p’er, Svan *p’er; Proto-Kartvelian *prin ‘to fly,’ Georgian *prin ~ *peng, Mingrelian *purin, Chan *purin. [KA 152, 190, IS 346]


Uralic: Yukaghir *perie ‘feathers,’ *perienze ‘feathered,’ perieui ‘have wings’; Proto-Uralic *parV ‘to fly,’ Ugric: Ostyak *par ~ *pur ‘to fly.’ [IS 346]

Draavidian: Proto-Draavidian *parV ~ *parV ‘to fly, run, jump,’ Tamil *para ‘to fly, hover,flutter, move with celerity,’ Malayalam *parakkala ‘to fly, flee,’ *para ‘bird,’ *par ‘flight,’ Kota *par- ‘to fly,’ Toda *pâr, Kannada *pâr ‘to leap up, run, jump, fly,’ Kodagu *pâr ‘to fly, leap,’ Telugu *para ‘to run away, flee,’ *para ‘a kind of bird,’ Kui *pâšk ‘to fly,’ Kuwi *pâad ‘to run away.’ [D 4020, NSD 27, IS 346]

?Tungus: Evenki *hâr ‘to soar.’ [IS 346]

Gilyak *parpar ‘to hover, fly about.’ [EU]

Caucasian: Proto-Caucasian *pîrV ‘to fly,’ Proto-West Caucasian *pær, Ubyx *pær, Abkhaz *pir; Proto-Lezghian *pVr-, Udi *pur, Archi *parx, Proto-Avar-Andi *par-pV-; Proto-Caucasian *pærVpālV ‘butterfly, moth,’
14. Global Etymologies

Proto-West Caucasian *para*palo ‘moth,’ Proto-Lezghian *pa(r)pal- ‘butterfly.’ [C 162, 167; KA 152, 190]

BASQUE pimpirina ‘butterfly’ (< *pir-pir-).


?INDO-PACIFIC: Baham paru-baru ‘bird,’ Kondo boro, Kare purupuru, Buna-piropir ‘butterfly.’ [FS 8, 135]

Nahali aphir ‘to fly.’ [NA 59; according to Kuiper this is a borrowing from Kurku]

AUSTROASIATIC: Munda: Proto-Munda *apir ‘to fly’; Mon-Khmer: Mon pau, Khmer par, Bahmar par, Jeh pal, Vietnamese bay. [PB 482]

DAIC: Tai: Proto-Tai *bhin ‘to fly,’ Dioi bin; Sek *bil ∼ *biil; Kam-Sui: Proto-Kam-Sui *pwen ∼ *bwen, Kam pen, Sui win ∼ vyen, Mak vin; Lakkia pow; Ong-Be vin. [PB 394]

AUSTRONESIAN: Proto-Formosan *(maq)baR ‘to fly,’ *(mi-)pæRpæR. [PB 394]

20 POKO ‘arm’

?KHOISAN: Hadza upukwa ‘leg, hind leg, foot,’ ufukwani ‘thigh.’ [BD 247, 249]


NILO-SAHARAN: Bagirmi boko ‘arm,’ Baka baka, Berta buá, Didinga iba. [KS 4, CN 3, UOL 194]

INDO-EUROPEAN: Proto-Indo-European *bhāghu(s) ‘arm, forearm, elbow’;

Indic: Sanskrit bāhiḥ ‘arm’; Iranian: Avestan bāzus; Armenian bazuk ‘forearm’ (a loan from Iranian languages, according to Pokorny); Tocharian: Tocharian A poke ‘arm,’ Tocharian B pauke; Greek pakhs ‘elbow, forearm’; Germanic: Old English bōg ‘arm, shoulder, bough,’ English bough. [IE 108, UOL 194]

DRAVIDIAN: Kurux pāknā ‘to take up into one’s arms,’ Malto pāke ‘to take in the lap.’ [D 4050]

MONGOLIAN: Proto-Mongolian *barγu- ‘upper arm.’ [AD 20]

BURUSHASKI: Hunza barγu ‘double armful,’ Werchikwar barγo ‘taking or embracing in two arms.’ [B 65, W 38]

YENISEIAN: Proto-Yeniseian *bog ‘hand, palm.’ [Y 28]

SINO-TIBETAN: Tibeto-Burman: Proto-Tibeto-Burman *pow ∼ *bow ‘arm’ (cf. English bough for a similar phonetic development). [TB 442]

?INDO-PACIFIC: Andaman Islands: Bea pag ‘claw,’ Bale poag; Tasmanian
14. Global Etymologies


NAHALI boko ∼ bokko ‘hand.’ [NA 74]

?AUSTROASIATIC: Semang pâk ‘hand,’ ta-pak ‘to slap.’ [NA 63]

DAIC: Tai: Proto-Tai *ʔba ‘shoulder’; Sek va; Kam-Sui: Mak ha; Ong-Be bea;

Li: Proto-Li *va; Laqa mua ‘shoulder’ (< *mb(γ)a). [PB 378]

AUSTRO-NESEANS: Proto-Austronesian *(ʔ)baɣa‘shoulder,’ Proto-Formosan *qa-baɣa-(a)n, Proto-Oceanic *(qa-)paɣa, Mukawa kabara, Palwa kavara. [AN 19, WW 187, PB 378]


21 PUTI ‘vulva’

NIGER-CONGO: Mande: Malinke butu ‘vulva,’ Guro buri, Bobo-Fing bido, Bisa bid; Bantu: Luganda -buta ‘womb,’ Kunda -budu, Swazi -go-basti, Ki-sikongo -buti. [HJ, M]

NILO-SAHARAN: Songhai: Gao buti ‘vulva,’ Djerma bute; Koman: Ganza pit, Koma bitt. [NS 145, NSD 59]

AFRO-ASIATIC: Proto-Afro-Asiatic *pwt ‘hole, anus, vulva’; Omotic: Ganjule pote ‘vagina’; Semitic: Hebrew pot ‘vulva’ (“secret parts” in the King James Version, Isaiah 3:17); Cushitic: Somali fiito ‘anus,’ Darasa fido
Global Etymologies

'genitals,' Oromo fuği 'vulva'; Chadic: Jegu paate, 'vulva,' paato 'penis,' Angas fut 'hole.' [CS 381, IS 340, WM 64]

KARTVELIAN: Proto-Kartvelian *put' 'hole,' Svan put'u. [IS 340]

INDO-EUROPEAN: Proto-Indo-European *puto 'cunnus'; Indic: Sanskrit puta 'girl,' Old French pute (mod. puteain) 'whore,' Provencal puta(-na), Spanish puta; Germanic: Old Icelandic fuð 'cunnus,' Middle High German vut 'vulva,' Swiss German fotz ∼ fotza, Swedish fitta, fod 'rear end' (dialectal). [WP II: 21, IE 848, SM 1013]


DRAVIDIAN: Brahui punđì 'anus, buttocks,' pōs 'vulva,' Tamil puṇṭai 'vulva,' pûru ∼ pûru 'anus,' poccu 'vulva, anus,' Malayalam pûru 'buttocks, vulva,' Kannada pucci 'vulva,' Telugu puđa 'anus,' Tulu puṭi 'vulva,' Kodagu putù, Kota pû, Toda piḍy 'penis,' Kuwi putki. [D 4273, 4379, 4476, NSD 59]

MONGOLIAN: Middle Mongolian hûtû-gûn 'vulva.'

JAPANESE-RYUKYUAN: Old Japanese pu-to 'vulva' (mod. hoto). [SY]

Eskimo-Aleut: Proto-Eskimo-Aleut *putu 'hole.' [EA]

CAUCASIAN: Proto-Caucasian *put'i 'genitals (mostly female),' Proto-Nax *but 'vulva,' Proto-Avar-Andi *but'a, Proto-Lak *put'i 'tube,' Proto-Dargi *put'i 'anus,' Proto-Lezghian *pot 'penis.' [C 168]

BASQUE poto-rro 'pubis, vulva.'

AUSTRALIAN: Luridya pudá 'vulva.' [VB]

AUSTRONESIAN: Proto-Austronesian *betik 'vagina,' *puki 'vulva' (< *puti ?; cf. East Rukai pati 'vulva'), Ami puki, Tsou buki 'penis.' [AN 121, WW 231, 233, PB 417]

AMERIND: Almosan-Keresiouan: Delaware saputti 'anus,' Mohegan sebud, Wiyot bes 'vagina,' Upper Chehalis -pš 'anus'; Penutian: Chinook puč, Yaudanchi poto 'penis,' San Juan Bautista lapus 'anus,' Southern Sierra Miwok potol; Hokan: Washo (d-)ibis 'vagina,' Karok viθ, Diegueño lapičatt, Tequistlatec (la-)bešu?, Chibchan-Paezan: Move butie, Paya peta-is-tapeca 'anus,' Chimú pot, Ayoman busi 'vagina,' Allentiac poru; Andean: Quechua upiti 'anus,' Yanama pûta 'hole,' Aymara put'hu; Macro-Tucanoan: Gamella sebu 'vulva,' Uaina mbitioko 'anus,' Uasôna lubitioko; Equatorial: Guahibo petu 'vagina,' Guayabero sil-fûta 'vulva,' Kandoshi apcir(-iç), Toyeri apuit 'vagina,' Wachipairi ped, Piapoco afýutani 'buttocks,' Tarianâ pãti-niawa 'vagina,' Warakena péde 'clitoris,' Caranga piče 'vulva,' Uro piši, (cf. also such Equatorial forms as Siusi t'uta-pote 'vagina,' Campa sibîči 'vulva,' sibîči 'penis,' Uro sapiri 'genital organ'); Macro-Carib: Jaricuna poita 'vagina,' Pimenteira pitte-anaung, Waiwai boçi 'pubic hair,' Motilón pirri 'penis'; Macro-Panoan: Cavineña
14. Global Etymologies

busu-kani ‘anus,’ Tagnani opet, Tiatinagua besi ‘penis,’ Panobo busi,’ Lule pesu; Macro-Ge: Mekran putote. [AM 263, EQ 121, AMN]

22 TEKU ‘leg, foot’

Niger-Congo: Konyagi -tak ‘heel,’ Gurmana -duge, Jarawa -dudug-ul, Ki-kuwu -togi. [HJ II]


Dravidian: Proto-Central Dravidian *tāk ‘foot,’ Parji tāk, Pengo tāŋ(g), Kui tāka. [D 3151, LN 255] Cf. also Telugu ēkka ‘hoof,’ Naikri ēkka, Konda ēka, Kuwi ēkka. [D 2970]

Caucasian: Proto-Caucasian *ṭ’H₂ālq’wV ‘part of the leg,’ Proto-East Caucasian *ṭ’wehwV ‘foot,’ Proto-Dido *ṭ’q’wV ‘sole of the foot,’ Proto-Lezghian *ṭ’elq’wI ‘shin, ankle.’ [C 196]

Na-Dene: Proto-Eyak-Athabaskan *t’ahx ~ *t’ah ‘foot.’ [DC]


23 TIK ‘finger; one’


Uralic: Votyak odik ‘1,’ Zyrian őtik. [U 138, EU]

Turkic: Chuvash tek ‘only, just,’ Uighur tek ‘only, merely,’ Chagatai tek ‘only, single,’ Turkish tek ‘only,’ teken ‘one by one.’ [EU]

Korean (t)tayki ‘1, thing,’ teki ‘1, guy, thing,’ Old Korean tek ‘10.’ [EU]

Japanese-Ryukyuan: Japanese te ‘hand.’ [UOL 195]

Ainu tek ~ teke ‘hand,’ atiki ‘five.’ [UOL 195, EU]

Gilyak ŋak ‘once.’ [EU]

Chukchi-Kamchatkan: Kamchadal itygin ‘foot, paw.’ [EU]

Eskimo-Aleut: Proto-Eskimo-Aleut *q(t)tiq ‘middle finger; Eskimo: Kuskokwim tik(-iq) ‘index finger,’ Greenlandic tik(-iq) ‘index finger,’ tikkuag-paa ‘he points to it; Aleut: Attu tik(-laq) ‘middle finger,’ atgu ‘finger,’ tɔɾ̂̄aataq ‘1,’ Atka atakan. [EU, EA 121]

Yeniseian: Proto-Yeniseian *tak ‘finger.’ [VT]

Sino-Tibetan: Archaic Chinese *t’iɛk ‘single, 1;’ Tibeto-Burman: Proto-Tibeto-Burman *tyik ‘1,’ Rai tik(-pu), Tibetan (g-)tṣig. [ST 94]

Na-Dene: Haida (s)-tl’a ‘with the fingers,’ Tlingit t’veeq ‘finger,’ t’ek ‘1;’ Eyak tikhi; Athabaskan: Sarsi tlik’-(aza), Kutchin (i-)laq, Hupa lаʔ, Navajo lаʔ. [ND]

14. Global Etymologies

AUSTROASIATIC: Proto-Austroasiatic *(k-)tig ‘arm, hand’; Munda: Kharia tiʔ, Mon-Khmer: Riang tiʔ, Wa taiʔ, Khmer ṭai, Vietnamese tay, Proto-Aslian *tik ∼ *tig. [PB 467, UOL 195]

MIAO-YAO: Proto-Miao-Yao *nto ‘finger’; Proto-Yao *doʔ; Proto-Miao *ntai ‘point with the finger.’ [PB 356]

DAIC: Proto-Li *dlia ‘finger,’ Northern Li tieŋ ∼ theŋ, Loi theŋ ∼ čiaŋ. [PB 356]

?AUSTRONESEAN: Proto-Austronesian *(tu-)di ‘point with the finger.’ [AN 140, WW 156, PB 356, UOL 195]


24 TIKA ‘earth’

?NIGER-CONGO: Proto-Bantu *tâkà ‘earth, mud, ground, soil,’ Swahili taka ‘dirt, refuse.’ [BA IV: 87]

KARTVELIAN: Proto-Kartvelian *tiqa ∼ *diqa ‘soil, clay,’ Georgian tixa ‘clay, dirt’ (< Old Georgian tiqa), Mingrelian dixa ∼ dexa ‘soil, earth,’ Chan (n)dixa ‘soil.’ [KA 94, N 69]


DRAVIDIAN: Tamil tukal ‘dust,’ Telugu d¯ugar¯a ‘dust, dirt,’ Kolami t¯uk ‘dust, earth, clay,’ Naikri tuk ‘earth, clay,’ Parji t¯ukud ‘earth, clay, soil,’ Gadba t¯ukur. [D 3283]

JAPANESE-RYUKYUAN: Old Japanese tuk¨ı ‘mud,’ tuki ‘land’ (mod. t¨usi ∼ t¨usCi). [SY]

BURUSHASKI t˚ık ∼ tik ‘earth, ground.’ [B 351]

YENISEIAN: Proto-Yeniseian *t˚ëq- ‘clay, dirt,’ Ket tag-ar ‘clay,’ Kot t˚hag-ar ‘dirt.’ [SC 76]

SINO-TIBETAN: Lushei diak ‘mud,’ Sho d˚ek ∼ dek ‘earth.’ [IST 221]

14. Global Etymologies

25 TSAKU ‘leg, foot’

?NIGER-CONGO: Bantu: Proto-Bantu *t‘akù ‘calf of the leg.’ [BA 79]


Afro-Asiatic: Cushitic: Beja sìkwìna ‘foot,’ Quara sukanà; Semitic: Hebrew šoq ‘leg,’ Arabic sàq; Berber: Shilha (ta-)zux(-t) ‘foot’; Chadic: Proto-West Chadic *sAkA ‘leg,’ Bolewa šëke ‘foot,’ Fali sika. [CS 265, AA 34, OS 292]


Uralic: Yukaghir t’oy(-ul) ‘foot, leg’; Ugric: Ostyak ság(-súl) ‘soh(-ált)’ ~ šog(-as) ‘back side of the leg from the heel to the bend of the knee (of a human being); back hoof (of a horse); Finnic: Saami çæwγ ‘hock of reindeer or other quadrupled.’ [U 92]

Chukchi-Kamchatkan: Kamchadal t‘k(-ana) ~ t‘ki ‘foot, leg, paw.’ [Swadesh 1962]

Caucasian: Proto-Caucasian *c‘Y[l]k’w ‘foot, hoof,’ Proto-Avar-Andi *c‘ik’wa ‘foot.’ [C 75, DC]

Burushaski: Hunza šak ‘arm, forearm (of a human being); thigh, upper part of the leg (of an animal),’ Werchikwar šak. [B 320, W 215]


Indo-Pacific: Andaman Islands: Onge t‘ige ‘leg,’ Biada t‘ag, Puchikwar t‘ok, Juwoi čok; Central New Guinea: Mikaru saga ‘foot,’ Grand Valley Dani (ne-)šok ‘(my) foot’; East New Guinea: Korona sogo ‘foot,’ Sikube suku, Mafulu soge, Kambisa suga. [IP 80, T 458]


26 TSUMA ‘hair’

KHOISAN: !Kung čum ‘shell,’ šum ‘skin,’ Eastern ≠Hua č’ü, t”ü, dt”ü ‘skin’; G//abake ča ~ čo ‘skin’; /Xam tu ‘shell.’ [SAX 597, 807]

NILO-SAHARAN: Nyangiya sim-at ‘hair,’ Nandi sum. [KER 445]


CAUCASIAN: Proto-Caucasian *tší wém ‘eyebrow,’ Proto-Lezghian *tšwem, Proto-Nax *tš’a-t’ém ‘hair’ [C 70]

BASQUE zam-ar(r) ‘lock of wool, shock of hair.’ [SC 12]

YENISEIAN: Proto-Yeniseian *tsoje ‘hair.’ [SC 12]

SINO-TIBETAN: Proto-Sino-Tibetan *tsüam ‘hair’; Archaic Chinese *sam ~ *šam ‘hair, feather’; Tibeto-Burman: Proto-Tibeto-Burman *tsam ‘hair,’ Lepcha ātsom, Tibetan (ʔag-)tsom ‘beard of the chin’ (= [mouth]-hair), Kanauri tsam ‘wool, fleece,’ (mik-)tsam ‘eyebrow’ (= [eye]-hair), Magari tšām ‘hair, wool,’ Burmese tsham, Lushei sam ‘hair (of the head),’ Dhimal tšam ‘hide, bark,’ Garo mik sam ‘eyebrow,’ Nung aŋsam ‘hide.’ [ST 73, 191, UOL 194, SS 23]

MIAO-YAO: Proto-Miao-Yao *sjām ~ *sjām ‘beard, moustache.’ [PB 307]

14. Global Etymologies

27 AQ’WA ‘water’

KHOISAN: Northern: !q’ !kung kā ’to rain,’ !kung k”ā ‘drink’; Central: Naron k”ā ‘drink’; Southern: /kam-ka !ke k”wā ∼ k”wē ‘drink,’ kā ’to rain,’ //ng k”ā ∼ k”ē ‘drink,’ kā ’to rain,’ Batwa k”ā ∼ k”ē ‘drink,’ /auni k”ā ‘drink,’ Masarwa k”ā ‘drink,’ //en k”ā ‘drink.’ [KE 261]


AFRO-ASIATIC: Proto-Afro-Asiatic (Illich-Svitych) *q(w) ‘water,’ (Ehret) *ak’w-; Omotic: Proto-North Omotic *ak’-; She kai ‘wet,’ Janjero ak(k)a ‘water,’ Kaffa ać’o, Mocha ać’o, Shinasha ać’o, Baditu wat’č; Cushitic: Proto-Cushitic (Ehret) *-k’w- ‘to be wet,’ (Illich-Svitych) *qw ‘water,’ Agaw aq, Namur aq, Damot ag‘o ‘water,’ Proto-East Cushitic (Ehret) *k’oy- ‘wet,’ Hadiyya wō ‘water,’ Tambaro waha, Sidamo waho, Iraqw aha ‘drink.’ [N 139, EU, AM 87, CE 348]


URALIC: Proto-Uralic (Rédéi) *yoka ‘river.’ [R 99–100]

JAPANESE aka ‘bilge water.’ [JP 100]

AINU wakka ‘water,’ ku ‘drink.’ [JP 100]

CAUCASIAN: Proto-Caucasian *-YqV ‘suck,’ Proto-Lezghian *pōq”a ‘drink,’ Lezghian χ”a-l, Agul qas, Proto-Lezghian *pōq”a- ‘rain,’ Lezghian q”a-z, Rutul haw”as, Tsakhur jow”i; Proto-Nax *aq- ‘suck(le),’ Chechen -aq- ‘suck,’ Proto-Dargi *-aq- ‘suck(le).’ [C 3, 16]

?BURUSHASKI haγ-um ‘wet.’ [AC]

SINO-TIBETAN: Proto-Sino-Tibetan *Ku ‘fluid, spill,’ Newari kwo ‘river,’ Khaling ku ‘water,’ Kachin k’u. [NSC 43]

INDO-PACIFIC: Aywu okho ‘water, river,’ Syiagha okho ‘water,’ Yareba ogo, Yonggom og, Ninggirum ok. [FS 96, 134]

AUSTRALIAN: Proto-Australian *gugu ‘water.’ [AC]

AMERIND: Alamosan-Keresian: Proto-Central Algonquian *akwa ‘from water,’ Kutenai -q- ‘in water,’ Quileute kwāya ‘water,’ Snohomish q”aʔ, Caddo koko; Penutian: Nass akj-s, Takelma ug’w ‘drink,’ Wintun wak’ai ‘creek,’ Zuni k’a ‘water,’ Atakapa ak, Yuki uk, Tetontepec uuʔk ‘drink,’ Yucatec uk ‘be thirsty,’ Hokan: Chimariko aqa ‘water,’ Kashaya ḫahqa ‘water,’ q’o ‘drink,’ Seri ḫax ‘water,’ Diegueno ḫax, Quinigua kwa, Tonkawa ḫix, Tequistlatec ḫaxaʔ, Central Amerind: Proto-Chinantece *γ”a
14. Global Etymologies

‘stream, river’; Chibchan-Paezan: Shiriana koa ‘drink,’ Chimila uk-, Bin-
ticua agua, Allentiac aka ‘water’; Andean: Iquito aqua, Quechua yaku, Ya-
manaka aka ‘lake’; Macro-Tucanoan: Auake okōa ‘water, river,’ Cubeo oko
‘water,’ Tucano axko; Equatorial: Amniape ākū, Quitemoako, Uaraicu
uaka ‘wash,’ Terena oko ‘rain,’ Chipaya ax’ ‘wash’; Macro-Carib: Yagua
xa ‘water,’ Witoto joko ‘wash,’ Macushi u-wuku ‘my drink,’ Waiwai
woku ‘drink,’ Taulipang ai’ku ‘wet’; Macro-Panoan: Lule uk ‘drink,’ May-
oruna uaka ‘water,’ Culinoyaku ‘water,’ waka ‘river,’ Huarayo hakua
‘wash’; Macro-Ge: Koraveka ako ‘drink,’ Fulnio waka ‘lake,’ Kamakan
kwa ‘drink,’ Chavante kō ‘water,’ Aponegicran waiko ‘drink.’ [AM 87,
AMN]

ABBREVIATIONS

A Andean, Greenberg 1987
AA Afro-Asiatic, Greenberg 1963
AAD Afro-Asiatic Dictionary, Diakonov 1981–
AB Allan Bomhard, 1987
AC A. Capell, 1956
AD Anna Dybo, 1988
AIW Mary Key, 1987
AK Almosan-Keresiouan, Greenberg 1987
AM Amerind, Greenberg 1987
AMN Amerindian Notebooks, 23 vols., Greenberg 1981
AN Austronesian, Dempwolff 1934–38
AT A. N. Tucker and M. A. Bryan, 1957
B Burushaski, Lorimer 1938
BA Bantu, Guthrie 1967
BB Barry Blake, 1988
BD Bushman Dictionary, Bleek 1956
C Caucasian, Nikolaev and Starostin 1992
CA Central Amerind, Greenberg 1987
CAN Central Amerind Notebook, Greenberg 1981
CCE Vladimir Orel and Olga Stolbova, 1988
CE Christopher Ehret, 1989
CK Chukchi-Kamchatkan, Mudrak 1990
CN Chari-Nile, Greenberg 1963
CP Chibchan-Paezan, Greenberg 1987
CS Marcel Cohen, 1947
D Dravidian, Burrow and Emeneau 1984
DB Dorothea Bleek, 1929
DC Dene-Caucasian, Nikolaev 1991
### 14. Global Etymologies

<table>
<thead>
<tr>
<th>Code</th>
<th>Etymology</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL</td>
<td>D. R. Leshchiner</td>
<td>1989</td>
</tr>
<tr>
<td>EA</td>
<td>Eskimo-Aleut, Mudrak</td>
<td>1989</td>
</tr>
<tr>
<td>EC</td>
<td>East Caucasian, Starostin and Nikolaev</td>
<td>1975</td>
</tr>
<tr>
<td>EQ</td>
<td>Equatorial, Greenberg</td>
<td>1987</td>
</tr>
<tr>
<td>ES</td>
<td>East Sudanic, Greenberg</td>
<td>1963</td>
</tr>
<tr>
<td>EU</td>
<td>Eurasian, Greenberg to appear</td>
<td></td>
</tr>
<tr>
<td>FS</td>
<td>F. Seto</td>
<td>1988</td>
</tr>
<tr>
<td>H</td>
<td>Hakan, Greenberg</td>
<td>1987</td>
</tr>
<tr>
<td>HF</td>
<td>Harold Fleming, ed., <em>Mother Tongue</em>, 1986–</td>
<td></td>
</tr>
<tr>
<td>HJ</td>
<td>Harry Johnston</td>
<td>1922</td>
</tr>
<tr>
<td>IE</td>
<td>Indo-European, Pokorny</td>
<td>1959</td>
</tr>
<tr>
<td>IP</td>
<td>Indo-Pacific, Greenberg</td>
<td>1971</td>
</tr>
<tr>
<td>IS</td>
<td>V. M. Illich-Svitych</td>
<td>1967</td>
</tr>
<tr>
<td>IST</td>
<td>Robert Shafer</td>
<td>1974</td>
</tr>
<tr>
<td>JB</td>
<td>John Bengtson</td>
<td>1986</td>
</tr>
<tr>
<td>JP</td>
<td>James Patrie</td>
<td>1982</td>
</tr>
<tr>
<td>JR</td>
<td>Johannes Rahder</td>
<td>1963</td>
</tr>
<tr>
<td>K</td>
<td>Khoisan, Greenberg</td>
<td>1963</td>
</tr>
<tr>
<td>KA</td>
<td>Kartvelian, Klimov</td>
<td>1964</td>
</tr>
<tr>
<td>KE</td>
<td>Khoisan Etymologies, Ruhlen</td>
<td>1987b</td>
</tr>
<tr>
<td>KER</td>
<td>Harold Fleming</td>
<td>1983a</td>
</tr>
<tr>
<td>KS</td>
<td>Kongo-Saharan, Gregersen</td>
<td>1972</td>
</tr>
<tr>
<td>LC</td>
<td>Morris Swadesh</td>
<td>1960</td>
</tr>
<tr>
<td>LN</td>
<td>Vaclav Blazek</td>
<td>1990</td>
</tr>
<tr>
<td>M</td>
<td>Mande, Mukarovsky</td>
<td>1966</td>
</tr>
<tr>
<td>MG</td>
<td>Macro-Ge, Greenberg</td>
<td>1987</td>
</tr>
<tr>
<td>MT</td>
<td>Macro-Tucanoan, Greenberg</td>
<td>1987</td>
</tr>
<tr>
<td>N</td>
<td>Nostratic, Illich-Svitych</td>
<td>1971–84</td>
</tr>
<tr>
<td>NA</td>
<td>Nahali, Kuiper</td>
<td>1962</td>
</tr>
<tr>
<td>NC</td>
<td>Niger-Congo, Greenberg</td>
<td>1963</td>
</tr>
<tr>
<td>ND</td>
<td>Na-Dene Notebook, Greenberg</td>
<td>1981</td>
</tr>
<tr>
<td>NK</td>
<td>Niger-Kordofanian, Greenberg</td>
<td>1963</td>
</tr>
<tr>
<td>NNG</td>
<td>North New Guinea, Greenberg</td>
<td>1971</td>
</tr>
<tr>
<td>NP</td>
<td>Norman-Paman, Black</td>
<td>1980</td>
</tr>
<tr>
<td>NS</td>
<td>Nilo-Saharan, Greenberg</td>
<td>1963</td>
</tr>
<tr>
<td>NSB</td>
<td>Nilo-Saharan, Bender</td>
<td>1980</td>
</tr>
<tr>
<td>NSC</td>
<td>Nostratic–Sino-Caucasian, Starostin</td>
<td>1991</td>
</tr>
<tr>
<td>NSD</td>
<td>Nilo-Saharan–Dravidian, Greenberg</td>
<td>1986</td>
</tr>
<tr>
<td>OL</td>
<td>Václav Blážek</td>
<td>1989</td>
</tr>
<tr>
<td>OS</td>
<td>Olga V. Stolbova</td>
<td>1987</td>
</tr>
<tr>
<td>P</td>
<td>Penutian, Greenberg</td>
<td>1987</td>
</tr>
</tbody>
</table>
### 14. Global Etymologies

<table>
<thead>
<tr>
<th>Code</th>
<th>Author</th>
<th>Year(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB</td>
<td>Paul Benedict</td>
<td>1975</td>
</tr>
<tr>
<td>PP</td>
<td>Paul Proulx</td>
<td>1984</td>
</tr>
<tr>
<td>R</td>
<td>Károly Rédei</td>
<td>1986–88</td>
</tr>
<tr>
<td>RB</td>
<td>Robert Blust</td>
<td>1980</td>
</tr>
<tr>
<td>RD</td>
<td>Robert Dixon</td>
<td>1980</td>
</tr>
<tr>
<td>SAK</td>
<td>Southern African Khoisan, Ruhlen</td>
<td>1987b</td>
</tr>
<tr>
<td>SB</td>
<td>S. Bhattacharya</td>
<td>1966</td>
</tr>
<tr>
<td>SC</td>
<td>Sino-Caucasian</td>
<td>1991a</td>
</tr>
<tr>
<td>SES</td>
<td>Southeast Surmic</td>
<td>1983b</td>
</tr>
<tr>
<td>SM</td>
<td>Stuart Mann</td>
<td>1984–88</td>
</tr>
<tr>
<td>SN</td>
<td>Sergei Nikolaev</td>
<td>1991</td>
</tr>
<tr>
<td>SNG</td>
<td>South New Guinea</td>
<td>1971</td>
</tr>
<tr>
<td>SS</td>
<td>Sergei Starostin</td>
<td>1984</td>
</tr>
<tr>
<td>ST</td>
<td>Sino-Tibetan</td>
<td>1972</td>
</tr>
<tr>
<td>SUL</td>
<td>Björn Collinder</td>
<td>1957</td>
</tr>
<tr>
<td>SWNG</td>
<td>Southwest New Guinea</td>
<td>1971</td>
</tr>
<tr>
<td>SY</td>
<td>S. Yoshitake</td>
<td>1934</td>
</tr>
<tr>
<td>T</td>
<td>Tasmanian</td>
<td>1976</td>
</tr>
<tr>
<td>TB</td>
<td>Tibeto-Burman</td>
<td>1985</td>
</tr>
<tr>
<td>U</td>
<td>Uralic</td>
<td>1977</td>
</tr>
<tr>
<td>UOL</td>
<td>Alfredo Trombetti</td>
<td>1905</td>
</tr>
<tr>
<td>VB</td>
<td>Václav Blažek</td>
<td>1988</td>
</tr>
<tr>
<td>VT</td>
<td>V. N. Toporov</td>
<td>1967</td>
</tr>
<tr>
<td>W</td>
<td>Werchikwar</td>
<td>1962</td>
</tr>
<tr>
<td>WM</td>
<td>Walter Müller</td>
<td>1975</td>
</tr>
<tr>
<td>WP</td>
<td>Alois Walde and Julius Pokorny</td>
<td>1930</td>
</tr>
<tr>
<td>WW</td>
<td>S. A. Wurm and B. Wilson</td>
<td>1975</td>
</tr>
<tr>
<td>Y</td>
<td>Yeniseian</td>
<td>1984</td>
</tr>
</tbody>
</table>

### REFERENCES


14. Global Etymologies


14. Global Etymologies


Key, Mary R. 1968. Comparative Tacanan Phonology. The Hague.


14. Global Etymologies


