

## BORROWED ENGLISH

DAN TILQUE

Beaverton, Oregon

English is known as a great borrowing language. Among words in the dictionary, roughly 25 percent are native English words, the rest derived from other languages. But how many languages has English borrowed from? The way to find out is to compile a list of words derived from as many different languages as possible.

Before counting, there are a number of issues to resolve. Different references have different etymologies for the same word. Should equal consideration be given to all of them? What exactly constitutes a language? How should language families, creoles, dialects, etc. be handled?

After much consideration, I settled on the following ground rules:

- Only one dictionary is to be used as reference. I chose Merriam-Webster's Collegiate Dictionary, Tenth Edition
- All words with etymologies are to be considered except the names of monetary units, ethnic groups, and languages
- Any language in the direct history of a word is allowed. However, only one language per word will be counted
- Languages of cognate words (indicated in etymologies by the phrase 'akin to') are not allowed
- English dialects are out, dialects of other languages are in
- Creoles, pidgins, and trade languages are to be counted
- Language families are included only where there is no word borrowed from any of the languages in that family. Excluded examples are Scandinavian, Algonquian, and Bantu
- Languages that have the same name as another language but with a chronological modifier in the name are considered part of that language. Examples are Old High German, Middle French, and New Greek

It turned out that the exception for English dialects was almost a nonissue. Except for words from Scots and one other word, English dialects are not identified by Merriam-Webster. That other word is LIMBO (the dance) which comes from the English of Trinidad and Barbados.

The rule allowing any language in a word's history leads to an unusual result. Gothic, an East Germanic language, is mentioned in several etymologies, but in all but one it is as a cognate. The exception is CZAR, a word normally associated with Russian.

However, the history of this word starts with Caesar which is borrowed via Greek into Gothic as kaiser and thence into Russian. So why is Gothic not in the etymology of kaiser? According to Merriam-Webster, kaiser comes from a "prehistoric German word" that was borrowed from Latin and Gothic does not appear there.

The restriction of only one language per word meant that four languages did not make it into my list. PARKA is the only word for Aleut as well as for Nenets, VINDALOO is from both Indo-Portuguese and Konkani, PUNKIE from New York Dutch and the Munsee dialect of Delaware, and VOODOO from Louisiana Creole and Ewe.

After I was well underway in researching this project, a book was published on exactly the same subject. That book was *The World in So Many Words* (1999, Houghton Mifflin) by Allan Metcalf. This article is not going to review that book, nor is it going to make a detailed comparison of the two lists. However, some points of comparison may be of interest.

Metcalf includes a number of words that come from English dialects, although he limits them to national level dialects. That is, he has words from American English, Australian English, etc. but not from, say, Cockney or Brooklynese. On the other hand, Merriam-Webster frequently identifies the dialects of many non-English languages where most other dictionaries do not.

Metcalf's research led him to disagree with Merriam-Webster about a number of words. For instance, Metcalf has MOJO as coming from Fula, a language spoken in Cameroon; Merriam-Webster merely says it's an African word and gives the Fulani word as a cognate. There are quite a few differences of this type, especially among words from African languages.

A different sort of disagreement is represented by BORA, a downslope wind in either the Adriatic Sea or Black Sea areas. Webster's derives it from the Trieste dialect of Italian which in turn gets it from Latin boreas (north wind). Metcalf acknowledges this as possible, but thinks it more likely that the source is Bulgarian.

Another difference is that Metcalf has several eponyms such as CASHMERE, SEQUOIA, BIKINI and HOOCH. Merriam-Webster generally does not mention languages in the etymologies of eponyms so they don't appear in the list below.

In light of the significant number of differences between the two lists, it is surprising that they have almost exactly the same length! Metcalf has 212 languages, this list has 216. But since they do have differences, it's possible to make a longer list by merging the two. Doing so gives about 240 languages, not counting English dialects or eponyms, and perhaps 255 with them.

In the list below, languages are the same as they are in the Tenth Collegiate, except that abbreviations have been spelled out. Dialects appear in parentheses immediately after the language name.



## ASIA

Akkadian: ziggurat  
 Ambonese: ailanthus  
 Arabic: xebec  
 Aramaic: abbot  
 Avestan: satem  
 Balti: polo  
 Bazaar Malay: orangutan  
 Bengali: dinghy  
 Bisayan: cogon  
 Chinese (Beijing): ginseng  
 Chinese (Fujian): bohea  
 Chinese (Guangdong): kumquat  
 Chinese (Xiamen): tea  
 Chinese pidgin English: chowchow  
 Deccan Hindi: ragi  
 Divehi: atoll  
 Evenki: shaman  
 Georgian: zerkova  
 Gujarati: banyan  
 Hebrew: behemoth  
 Hindi: bungalow  
 Indonesian Malay: upas  
 Japanese: haiku  
 Japanese (Okinawa): nunchaku  
 Javanese: lahar  
 Kannada: dhole  
 Kazakh: barchan  
 Khmer: kouprey  
 Konkani: vindaloo  
 Korean: kimchi  
 Lepcha: serow  
 Malay: bamboo  
 Malay (Java): gourami  
 Malayalam: teak  
 Marathi: mongoose  
 Mishmi: takin  
 Mongolian: argali  
 Nenets: parka  
 Nepali: tahr  
 Pali: ginger  
 Panjabi: urial  
 Persian: bazaar  
 Philippine Spanish: bolo  
 Sanskrit: avatar  
 Sinhalese: beriberi  
 Sundanese: muntjac  
 Syriac: arsenic  
 Tagalog: boondocks  
 Tamil: betel  
 Telugu: bandicoot  
 Thai: bong (waterpipe)  
 Tibetan: yak  
 Turkish: yogurt  
 Urdu: nabob  
 Vietnamese: Tet

## OCEANIA

Adnyamadhanha: euro (wallaroo)  
 Dharuk: wombat  
 Fore: kuru  
 Guugu Yimidhirr: kangaroo  
 Hawaiian: luau  
 Jagara: dilly bag  
 Maori: kiwi  
 Marquesan: tiki  
 Nyungar: jarrah  
 Samoan: lavalava  
 Tahitian: tattoo  
 Tongan: taboo  
 Wik Munkan: taipan (snake)  
 Wiradhuri: kookaburra  
 Wuywurung: yabber  
 Yolngu: didgeridoo  
 Yuwaalaraay: budgerigar

## AMERICAS

American Yiddish: boychick  
 Araucanian: poncho  
 Arawak: iguana  
 Arawak of Lesser Antilles: anole  
 Aymara: alpaca  
 Brazilian Portuguese: maxixe  
 Cahuilla: chuckwalla  
 Canadian French: lacrosse  
 Carib: caiman  
 Catawba: yaupon  
 Chinook Jargon: camas  
 Choctaw: bayou

|                                     |                                    |
|-------------------------------------|------------------------------------|
| Cree: pemmican                      | Miskito: dory                      |
| Creek: tupelo                       | Nahuatl: atlatl                    |
| Creek (Florida): coontie            | Narraganset: quahog                |
| Cumanagoto: divi-divi               | Navajo: hogan                      |
| Cuna: guan                          | Nootka: potlatch                   |
| Dakota: tepee                       | North Carolina Algonquian: pocosin |
| Delaware (Munsee): punkie           | Northern Straits: sockeye          |
| Delaware (Unami): shoepac           | Ojibwa: totem                      |
| Eastern Abenaki: wigwam             | Pennsylvania German: dunk          |
| Fox: wickiup                        | Quechua: puma                      |
| Guarani: jaguarundi                 | Rumsen: abalone                    |
| Gullah: tabby (cement, from shells) | Shawnee: wapiti                    |
| Haitian Creole: merengue            | Shoshone: pogonip                  |
| Halkomelen: coho                    | Shuswap: kokanee                   |
| Hopi: piki                          | Southern Paiute: sego lily         |
| Inuit: igloo                        | Taino: hammock                     |
| Inuit (Greenland): anorak           | Tarascan: huarache                 |
| Louisiana French: etouffee          | Tupi: tapioca                      |
| Lower Chehalis: chinook (salmon)    | Unquachog: sea puss                |
| Lower Chinook: salal                | Virginia Algonquian: raccoon       |
| Lushootseed: geoduck                | Western Abenaki: hackmatack        |
| Massachuset: wampum                 | Yucatec: cenote                    |
| Mexican Spanish: saguaro            | Yupic: mukluk                      |
| Micmac: caribou                     |                                    |