

NAME

I18N::LangTags::List -- tags and names for human languages

SYNOPSIS

```
use I18N::LangTags::List;
print "Parlez-vous... ", join(' ',
I18N::LangTags::List::name('elx') || 'unknown_language',
I18N::LangTags::List::name('ar-Kw') || 'unknown_language',
I18N::LangTags::List::name('en') || 'unknown_language',
I18N::LangTags::List::name('en-CA') || 'unknown_language',
), "?\n";
```

prints:

Parlez-vous... Elamite, Kuwait Arabic, English, Canadian English?

DESCRIPTION

This module provides a function `I18N::LangTags::List::name(langtag)` that takes a language tag (see `I18N::LangTags`) and returns the best attempt at an English name for it, or `undef` if it can't make sense of the tag.

The function `I18N::LangTags::List::name(...)` is not exported.

This module also provides a function `I18N::LangTags::List::is_decent(langtag)` that returns true iff the language tag is syntactically valid and is for general use (like “fr” or “fr-ca”, below). That is, it returns false for tags that are syntactically invalid and for tags, like “aus”, that are listed in brackets below. This function is not exported.

The map of tags-to-names that it uses is accessible as `%I18N::LangTags::List::Name`, and it's the same as the list that follows in this documentation, which should be useful to you even if you don't use this module.

ABOUT LANGUAGE TAGS

Internet language tags, as defined in RFC 3066, are a formalism for denoting human languages. The two-letter ISO 639-1 language codes are well known (as “en” for English), as are their forms when qualified by a country code (“en-US”). Less well-known are the arbitrary-length non-ISO codes (like “i-mingo”), and the recently (in 2001) introduced three-letter ISO-639-2 codes.

Remember these important facts:

- Language tags are not locale IDs. A locale ID is written with a “_” instead of a “-”, (almost?) always matches `m/\w\w_\w\w\b/`, and *means* something different than a language tag. A language tag denotes a language. A locale ID denotes a language *as used in* a particular place, in combination with non-linguistic location-specific information such as what currency is used there. Locales *also* often denote character set information, as in “en_US.ISO8859-1”.
- Language tags are not for computer languages.
- “Dialect” is not a useful term, since there is no objective criterion for establishing when two language-forms are dialects of each other, or are separate languages.
- Language tags are not case-sensitive. en-US, en-us, En-U_s, etc., are all the same tag, and denote the same language.
- Not every language tag really refers to a single language. Some language tags refer to conditions: i-default (system-message text in English plus maybe other languages), und (undetermined language). Others (notably lots of the three-letter codes) are bibliographic tags that classify whole groups of languages, as with cus “Cushitic (Other)” (i.e., a language that has been classed as Cushtic, but which has no more specific code) or the even less linguistically coherent sai for “South American Indian (Other)”. Though useful in bibliography, **SUCH TAGS ARE NOT FOR GENERAL USE**. For further guidance, email me.

- Language tags are not country codes. In fact, they are often distinct codes, as with language tag `ja` for Japanese, and ISO 3166 country code `.jp` for Japan.

LIST OF LANGUAGES

The first part of each item is the language tag, between `{...}`. It is followed by an English name for the language or language-group. Language tags that I judge to be not for general use, are bracketed.

This list is in alphabetical order by English name of the language.

`{ab}` : Abkhazian
eq Abkhaz

`{ace}` : Achinese

`{ach}` : Acoli

`{ada}` : Adangme

`{ady}` : Adyghe

eq Adygei

`{aa}` : Afar

`{afh}` : Afrihili

(Artificial)

`{af}` : Afrikaans

`[{afa}` : Afro-Asiatic (Other)]

`{ak}` : Akan

(Formerly “aka”.)

`{akk}` : Akkadian

(Historical)

`{sq}` : Albanian

`{ale}` : Aleut

`[{alg}` : Algonquian languages]

NOT Algonquin!

`[{tut}` : Altaic (Other)]

`{am}` : Amharic

NOT Aramaic!

`{i-ami}` : Ami

eq Amis. eq 'Amis. eq Pangca.

`[{apa}` : Apache languages]

`{ar}` : Arabic

Many forms are mutually un-intelligible in spoken media. Notable forms: `{ar-ae}` UAE Arabic; `{ar-bh}` Bahrain Arabic; `{ar-dz}` Algerian Arabic; `{ar-eg}` Egyptian Arabic; `{ar-iq}` Iraqi Arabic; `{ar-jo}` Jordanian Arabic; `{ar-kw}` Kuwait Arabic; `{ar-lb}` Lebanese Arabic; `{ar-ly}` Libyan Arabic; `{ar-ma}` Moroccan Arabic; `{ar-om}` Omani Arabic; `{ar-qa}` Qatari Arabic; `{ar-sa}` Sauda Arabic; `{ar-sy}` Syrian Arabic; `{ar-tn}` Tunisian Arabic; `{ar-ye}` Yemen Arabic.

`{arc}` : Aramaic

NOT Amharic! NOT Samaritan Aramaic!

`{arp}` : Arapaho

`{arn}` : Araucanian

`{arw}` : Arawak

`{hy}` : Armenian

`{an}` : Aragonese

`[{art}` : Artificial (Other)]

{ast} : Asturian
 eq Bable.

{as} : Assamese
 [{ath} : Athapascan languages]
 eq Athabaskan. eq Athapaskan. eq Athabascan.

[{aus} : Australian languages]
 [{map} : Austronesian (Other)]

{av} : Avaric
 (Formerly “ava”.)

{ae} : Avestan
 eq Zend

{awa} : Awadhi
 {ay} : Aymara
 {az} : Azerbaijani
 eq Azeri

 Notable forms: {az-Arab} Azerbaijani in Arabic script; {az-Cyrl} Azerbaijani in Cyrillic script; {az-Latn} Azerbaijani in Latin script.

{ban} : Balinese
 [{bat} : Baltic (Other)]

{bal} : Baluchi
 {bm} : Bambara
 (Formerly “bam”.)

[{bai} : Bamileke languages]
 {bad} : Banda
 [{bnt} : Bantu (Other)]

{bas} : Basa
 {ba} : Bashkir
 {eu} : Basque
 {btk} : Batak (Indonesia)
 {bej} : Beja
 {be} : Belarusian
 eq Belarussian. eq Byelarussian. eq Belorussian. eq Byelorussian. eq White Russian. eq White Ruthenian. NOT Ruthenian!

{bem} : Bemba
 {bn} : Bengali
 eq Bangla.

[{ber} : Berber (Other)]
 {bho} : Bhojpuri
 {bh} : Bihari
 {bik} : Bikol
 {bin} : Bini
 {bi} : Bislama
 eq Bichelamar.

{bs} : Bosnian
 {bra} : Braj
 {br} : Breton
 {bug} : Buginese
 {bg} : Bulgarian

{i-bnn} : Bunun
 {bua} : Buriat
 {my} : Burmese
 {cad} : Caddo
 {car} : Carib
 {ca} : Catalan
 eq Catalán. eq Catalanian.
 [{cau} : Caucasian (Other)]
 {ceb} : Cebuano
 [{cel} : Celtic (Other)]
 Notable forms: {cel-gaulish} Gaulish (Historical)
 [{cai} : Central American Indian (Other)]
 {chg} : Chagatai
 (Historical?)
 [{cmc} : Chamic languages]
 {ch} : Chamorro
 {ce} : Chechen
 {chr} : Cherokee
 eq Tsalagi
 {chy} : Cheyenne
 {chb} : Chibcha
 (Historical) NOT Chibchan (which is a language family).
 {ny} : Chichewa
 eq Nyanja. eq Chinyanja.
 {zh} : Chinese
 Many forms are mutually un-intelligible in spoken media. Notable forms: {zh-Hans} Chinese, in simplified script; {zh-Hant} Chinese, in traditional script; {zh-tw} Taiwan Chinese; {zh-cn} PRC Chinese; {zh-sg} Singapore Chinese; {zh-mo} Macau Chinese; {zh-hk} Hong Kong Chinese; {zh-guoyu} Mandarin [Putonghua/Guoyu]; {zh-hakka} Hakka [formerly “i-hakka”]; {zh-min} Hokkien; {zh-min-nan} Southern Hokkien; {zh-wuu} Shanghaiese; {zh-xiang} Hunanese; {zh-gan} Gan; {zh-yue} Cantonese.
 {chn} : Chinook Jargon
 eq Chinook Wawa.
 {chp} : Chipewyan
 {cho} : Choctaw
 {cu} : Church Slavic
 eq Old Church Slavonic.
 {chk} : Chuukese
 eq Trukese. eq Chuuk. eq Truk. eq Ruk.
 {cv} : Chuvash
 {cop} : Coptic
 {kw} : Cornish
 {co} : Corsican
 eq Corse.
 {cr} : Cree
 NOT Creek! (Formerly “cre”.)
 {mus} : Creek
 NOT Cree!

[{cpe} : English-based Creoles and pidgins (Other)]
 [{cpf} : French-based Creoles and pidgins (Other)]
 [{cpp} : Portuguese-based Creoles and pidgins (Other)]
 [{crp} : Creoles and pidgins (Other)]
 {hr} : Croatian
 eq Croat.

 [{cus} : Cushitic (Other)]
 {cs} : Czech
 {dak} : Dakota
 eq Nakota. eq Latoka.

 {da} : Danish
 {dar} : Dargwa
 {day} : Dayak
 {i-default} : Default (Fallthru) Language
 Defined in RFC 2277, this is for tagging text (which must include English text, and
 might/should include text in other appropriate languages) that is emitted in a context where
 language-negotiation wasn't possible — in SMTP mail failure messages, for example.

 {del} : Delaware
 {din} : Dinka
 {dv} : Divehi
 eq Maldivian. (Formerly “div”.)

 {doi} : Dogri
 NOT Dogrib!

 {dgr} : Dogrib
 NOT Dogri!

 [{dra} : Dravidian (Other)]
 {dua} : Duala
 {nl} : Dutch
 eq Netherlander. Notable forms: {nl-nl} Netherlands Dutch; {nl-be} Belgian Dutch.

 {dum} : Middle Dutch (ca.1050-1350)
 (Historical)

 {dyu} : Dyula
 {dz} : Dzongkha
 {efi} : Efik
 {egy} : Ancient Egyptian
 (Historical)

 {eka} : Ekajuk
 {elx} : Elamite
 (Historical)

 {en} : English
 Notable forms: {en-au} Australian English; {en-bz} Belize English; {en-ca} Canadian
 English; {en-gb} UK English; {en-ie} Irish English; {en-jm} Jamaican English; {en-nz} New
 Zealand English; {en-ph} Philippine English; {en-tt} Trinidad English; {en-us} US English;
 {en-za} South African English; {en-zw} Zimbabwe English.

 {enm} : Old English (1100-1500)
 (Historical)

 {ang} : Old English (ca.450-1100)
 eq Anglo-Saxon. (Historical)

{i-enochian} : Enochian (Artificial)
 {myv} : Erzya
 {eo} : Esperanto
 (Artificial)
 {et} : Estonian
 {ee} : Ewe
 (Formerly “ewe”.)
 {ewo} : Ewondo
 {fan} : Fang
 {fat} : Fanti
 {fo} : Faroese
 {fj} : Fijian
 {fi} : Finnish
 [{fiu} : Finno-Ugrian (Other)]
 eq Finno-Ugric. NOT Ugaritic!
 {fon} : Fon
 {fr} : French
 Notable forms: {fr-fr} France French; {fr-be} Belgian French; {fr-ca} Canadian French; {fr-
 ch} Swiss French; {fr-lu} Luxembourg French; {fr-mc} Monaco French.
 {frm} : Middle French (ca.1400-1600)
 (Historical)
 {fro} : Old French (842-ca.1400)
 (Historical)
 {fy} : Frisian
 {fur} : Friulian
 {ff} : Fulah
 (Formerly “ful”.)
 {gaa} : Ga
 {gd} : Scots Gaelic
 NOT Scots!
 {gl} : Gallegan
 eq Galician
 {lg} : Ganda
 (Formerly “lug”.)
 {gay} : Gayo
 {gba} : Gbaya
 {gez} : Geez
 eq Ge'ez
 {ka} : Georgian
 {de} : German
 Notable forms: {de-at} Austrian German; {de-be} Belgian German; {de-ch} Swiss German;
 {de-de} Germany German; {de-li} Liechtenstein German; {de-lu} Luxembourg German.
 {gmh} : Middle High German (ca.1050-1500)
 (Historical)
 {goh} : Old High German (ca.750-1050)
 (Historical)
 [{gem} : Germanic (Other)]

{gil} : Gilbertese
{gon} : Gondi
{gor} : Gorontalo
{got} : Gothic
 (Historical)

{grb} : Grebo
{grc} : Ancient Greek
 (Historical) (Until 15th century or so.)

{el} : Modern Greek
 (Since 15th century or so.)

{gn} : Guarani
 Guaraní

{gu} : Gujarati
{gwi} : Gwich'in
 eq Gwichin

{hai} : Haida
{ht} : Haitian
 eq Haitian Creole

{ha} : Hausa
{haw} : Hawaiian
 Hawai'ian

{he} : Hebrew
 (Formerly "iw".)

{hz} : Herero
{hil} : Hiligaynon
{him} : Himachali
{hi} : Hindi
{ho} : Hiri Motu
{hit} : Hittite
 (Historical)

{hmn} : Hmong
{hu} : Hungarian
{hup} : Hupa
{iba} : Iban
{is} : Icelandic
{io} : Ido
 (Artificial)

{ig} : Igbo
 (Formerly "ibo".)

{ijo} : Ijo
{ilo} : Iloko
[inc] : Indic (Other)
[ine] : Indo-European (Other)
{id} : Indonesian
 (Formerly "in".)

{inh} : Ingush
{ia} : Interlingua (International Auxiliary Language Association)
 (Artificial) NOT Interlingue!

{ie} : Interlingue
(Artificial) NOT Interlingua!

{iu} : Inuktitut
A subform of “Eskimo”.

{ik} : Inupiaq
A subform of “Eskimo”.

[[ira] : Iranian (Other)]

{ga} : Irish

{mga} : Middle Irish (900-1200)
(Historical)

{sga} : Old Irish (to 900)
(Historical)

[[iro] : Iroquoian languages]

{it} : Italian
Notable forms: {it-it} Italy Italian; {it-ch} Swiss Italian.

{ja} : Japanese
(NOT “jp”!)

{jv} : Javanese
(Formerly “jw” because of a typo.)

{jrb} : Judeo-Arabic

{jpr} : Judeo-Persian

{kbd} : Kabardian

{kab} : Kabyle

{kac} : Kachin

{kl} : Kalaallisut
eq Greenlandic “Eskimo”

{xal} : Kalmyk

{kam} : Kamba

{kn} : Kannada
eq Kanarese. NOT Canadian!

{kr} : Kanuri
(Formerly “kau”.)

{krc} : Karachay-Balkar

{kaa} : Kara-Kalpak

{kar} : Karen

{ks} : Kashmiri

{csb} : Kashubian
eq Kashub

{kaw} : Kawi

{kk} : Kazakh

{kha} : Khasi

{km} : Khmer
eq Cambodian. eq Kampuchean.

[[khi] : Khoisan (Other)]

{kho} : Khotanese

{ki} : Kikuyu
eq Gikuyu.

{kmb} : Kimbundu
{rw} : Kinyarwanda
{ky} : Kirghiz
{i-klingon} : Klingon
{kv} : Komi
{kg} : Kongo
(Formerly “kon”.)

{kok} : Konkani
{ko} : Korean
{kos} : Kosraean
{kpe} : Kpelle
{kro} : Kru
{kj} : Kuanyama
{kum} : Kumyk
{ku} : Kurdish
{kru} : Kurukh
{kut} : Kutenai
{lad} : Ladino
eq Judeo-Spanish. NOT Ladin (a minority language in Italy).

{lah} : Lahnda
NOT Lamba!

{lam} : Lamba
NOT Lahnda!

{lo} : Lao
eq Laotian.

{la} : Latin
(Historical) NOT Ladin! NOT Ladino!

{lv} : Latvian
eq Lettish.

{lb} : Letzeburgesch
eq Luxemburgian, eq Luxemburger. (Formerly “i-lux”.)

{lez} : Lezghian
{li} : Limburgish
eq Limburger, eq Limburgan. NOT Letzeburgesch!

{ln} : Lingala
{lt} : Lithuanian
{nds} : Low German
eq Low Saxon. eq Low German. eq Low Saxon.

{art-lojban} : Lojban (Artificial)
{loz} : Lozi
{lu} : Luba-Katanga
(Formerly “lub”.)

{lua} : Luba-Lulua
{lui} : Luiseno
eq Luiseño.

{lun} : Lunda
{luo} : Luo (Kenya and Tanzania)
{lus} : Lushai

{mk} : Macedonian
eq the modern Slavic language spoken in what was Yugoslavia. NOT the form of Greek spoken in Greek Macedonia!

{mad} : Madurese
{mag} : Magahi
{mai} : Maithili
{mak} : Makasar
{mg} : Malagasy
{ms} : Malay
NOT Malayalam!

{ml} : Malayalam
NOT Malay!

{mt} : Maltese
{mnc} : Manchu
{mdr} : Mandar
NOT Mandarin!

{man} : Mandingo
{mni} : Manipuri
eq Meithei.

[[mno] : Manobo languages]
{gv} : Manx
{mi} : Maori
NOT Mari!

{mr} : Marathi
{chm} : Mari
NOT Maori!

{mh} : Marshall
eq Marshallese.

{mwr} : Marwari
{mas} : Masai
[[myn] : Mayan languages]
{men} : Mende
{mic} : Micmac
{min} : Minangkabau
{i-mingo} : Mingo
eq the Iroquoian language West Virginia Seneca. NOT New York Seneca!

[[mis] : Miscellaneous languages]
Don't use this.

{moh} : Mohawk
{mdf} : Moksha
{mo} : Moldavian
eq Moldovan.

[[mkh] : Mon-Khmer (Other)]
{lol} : Mongo
{mn} : Mongolian
eq Mongol.

{mos} : Mossi
[[mul] : Multiple languages]
Not for normal use.

[{mun} : Munda languages]
 {nah} : Nahuatl
 {nap} : Neapolitan
 {na} : Nauru
 {nv} : Navajo
 eq Navaho. (Formerly “i-navajo”.)

 {nd} : North Ndebele
 {nr} : South Ndebele
 {ng} : Ndonga
 {ne} : Nepali
 eq Nepalese. Notable forms: {ne-np} Nepal Nepali; {ne-in} India Nepali.

 {new} : Newari
 {nia} : Nias
 [{nic} : Niger-Kordofanian (Other)]
 [{ssa} : Nilo-Saharan (Other)]
 {niu} : Niuean
 {nog} : Nogai
 {non} : Old Norse
 (Historical)

 [{nai} : North American Indian]
 Do not use this.

 {no} : Norwegian
 Note the two following forms:

 {nb} : Norwegian Bokmal
 eq Bokmål, (A form of Norwegian.) (Formerly “no-bok”.)

 {nn} : Norwegian Nynorsk
 (A form of Norwegian.) (Formerly “no-nyn”.)

 [{nub} : Nubian languages]
 {nym} : Nyamwezi
 {nyn} : Nyankole
 {nyo} : Nyoro
 {nzi} : Nzima
 {oc} : Occitan (post 1500)
 eq Provençal, eq Provençal

 {oj} : Ojibwa
 eq Ojibwe. (Formerly “oji”.)

 {or} : Oriya
 {om} : Oromo
 {osa} : Osage
 {os} : Ossetian; Ossetic
 [{oto} : Otomian languages]
 Group of languages collectively called “Otomi”.

 {pal} : Pahlavi
 eq Pahlevi

 {i-pwn} : Paiwan
 eq Pariwan

 {pau} : Palauan
 {pi} : Pali
 (Historical?)

{pam} : Pampanga
 {pag} : Pangasinan
 {pa} : Panjabi
 eq Punjabi

 {pap} : Papiamentu
 eq Papiamentu.

 [{paa} : Papuan (Other)]
 {fa} : Persian
 eq Farsi. eq Iranian.

 {peo} : Old Persian (ca.600-400 B.C.)
 [{phi} : Philippine (Other)]
 {phn} : Phoenician
 (Historical)

 {pon} : Pohnpeian
 NOT Pompeiian!

 {pl} : Polish
 {pt} : Portuguese
 eq Portugese. Notable forms: {pt-pt} Portugal Portuguese; {pt-br} Brazilian Portuguese.

 [{pra} : Prakrit languages]
 {pro} : Old Provençal (to 1500)
 eq Old Provençal. (Historical.)

 {ps} : Pushto
 eq Pashto. eq Pushtu.

 {qu} : Quechua
 eq Quecha.

 {rm} : Raeto-Romance
 eq Romansh.

 {raj} : Rajasthani
 {rap} : Rapanui
 {rar} : Rarotongan
 [{qaa - qtz} : Reserved for local use.]
 [{roa} : Romance (Other)]
 NOT Romanian! NOT Romany! NOT Romansh!

 {ro} : Romanian
 eq Rumanian. NOT Romany!

 {rom} : Romany
 eq Rom. NOT Romanian!

 {rn} : Rundi
 {ru} : Russian
 NOT White Russian! NOT Rusyn!

 [{sal} : Salishan languages]
 Large language group.

 {sam} : Samaritan Aramaic
 NOT Aramaic!

 {se} : Northern Sami
 eq Lappish. eq Lapp. eq (Northern) Saami.

{sma} : Southern Sami
 {smn} : Inari Sami
 {smj} : Lule Sami
 {sms} : Skolt Sami
 [{smi} : Sami languages (Other)]
 {sm} : Samoan
 {sad} : Sandawe
 {sg} : Sango
 {sa} : Sanskrit
 (Historical)
 {sat} : Santali
 {sc} : Sardinian
 eq Sard.
 {sas} : Sasak
 {sco} : Scots
 NOT Scots Gaelic!
 {sel} : Selkup
 [{sem} : Semitic (Other)]
 {sr} : Serbian
 eq Serb. NOT Sorbian.
 Notable forms: {sr-Cyrl} : Serbian in Cyrillic script; {sr-Latn} : Serbian in Latin script.
 {srr} : Serer
 {shn} : Shan
 {sn} : Shona
 {sid} : Sidamo
 {sgn-...} : Sign Languages
 Always use with a subtag. Notable forms: {sgn-gb} British Sign Language (BSL); {sgn-ie} Irish Sign Language (ESL); {sgn-ni} Nicaraguan Sign Language (ISN); {sgn-us} American Sign Language (ASL).
 (And so on with other country codes as the subtag.)
 {bla} : Siksika
 eq Blackfoot. eq Pikanii.
 {sd} : Sindhi
 {si} : Sinhalese
 eq Sinhala.
 [{sit} : Sino-Tibetan (Other)]
 [{sio} : Siouan languages]
 {den} : Slave (Athapascan)
 (“Slavey” is a subform.)
 [{sla} : Slavic (Other)]
 {sk} : Slovak
 eq Slovakian.
 {sl} : Slovenian
 eq Slovene.
 {sog} : Sogdian
 {so} : Somali
 {son} : Songhai

{snk} : Soninke
 {wen} : Sorbian languages
 eq Wendish. eq Sorb. eq Lusatian. eq Wend. NOT Venda! NOT Serbian!
 {nso} : Northern Sotho
 {st} : Southern Sotho
 eq Sutu. eq Sesotho.
 [{sai} : South American Indian (Other)]
 {es} : Spanish
 Notable forms: {es-ar} Argentine Spanish; {es-bo} Bolivian Spanish; {es-cl} Chilean Spanish;
 {es-co} Colombian Spanish; {es-do} Dominican Spanish; {es-ec} Ecuadorian Spanish; {es-es}
 Spain Spanish; {es-gt} Guatemalan Spanish; {es-hn} Honduran Spanish; {es-mx} Mexican
 Spanish; {es-pa} Panamanian Spanish; {es-pe} Peruvian Spanish; {es-pr} Puerto Rican
 Spanish; {es-py} Paraguay Spanish; {es-sv} Salvadoran Spanish; {es-us} US Spanish; {es-uy}
 Uruguayan Spanish; {es-ve} Venezuelan Spanish.
 {suk} : Sukuma
 {sux} : Sumerian
 (Historical)
 {su} : Sundanese
 {sus} : Susu
 {sw} : Swahili
 eq Kiswahili
 {ss} : Swati
 {sv} : Swedish
 Notable forms: {sv-se} Sweden Swedish; {sv-fi} Finland Swedish.
 {syr} : Syriac
 {tl} : Tagalog
 {ty} : Tahitian
 [{tai} : Tai (Other)]
 NOT Thai!
 {tg} : Tajik
 {tmh} : Tamashek
 {ta} : Tamil
 {i-tao} : Tao
 eq Yami.
 {tt} : Tatar
 {i-tay} : Tayal
 eq Atayal. eq Atayan.
 {te} : Telugu
 {ter} : Tereno
 {tet} : Tetum
 {th} : Thai
 NOT Tai!
 {bo} : Tibetan
 {tig} : Tigre
 {ti} : Tigrinya
 {tem} : Timne
 eq Themne. eq Timene.
 {tiv} : Tiv

{tli} : Tlingit
{tpi} : Tok Pisin
{tkl} : Tokelau
{tog} : Tonga (Nyasa)
NOT Tsonga!

{to} : Tonga (Tonga Islands)
(Pronounced “Tong-a”, not “Tong-ga”)
NOT Tsonga!

{tsi} : Tsimshian
eq Sm’algyax

{ts} : Tsonga
NOT Tonga!

{i-tsu} : Tsou
{tn} : Tswana
Same as Setswana.

{tum} : Tumbuka
[tup] : Tupi languages
{tr} : Turkish
(Typically in Roman script)

{ota} : Ottoman Turkish (1500-1928)
(Typically in Arabic script) (Historical)

{crh} : Crimean Turkish
eq Crimean Tatar

{tk} : Turkmen
eq Turkmeni.

{tvl} : Tuvalu
{tyv} : Tuvinian
eq Tuvan. eq Tuvin.

{tw} : Twi
{udm} : Udmurt
{uga} : Ugaritic
NOT Ugric!

{ug} : Uighur
{uk} : Ukrainian
{umb} : Umbundu
{und} : Undetermined
Not a tag for normal use.

{ur} : Urdu
{uz} : Uzbek
eq Ozbek
Notable forms: {uz-Cyrl} Uzbek in Cyrillic script; {uz-Latn} Uzbek in Latin script.

{vai} : Vai
{ve} : Venda
NOT Wendish! NOT Wend! NOT Avestan! (Formerly “ven”.)

{vi} : Vietnamese
eq Viet.

{vo} : Volapuk
 eq Volapük. (Artificial)

{vot} : Votic
 eq Votian. eq Vod.

[[wak] : Wakashan languages]

{wa} : Walloon

{wal} : Walamo
 eq Wolaytta.

{war} : Waray
 Presumably the Philippine language Waray-Waray (Samareño), not the smaller Philippine language Waray Sorsogon, nor the extinct Australian language Waray.

{was} : Washo
 eq Washoe

{cy} : Welsh

{wo} : Wolof

{x-...} : Unregistered (Semi-Private Use)
 “x-” is a prefix for language tags that are not registered with ISO or IANA. Example, x-double-dutch

{xh} : Xhosa

{sah} : Yakut

{yao} : Yao
 (The Yao in Malawi?)

{yap} : Yapese
 eq Yap

{ii} : Sichuan Yi

{yi} : Yiddish
 Formerly “ji”. Usually in Hebrew script.
 Notable forms: {yi-latn} Yiddish in Latin script

{yo} : Yoruba

[[ypk] : Yupik languages]
 Several “Eskimo” languages.

{znd} : Zande

[[zap] : Zapotec]
 (A group of languages.)

{zen} : Zenaga
 NOT Zend.

{za} : Zhuang

{zu} : Zulu

{zun} : Zuni
 eq Zuñi

SEE ALSO

[I18N::LangTags](#) and its “See Also” section.

COPYRIGHT AND DISCLAIMER

Copyright (c) 2001+ Sean M. Burke. All rights reserved.

You can redistribute and/or modify this document under the same terms as Perl itself.

This document is provided in the hope that it will be useful, but without any warranty; without even the implied warranty of accuracy, authoritativeness, completeness, merchantability, or fitness for a particular purpose.

Email any corrections or questions to me.

AUTHOR

Sean M. Burke, sburke@cpan.org