

American Community Survey Redesign of Language-Spoken-at-Home Data, 2016

Christine P. Gambino
Social, Economic and Housing Statistics Division
U.S. Census Bureau

SEHSD Working Paper Number 2018-31

This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on methodological or operational issues are those of the author and are not necessarily those of the U.S. Census Bureau. Any error or omissions are the sole responsibility of the author. All data are subject to error arising from a variety of sources, including sampling error, non-sampling error, modeling error, and other sources of error.

The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. CBDRB-FY19-ROSS-B0027

Introduction

The [American Community Survey](#) (ACS) is the most widely used source of language data in the United States, largely because it is the only survey to provide language and English-speaking ability data for local communities. Language data users are a diverse group and include:

- Government agencies complying with laws protecting non-English language speakers or planning for future language resource needs.
- Businesses who utilize language data to make marketing plans or hiring decisions.
- Linguists, demographers, and other researchers.

In an effort to better serve these data users' needs, the U.S. Census Bureau made some language data coding and presentation improvements in 2016. The goals of this overhaul included the following:

- 1) Produce more precise and granular language data without placing additional burden on ACS respondents or data collection operations.
- 2) Provide better data products for a variety of users across different levels of detail.
- 3) Conform ACS language data to the standards most widely used by translators.
- 4) Maintain comparability over time with past language data, whenever possible.

The changes seen in the 2016 ACS language data and beyond were a result of a process that took more than four years. The need for these updates was first noticed after the 2010 Census, when subject matter experts in the Education and Social Stratification Branch of the Social (ESSB), Economic and Housing Statistics Division identified the need for more uniform, standardized coding. Between 2013 and 2015, the new language code list was created. Between 2015 and 2017, implementation of the changes into ACS production took place. New data products for individual and household language data were designed in 2016, and were released to the public in 2017 and early 2018. This report describes the processes, decision points, and challenges involved in this redesign.

Making Language Labels More Precise and Understandable

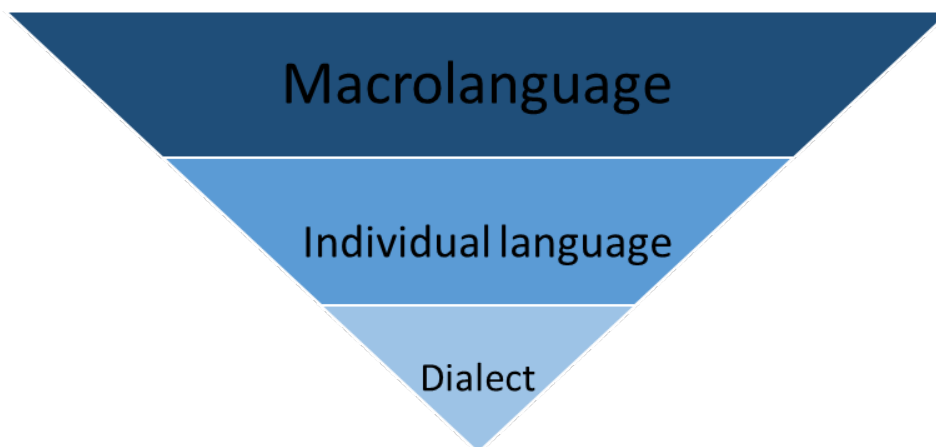
One of goals of the redesign was to have a more standard and transparent way of deciding what counts as a language, and what to call that language. American Community Survey measures language use with a series of three questions:

- Do you speak a language other than English at home?
- What is this language?
 - Specify: _____
- How well do you speak English?
 - Very well
 - Well
 - Not well
 - Not at all

Respondents who speak a language other than English at home specify the language in writing. The responses are recorded and then coded at Census Bureau headquarters, first by computer-matching responses to a compiled list of previous responses (called the “autocoder”), then, if there is no match, by staff who examine the write-ins and assign codes (“clerical coding”). If multiple languages are listed, only the first language is coded.

The [International Organization for Standardization’s ISO-639-3 standard](http://www-01.sil.org/iso639-3/standard) was chosen as the basis for the updated coding scheme because it focuses on the individual language level, yet also includes macrolanguages (see Figure 1). The ISO-639-3 standard does not include dialects. Languages are followed with a three-letter unique identifier, a convention which has been adopted by the ACS, as it is helpful in differentiating languages that with similar or the same name. Other related standards such as ISO-629-2 and ISO-639-5 focus on higher or lower levels of classification than the individual language level. For more information, see <http://www-01.sil.org/iso639-3/scope.asp>.

Figure 1. Hierarchy of linguistic categories.



A macrolanguage is a group of individual languages which are closely related to each other and are considered as a single language in certain contexts. Individual languages are considered to be closely related, if not always mutually intelligible. For example, Chinese (chi) is a macrolanguage that contains individual languages such as Mandarin Chinese (man), Min Nan Chinese (spoken in Taiwan), and Yue Chinese (Cantonese). Dialects within an individual language are considered to be mutually intelligible with each other.

Beginning with 2016 1-year and 2012-2016 5-year ACS data, language coding was changed so that individual language codes reflect languages in accord with the ISO-639-3 standard. This enables the Census Bureau to present languages in terms that are understood by linguists and translators, and also helps ensure that the specificity and precision of the data coding process is the same whether a language has millions of speakers or just a few. All ACS language codes reflect standard individual languages or macrolanguages, except for a few that reflect common responses that cannot be classified as individual languages (i.e., country names or language families). Dialects below the individual language

level are not recorded separately, but within the appropriate ISO-639-3 language. The full list of language codes is in Appendix A – American Community Survey Language Code List. Macrolanguages on the 2016 list in Appendix A are noted as such.

At first, it was suggested that all codes should be kept at the individual language level. It quickly became apparent that many people report speaking macrolanguages, rather than individual languages. Depending on the macrolanguage, this is more or less true. In 2016, 95% of speakers of any Arabic language could not be coded as an individual Arabic language, but could be coded as the Arabic macrolanguage. About 63% of Chinese speakers could not be coded as an individual language such as Mandarin or Yue (Cantonese) Chinese, and were instead coded as the Chinese macrolanguage in 2016.

The finding that the names of these macrolanguages often correspond well to the general terms often seen as actual ACS responses may help demonstrate the utility of the macrolanguage-level unit in language access planning, especially for written materials. For example, although individual spoken Chinese languages may be very different, they usually use a common writing system. Written materials, including ACS materials, are often translated into “Simplified Chinese,” which corresponds with the macrolanguage Chinese.

Whenever possible, language responses are coded as an ISO-639-3 language. Other codes have been added for common responses that can only be classified within a language family (i.e. Berber languages, Karen languages), or within a geographical region (i.e. Europe N.E.C., Nigeria N.E.C.). Similarly, languages within a macrolanguage are coded at the individual language level whenever possible, and at the macrolanguage level (i.e. Chinese, Arabic, Persian) when it is not possible to determine the specific individual language.

In real linguistic hierarchies, macrolanguages are inclusive of all the individual languages that make up the macrolanguage. However, in ACS coding, we needed our codes to be mutually exclusive in order to stay consistent with previous years’ data and to ensure that no person is double-counted. Cases that are coded to an individual language are not also coded to the macrolanguage. Only cases that cannot be coded down to an individual language are given the macrolanguage code. Therefore, researchers need to add together the macrolanguage and related individual languages to get the total estimate for the macrolanguage.

Updating the master file

The master file is a comprehensive list of responses with associated language codes. It has been assembled over the years from the basic coding process. Starting in 2013, analysts went through the existing master file, which contained every language that was reported on the 1980, 1990, and 2000 Censuses, and the American Community Surveys from 2001 forward. There were 187,000 unique responses in this file.

Analyzing the master file of responses required hundreds of hours of work divided among four analysts from the Social, Economic and Housing Statistics Division (SEHSD). After all 187,000 responses were coded as an ISO-standardized language, we had both an old language code as well as a new language

code for each individual response. For example, for a response of “AF SOMALI,” the old language would be Cushite, and the new language would be Somali (som).

A major result of this effort was identifying all the ISO-639-3 standardized languages that would go into the 2016 code list. There are over 7000 ISO-639-3 standardized languages active in the world, and part of this process was to decide which would be given a code in the American Community Survey. Any language that was written in at least once between 1980 and 2015 was given a code. This resulted in 1,334 language codes. While 1,334 is a large number of languages, there are thousands of active languages that are not on the ACS language list. It can be inferred that any language not on the 2016 code list had not been written in between 1980 and 2015, or could not be identified by coders. New languages that show up in the survey in 2016 and beyond will be flagged and analyzed to see if they should be added to the list later.

Alternate names for languages were also catalogued during this effort of combing through the 187,000 responses. Ethnologue (Simons and Fennig, 2017) has been our primary resource for finding alternate names and matching them with the appropriate ISO-standardized languages. As coders encounter alternate names for languages, they are entered into an internal database. Respondents typically do not know they are speaking “Eastern Oromo (hae),” and give language names like, “HARAR”/“HARER,” “ITTU,” or “QUOTTU”/“QWOTTU.” Starting in 2016, the coding system also shows language coders the respondent’s country of birth, in order to help confirm this code choice.

Creating the 2016 Language List

A main goal of the American Community Survey language program is to enumerate the languages spoken at home in the United States. For this reason, data collection and processing decisions reflect the distribution of languages spoken in the U.S., as opposed to the entire world.

In 2015, the list of languages recorded in the ACS was already the most detailed language list ever used by the Census Bureau, and with 384 languages, was one of the most detailed language lists used in any Census or survey in the world. However, the level of detail for language families was greater for languages that had been more prevalent in the United States in previous decades, such as Western Indo-European and Native North American languages, and lacked as great of detail for languages from Africa and Asia.

The language questions that now appear on the ACS were originally devised for the 1980 census and have remained unchanged ever since. Data from the 1980 census were used to create a detailed language code list that was later adopted by the ACS. Some codes have been added to the ACS language list over the years, but in 2015, the structure of the ACS language list largely reflected characteristics of the United States population in 1980.

Between 2013 and 2015, Census researchers analyzed every language that was reported on the 1980, 1990, and 2000 Censuses, and the 2001 to 2015 American Community Surveys, with the goal of expanding the original language code list to encompass the languages seen in the U.S. by 2016. This process resulted in the addition of over 900 new language codes available in ACS data. In data year 2016, the ACS language list had increased to 1,334 language codes. Table 1 shows the number of language codes added in different linguistic categories, along with the number of those who spoke

Spanish, Other Indo-European languages, Asian and Pacific Islander languages, and Other languages in 2016.

In updating the existing code list, many languages that were already on the 2015 code list needed to receive updated definitions or labels that were more transparent to users. For example, all speakers of French-based creole languages (863 thousand people in 2015) were coded as “French Creole.” There was a lot of interest and confusion from data users about what languages made up this group. The majority of French Creole speakers in the United States speak Haitian¹, and in 2016, Census products instead give estimates for number of Haitian speakers (856 thousand people in 2016). Even though they had small populations, most Native North American languages were already on the code list, as were many Creole languages inherent to the U.S. (e.g. Gullah, Cajun, and Hawaiian Pidgin). Increased immigration from Asia, the Pacific Islands, and Africa has led to growth in languages from those areas. Most of the languages added to the list were from those world regions, and Africa received the most new codes.

Prior to 2016, most languages of Africa were aggregated into codes such as “Cushite” which described a language family rather than an individual language such as Somali (som). Other African languages were grouped into codes like “Kru, Ibo, Yoruba,” which reflected a geographical area stretching from Liberia to Nigeria. This lack of specificity in the previous list is attributable to the fact that in 1980, very few people in the U.S. spoke a language from Africa other than Arabic – only about 50 thousand foreign-born speakers and an unknown number of native-born speakers (Gibson and Jung, 2006). In 2016 there were 1.2 million speakers of a language of Africa (excluding Arabic).²

Table 1. Number of Language Codes and Speakers by Language Category in the 2015 and 2016 American Community Survey Data

Language category	Number of language codes in 2015 ACS	Number of language codes in 2016 ACS	Number of speakers in 2016	Margin of Error (+/-)
Spanish	3	4	40,490,000	106,800
Other Indo-European languages	79	173	11,090,000	86,400
Primarily spoken in Europe	62	88		
Primarily spoken in Southern and Western Asia	17	85		
Asian and Pacific Island languages	91	280	10,600,000	63,500
Asian continent	44	168		

¹ Haitian and Haitian Creole are two terms for the same individual language. Haitian (hat) is the ISO-639-3 standardized label for the Haitian Creole language. Other Census Bureau public-facing documents use the more common label “Haitian Creole,” while American Community Survey tends to use “Haitian (hat)” to conform with the ISO standard.

² This estimate includes the categories “Amharic, Somali, or other Afro-Asiatic languages,” “Swahili or other languages of Central, Eastern, and Southern Africa,” and “Yoruba, Twi, Igbo, or other languages of Western Africa” listed in our [published table B16001](#), but subtracts estimates for two Afro-Asiatic languages from Asia, Chaldean neo-Aramaic and Assyrian neo-Aramaic, which are available from [public use microdata files](#).

Pacific Islands (including Philippine languages)	47	112		
Other languages	211	877	3,330,000	50,000
Afro-Asiatic (including Semitic) languages	7	67		
Languages of Western Africa	6	189		
Languages of Central, Eastern and Southern Africa	10	204		
Native North American languages	161	217		
Central and South American languages	12	87		
Creole languages	6	41		
Other (including Uralic and Caucasian languages and language isolates)	9	72		
Total	384	1334	65,520,000	154,000

Source: U.S. Census Bureau, American Community Survey, 2016, 1-year estimates. For more information, see [census.gov/acs](https://www.census.gov/acs).

The full list of 1,334 languages is not available in data products or public use files, due to confidentiality restrictions that apply to all data released by the Census Bureau. The most detailed table released for 2016 contains 42 categories, rather than the 39 categories in 2015. The 2016 public-use microdata file includes 132 languages or language groupings, up from 110 in 2015. For more information, see the section on “Data products,” below.

The number of languages available to users under the new coding scheme did not expand as dramatically as the number of languages given codes, as shown in the table above. The main goals of the coding redesign was not to increase detail, but to improve coding. The new coding system:

- increases the accuracy of coding by reducing guesswork in the assignment of responses to codes,
- provides flexibility to combine individual languages into useful groupings now and in the future,
- aligns with language classification standards, and
- provides a means to capture new and emerging languages.

Future data products may include an updated detailed language list similar to that released in 2015. While data releases for languages with a small number of speakers are limited by disclosure rules, there are some languages where release of specific information was limited by classification rules instead. As a result, future detailed language releases may contain more than the 350 languages included in the 2015 table release³. In addition, researchers from the Census Bureau and others with authorized access to the internal Census Bureau files may be able to study levels and trends in specific languages with a number of speakers too small to allow public release.

Finalizing the code list

³ Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 Years and Over: 2009-2013 is available at <<https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>>.

Table 2 shows a segment of the code list from codes 4790 to 4880, the Cushitic and Omotic languages, which are in the Afro-Asiatic language family. Previously, only one code existed (783) for all these languages. In 2016, these sixteen codes were added, which include three Omotic languages, 12 Cushitic languages and one macrolanguage (4830 - Oromo). These are not the only Cushitic and Omotic languages in the world, but they are the only ones ever reported on the Census or American Community Survey. Similarity between languages was determined primarily by closeness in the linguistic hierarchies in Ethnologue and Glottolog. When there was a disagreement between Ethnologue and Glottolog, we used the Ethnologue hierarchy. The similarity in language hierarchy is shown in the rightmost columns headed “Level 1” through “Level 7.”

Table 2.
Excerpt from the 2016 internal language code list showing sorting and classification of languages previously coded as Cushite

2016 code	2016 Label	Top 42 language category	PUMS category	2015 code	2015 label	Language hierarchy level 1	Language hierarchy level 2	Language hierarchy level 3	Language hierarchy level 4	Language hierarchy level 5	Language hierarchy level 6	Language hierarchy level 7
4790	Bilin (byn)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	Central	Northern			
4800	Afar (aar)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Saho-Afar			
4805	Saho (ssy)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Saho-Afar			
4815	Hadiyya (hdy)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Highland			
4825	Sidamo (sid)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Highland			
4830	Oromo (orm)	Other Afro-Asiatic languages	Oromo	783	Cushite	Afro-Asiatic	Cushitic	East	Oromo			
4832	Borana-Arsi-Guji Oromo (gax)	Other Afro-Asiatic languages	Oromo	783	Cushite	Afro-Asiatic	Cushitic	East	Oromo	Oromo (macrolanguage)		
4834	Eastern Oromo (hae)	Other Afro-Asiatic languages	Oromo	783	Cushite	Afro-Asiatic	Cushitic	East	Oromo	Oromo (macrolanguage)		
4836	West Central Oromo (gaz)	Other Afro-Asiatic languages	Oromo	783	Cushite	Afro-Asiatic	Cushitic	East	Oromo	Oromo (macrolanguage)	Nuclear Oromo	
4840	Somali (som)	Other Afro-Asiatic languages	Somali	783	Cushite	Afro-Asiatic	Cushitic	East	Somali			
4845	Garre (gex)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Somali			
4850	Maay (ymm)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	East	Somali			
4855	Beja (bej)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Cushitic	North				
4860	Aari (aiw)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Omotic	South				
4870	Kafa (kbr)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Omotic	North	Gonga-Gimojan	Gonga	South	
4880	Wolaytta (wal)	Other Afro-Asiatic languages	Other Afro-Asiatic languages	783	Cushite	Afro-Asiatic	Omotic	North	Gonga-Gimojan	Gimojan	Ometo-Gimira	Ometo

Note: Language hierarchy shown is from Ethnologue, 15th edition. Seven levels of linguistic hierarchy are shown in this table because that was as many as were represented in the languages used for this example. In the full code list, up to 15 levels of hierarchy (“L1” through “L15”) were used to classify all the languages.
Source: Education and Social Stratification Branch internal documentation.

In the list, the hierarchy lists the classification level that contains the listed language, with the language in the “2016 label” column being the specific one that was included in the final list.

Once the languages were determined, code numbers were assigned. To assign the 2016 code numbers, the languages first had to be sorted by similarity, and then space was added between codes to allow for the addition of future codes. We needed to create a code list that would not have to be changed much

in the future, in order to minimize the changing of technical specifications for tables and applications that use LAN and HHLAN code ranges in their sourcing.

Another requirement of the code numbering was to leave sufficient room in between numerical codes for adding new languages to the list without rearranging these codes in the future. Enough space had to be added between code numbers so that in the future, extant codes could be added in to the code list where they should fit in the linguistic hierarchy. The possible number of languages in each family, as well as the number of speakers in the world within each language family, were taken into consideration.⁴

When assigning a code number, the goal was to keep languages close to other linguistically similar languages. To sort the languages, they were grouped within the Level 1 hierarchy as shown in Table 2 (i.e. Afro-Asiatic languages grouped together), then in sub-hierarchies in level 2 and beyond. For example, Omotic languages were grouped together, then North Omotic, and so on.

Keeping similar languages together facilitated the creation of numerical ranges to capture a group of languages, which are used in computer programs and data products. For example, Oromo languages are in codes 4830 to 4839. Please see the headers in Appendix A for other suggested ranges that capture groups of languages and language families.

How do coders handle so many codes? Although there are many more codes to choose from, the new coding system is easier for coders overall, because they now rarely have to figure out which language family or category that a response falls in. For example, for a response of "SOMALI," coders used to have to look up the language family that Somali falls in and choose the code "Cushite." Now, coders would just choose the code for "Somali," which can take less time.

American Community Survey Production Language Coding Process

ACS questionnaires are administered in paper forms, Internet forms, and computer-assisted interview forms. Paper and Internet forms are self-response modes, while the computer-assisted forms are used by Census Bureau interviewers by telephone or in person. During data processing, punctuation, and extraneous characters are removed from the write-ins, and the first 20 characters are provided for coding. The responses are then compared with the computer list of write-ins that have been coded in the past (the master file), and if there is an exact match, the code is copied to the file for that entry. All other write-ins are sent to clerical coding.

Analysts take the responses, examine them against reference sources, and assign a code. Rules of thumb have been developed for coding macrolanguages and individual languages. However, many responses still require the coder to use their best judgment.

The coding efforts of the human coders and autocoder create the coded data in the unedited ACS data file. After coding, the editing and imputation process takes place, using logical assignment rules and hot-deck imputation for missing data.⁵ Any language coded as Uncodable, Unspecified, or missing in the

⁴ These numbers of possible living languages were taken from Ethnologue (<<https://www.ethnologue.com/statistics/family>>) in December 2015.

⁵ Table B99162 gives the number of specific languages allocated through the edits and imputation process each year, and is available at <<https://factfinder.census.gov>> and <<https://data.census.gov>>

unedited file is likely to be replaced with a real language in the editing and imputation process. For example, for the response of “SAME AS BEFORE,” a coder would code it as “Unspecified,” but logic rules within the editing process may assign a language if others in the household have a valid language code from which to draw.

Once a code is assigned, that particular response and code are stored in the master file. Then, when new responses are received in subsequent months, they are run through a computer program that assigns codes to any response that exactly matches one in the master file that has previously been coded. For example, if someone wrote in “SSPANESH,” and it had already been coded by a human coder as “1200 (Spanish),” the autocoder assigns it code 1200. The master file, when it is applied to coding operations in this way, is referred to as the “autocoder.” In 2017, approximately 99 percent of all responses in the unedited ACS were coded by using the autocoder, with the rest handled by subject matter experts through a system known as the Residual Coding System (RCS).

Coding challenges

ACS language codes are mutually exclusive, meaning each response of “language spoken at home” can only receive one numerical code. In actual linguistic hierarchies, macrolanguages are inclusive of all the languages that make up the macrolanguage. However, in ACS coding, they are not. For example, the code 1970 (Chinese) does not include all the individual Chinese languages in codes 1975 through 2050, and the same applies to 4500 (Arabic), 1290 (Persian), 2710 (Malay) and other macrolanguages (see Appendix A). The most common macrolanguages are Chinese and Arabic, which share the same coding challenges. If a coder has the information to code a response as an individual language (1975 through 2050 for Chinese; 4502 through 4540 for Arabic), they will. If not (for example if the write in was just “CHINESE” or just “ARABIC,” it is coded as the macrolanguage 1970 (Chinese) or 4500 (Arabic).

Effect of the autocoder

As discussed earlier, the ACS coding process uses an autocoder for text responses that checks to see if that exact response has been coded in the past, and if so, it assigns the same code that was given in the past. The use of the autocoder factors greatly into some coding rules. There is a continuing effort to monitor whether the coding and autocoding of languages is being done correctly. The autocoder is necessary to create efficiencies, but in some cases it makes coding decisions tricky. Sometimes, languages are coded at the macrolanguage level even when it is possible for the coder to guess the individual language, because of the autocoder use. For example, if a person wrote in “ARABIC,” and their place of birth was Lebanon, it would be coded as the macrolanguage “Arabic” as opposed to an individual language spoken in Lebanon in order to ensure all future exact-match responses of “ARABIC” were coded as Arabic. However, if someone wrote in “LEBANESE ARABIC,” that could be coded as “North Levantine Spoken Arabic,” as could all future exact-match responses.

To address the above challenges, more uniform coding rules were created to ensure less variation in coding. These included rules for incorporating other information available from the ACS such as Place of Birth, keeping in mind the use of an autocoder.

Formal coding rules had to be devised for responses that were just the name of a country. Devising these rules was simple for some country names and more difficult for others. In countries where a single

language predominates, a language can be assigned directly from a country name. For example, a response of "SOMALIA" is coded as Somali. For countries where a multitude of languages were spoken, deciding on which individual language to choose required the development of some decision rules:

- The language that had the largest number of speakers in the country in the most recent available data was usually chosen.
- However, if English was the most common language for a country, it was not chosen, predicated on the assumption that the person had correctly indicated that he or she spoke a language other than English.
- If there was a de facto language of national identity, it was sometimes chosen over a language that was more populous, when the more populous language was associated with a European country instead of a language native to the country in question.
- If more than one language would make an appropriate choice, a code for "<Country Name> N.E.C." was created, along with a coding rule. For example, a response of "NIGERIA" is coded as "Nigeria N.E.C." because there are many widely-spoken languages associated with Nigerian identity.

Machine learning feasibility project

A feasibility project was undertaken to see if the coding process could be augmented with text mining and machine learning.⁶ This project showed that responses of language names are less amenable to machine classification than other responses (such as addresses, occupations, or fields of degree) might be. This is because of the nature of the language responses, which come in many languages and alphabets.

This exploratory project has helped inform many current efforts at the Census Bureau to use machine learning and text mining to process responses. In the future, further developments may allow the Census Bureau to make use of these techniques.

Multiyear microdata crosswalk for the 2012-2016 5-year microdata

The above discussion applies to the 2016 one-year data. Previous years' 1-year data are unchanged. However, the five-year data file released by the Census Bureau contained data from both the old and new coding systems. In order to create a five year file using the new system, many LAN codes received a probabilistic recode for the 2012 through 2015 portion of the 2012-2016 5-year microdata. We have added over 900 individual codes, so the microdata crosswalk was mostly one-to-many matches (i.e. Arabic split into a dozen types of Arabic). First, cases from 2012 to 2015 were coded with the old and new LAN code, and then the percents of each old LAN code that went into each new LAN code were used to make "conversion rates" by nation and state that were used to recode LAN for the 2012-2016 5-year file. This process is similar to what is typically done for the ACS Industry variable when there is a shift in industry classifications.

⁶ "Text mining" identifies patterns and trends through means such as statistical pattern learning, usually by structuring the input text and deriving patterns within the structured data. "Machine learning" uses statistical techniques to give computers the ability to progressively improve performance on a specific data task, without being explicitly programmed each step of the way.

The specification provided by subject matter experts in ESSB said how to allocate each code - 100 percent to a new code, probabilistic using national-level conversion rates, or probabilistic using state conversion rates. These decisions were made through categorical data analysis of distributions of languages at the national and state levels.

Expanded household-level data and added variable - HHLAN

In 2015, person-level data in the microdata file were assigned individual languages (the variable LAN). However, the specific household-language variable, HHL, presented household languages in five categories: English only, Spanish, Other Indo-European languages, Asian and Pacific Islander languages, and other languages. In the ACS 1-year data beginning in 2016, the new variable HHLAN specifies the language spoken at home as a numeric code (1000-9999) by selecting one of 1,333 language codes on the edited LAN variable, based on responses from members of the household (see the “Household Language Algorithm” box below). The detailed codes are then collapsed into just over 130 categories that are available to the public in the PUMS datasets. The household language variables HHLAN and HHL are available in the PUMS for housing units only, as group quarters and vacant units are not included.

Each household language is defined by a discrete language code. HHLAN has the same codes as LAN (1000 through 9999), with one extra – 9500-English-only households. This 9500 value means that there is no one in the household who reported speaking a language other than English at home. There is no equivalent English-only code for the LAN variable, as English-only speakers are out of universe. The HHLAN variable will not be available in the 5-year ACS data until 2016-2020.

Household Language Algorithm

In households where one or more persons (age 5 years old and over) speak a language other than English, the household language assigned to all household members is the non-English language spoken by the first person with a non-English language in the following order: reference person, spouse, parent, sibling, child, grandchild, in-law, other relative, unmarried partner, housemate/roommate, roomer/boarder, foster child, or other nonrelative. Thus, persons who speak only English may have a non-English household language assigned to them in tabulations of persons by household language.

If the answer for all members of the household (age 5 years old and over) to the question “Do you speak a language other than English at home?” is “No,” then household language is assigned as 9500 - English only. Household language is missing for Group Quarters and vacant units.

DATA PRODUCTS

Language categories

Due to small sample counts, data tabulations are not available for all 1300+ languages. Instead, the Census Bureau collapses languages into smaller sets. In tabulations, individual languages are usually combined to create a category that avoids presenting data for small groups.

In general, data users would like to receive information on as many languages as possible, for the smallest geographies possible. However, the concentrations of language-other-than-English speakers vary widely across the country and between urban, suburban and rural areas. For this reason, ACS language estimates have always been presented within a range of nested language categorizations. More detailed breakdowns are generally needed in cities and in states with high proportions of language-other-than-English speakers, while more rural areas often just need to know estimates of English only, Spanish, and other languages.

High-level breakdowns are used in the majority of language data products because, in most geographies for which we publish standardized tables, the number of those speaking a language other than English or Spanish at home is small. We re-examined two of our commonly used high-level groupings: English Only/Spanish/Other languages, and English Only/Spanish/Other Indo-European languages/Asian and Pacific Island languages/Other languages. These breakdowns still seem to have the most utility for users in state and local agencies who are seeking to comply with civil rights laws protecting linguistic minorities. We specifically looked at whether to break apart the categories of “Other Indo-European,” “Asian and Pacific Islander,” and “Other languages” (which includes a mix of unrelated languages, i.e., African and Afro-Asiatic languages, Uralic languages, and Creole languages). The frequencies did not support doing this, so we kept these high-level breakdowns the same.

These collapsing standards are useful because they can be readily used in crosstabulations in other Census products or applications. The simplest collapse recodes languages other than English into four major language groups: Spanish, Other Indo-European languages, Asian and Pacific Island languages, and Other languages. More detailed collapsing uses 42 non-English languages and language groups (the comparable table for 2015 had 39 languages or language groups). An even more detailed collapsing scheme put the languages into 133 codes available in the Public Use Microdata Sample (PUMS) compared with 110 in 2015. Please see Appendix B for a list of the final PUMS codes for the individual (LANP) and household (HHLANP) variables, along with comparability to previous years, and the individual language codes (from variables LAN and HHLAN, respectively) that are collapsed into each code.

In American Factfinder and data.census.gov, Table B16001 provides the most detail for individual languages, tabulated by English-speaking ability. A top-42 language category breakdown is available in table B16001 for individuals. This breakdown is not available for households.

Table C16001 is a collapsed specific-language table with fewer languages. The categories in B16001 and C16001 were revised beginning with 2016 1-year and 2012-2016 5-year data, to better reflect the most commonly spoken languages in the United States. A 12 language category breakdown is available in table C16001 for individuals, and table B16002 for households (not yet available in 5-year datasets).

The finest level of detail available to the public for 2016 is in the Public Use Microdata Sample (PUMS). Data for smaller languages with estimated populations of 10,000 speakers or more nationwide are available in the 2016 1-year and 2012-2016 5-year Public Use Microdata Sample datasets.⁷

In 2015, the Census Bureau released a detailed table package listing 350 languages, using the 2009-2013 five-year data file.⁹ Future releases of a similar table using the updated coding system are likely in coming years. Finally, for internal Census and Research Data Center (RDC) users, the ACS microdata contains all 1334 codes for LAN in the 1-year and 5-year datasets and HHLAN in the 1-year dataset.

Updating the 12-language and 42-language category breakdowns

The determination of whether to show an individual language or collapse it into an aggregated category depends chiefly on the size of the population in the United States speaking that language at home. In tabulations, smaller languages are aggregated with other languages in a way that meets a certain population threshold, but has some utility for translators or researchers. Frequencies to inform table decisions were examined, including trends over the past 10 years. For table B16001, all categories except for Navajo and Other Native North American met a 200,000 speaker nationwide threshold.

Some languages were added, and some were deleted/combined into existing “other” categories. Rows for Scandinavian languages, Yiddish, and Hungarian were deleted from the table, but are still available in the PUMS and microdata. Thai and Lao were collapsed into “Thai, Lao, or other Tai-Kadai languages,” but are still available individually in the PUMS. “African languages” was split into new categories (see below).

Additions to table B16001 included Haitian, Bengali, Punjabi, Telugu, and Tamil added as individual languages. New categories were also added: “Malayalam, Kannada, or other Dravidian languages,” “Amharic, Somali, or other Afro-Asiatic languages,” “Yoruba, Twi, Igbo, or other languages of Western Africa,” and “Swahili or other languages of Central, Eastern, or Southern Africa.” Table 3 shows more detail on the 42 language groups, while Table 4 lists languages that can be compared between 2016 and previous years either directly by use of tables or by combining categories in the PUMS files. Following that is a description of languages where definitional changes lead us to recommend making comparisons with caution.

⁷ For information on the PUMS and other ways of obtaining data for individual languages such as those in the “Other” categories in American Factfinder tables, please see <http://www.census.gov/topics/population/language-use/about/faqs.html>.

⁹ <https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>.

Table 3. Four and Forty-Two Group Classifications of Languages Spoken at Home with Examples

<i>Four group classification</i>	<i>Forty-two group classification</i>	<i>Examples</i>
Spanish	Spanish	Spanish, Ladino
Other Indo-European languages	French (incl. Cajun)	French, Cajun
	Haitian	Haitian
	Italian	Italian, Sicilian
	Portuguese	Portuguese, Kabuverdianu
	German	German, Luxembourgish
	Yiddish, Pennsylvania Dutch or other West Germanic languages	Dutch, Yiddish
	Greek	Greek
	Russian	Russian
	Polish	Polish
	Serbo-Croatian	Bosnian, Croatian, Serbian
	Ukrainian or other Slavic languages	Bulgarian, Czech, Ukrainian
	Armenian	Armenian
	Persian (incl. Farsi, Dari)	Iranian Persian (Farsi), Dari
	Gujarati	Gujarati
	Hindi	Hindi
	Urdu	Urdu
	Punjabi	Punjabi (Panjabi)
	Bengali	Bengali
	Nepali, Marathi, or other Indic languages	Nepali, Marathi, Konkani
	Other Indo-European languages	Albanian, Lithuanian, Pashto (Pushto), Romanian, Swedish
	Telugu	Telugu
	Tamil	Tamil
	Malayalam, Kannada, or other Dravidian languages	Malayalam, Kannada

Asian and Pacific Island languages	Chinese (incl. Mandarin, Cantonese)	Mandarin Chinese, Min Nan Chinese (incl. Taiwanese), Yue Chinese (Cantonese)
	Japanese	Japanese
	Korean	Korean
	Hmong	Hmong
	Vietnamese	Vietnamese
	Khmer	Central Khmer (Cambodian)
	Thai, Lao, or other Tai-Kadai languages	Thai, Lao
	Other languages of Asia	Burmese, Karen, Turkish, Uzbek
	Tagalog (incl. Filipino)	Tagalog, Filipino
	Ilocano, Samoan, Hawaiian, or other Austronesian languages	Cebuano (Bisayan), Hawaiian, Iloko (Ilocano), Indonesian, Samoan
All other languages	Arabic	Arabic languages
	Hebrew	Hebrew
	Amharic, Somali, or other Afro-Asiatic languages	Amharic, Chaldean Neo-Aramaic, Somali, Tigrinya
	Yoruba, Twi, Igbo, or other languages of Western Africa	Akan (incl. Twi), Igbo (Ibo), Wolof, Yoruba
	Swahili or other languages of Central, Eastern, and Southern Africa	Ganda, Kinyarwanda, Lingala, Swahili
	Navajo	Navajo
	Other Native languages of North America	Apache languages, Cherokee, Lakota, Tohono O'odham, Yupik languages
	Other and unspecified languages	Hungarian, Jamaican Creole English, Unspecified

Table 4. Detailed Table B16001 row comparability between 2015 and 2016

Language categories in 2005 to 2015	Language categories in 2016 B16001
Speak only English	Speak only English
Spanish or Spanish Creole	Spanish
French (incl. Patois, Cajun)	Add B16001 French (incl. Cajun) estimate to PUMS estimate of Jamaican Creole English See "Compare with caution" note below.
May be comparable to French (incl. Patois, Cajun) in some geographies. See "Compare with caution" note below.	French (incl. Cajun)
French Creole See "Compare with caution" note below.	Haitian See "Compare with caution" note below.
Italian	Italian
Portuguese or Portuguese Creole	Portuguese
German	German
Yiddish	Use PUMS estimates
Other West Germanic languages	Subtract PUMS estimate for Yiddish from B16001 estimate for Yiddish, Pennsylvania Dutch or other West Germanic languages
Add B16001 estimates for Yiddish and Other West Germanic languages	Yiddish, Pennsylvania Dutch or other West Germanic languages
Scandinavian languages	Use PUMS estimates
Greek	Greek
Russian	Russian
Polish	Polish
Serbo-Croatian	Serbo-Croatian
Other Slavic Languages	Ukrainian or Other Slavic Languages
Armenian	Armenian
Persian	Persian (incl. Farsi, Dari)
Gujarati	Gujarati
Hindi	Hindi
Urdu	Urdu
Other Indic languages	Add B16001 rows for Punjabi, Bengali, and Nepali, Marathi, or other Indic languages
Use PUMS estimates for Panjabi	Punjabi
Use PUMS estimates	Bengali
B16001 estimate for Other Indic languages minus PUMS estimates for Bengali and Panjabi	Nepali, Marathi, or other Indic languages
Other Indo-European languages - Not comparable to 2016	Other Indo-European languages - Not comparable to 2015
Use PUMS estimate	Telugu
Use PUMS estimate	Tamil
Not comparable to 2016, though most speakers would be included within PUMS estimates for Malayalam and Kannada	Malayalam, Kannada, or other Dravidian languages
Chinese	Chinese (incl. Mandarin, Cantonese)
Japanese	Japanese
Korean	Korean
Hmong	Hmong
Vietnamese	Vietnamese

Mon-Khmer, Cambodian	Khmer
Thai	Use PUMS estimates
Laotian	Use PUMS estimates
Add B16001 estimates for Thai and Laotian	Thai, Lao, and other Tai-Kadai languages
Other languages of Asia - Not comparable to 2016	Other languages of Asia - Not comparable to 2015
Tagalog	Tagalog (incl. Filipino)
Other Pacific Island Languages	Ilocano, Samoan, Hawaiian, or other Austronesian languages
Arabic	Arabic
Hebrew	Hebrew
African languages	<ol style="list-style-type: none"> 1. Add B16001 estimates for Amharic, Somali, or other Afro-Asiatic languages, Swahili or other languages of Central, Eastern, and Southern Africa, and Yoruba, Twi, Igbo, or other languages of Western Africa. 2. Subtract PUMS estimates for Chaldean Neo-Aramaic and Assyrian Neo-Aramaic (Afro-Asiatic languages from Asia)
Not comparable to 2016	Amharic, Somali, or other Afro-Asiatic languages
Not comparable to 2016	Yoruba, Twi, Igbo, or other languages of Western Africa
Not comparable to 2016	Swahili or other languages of Central, Eastern, and Southern Africa
Navajo	Navajo
Other Native North American languages	Other Native languages of North America
Hungarian	Use PUMS estimates
Other and unspecified languages - Not comparable to 2016	Other and unspecified languages - Not comparable to 2015

Compare with caution between 2015 and 2016:

- French (incl. Cajun): Beginning in 2016, responses of “PATOIS” are coded as Jamaican Creole English. Prior to this, they were included in “French.” Detailed estimates for Jamaican Creole English are available in the PUMS files. For most geographies, this category should be comparable to previous estimates of French (incl. Patois, Cajun). A decline between 2015 and 2016 may be seen in geographies where there are many speakers of Patois.
- Creoles other than Haitian and Kabuverdianu have been moved to “Other and Unspecified languages.”
- French Creole: Although other ISO-standardized individual French Creole language codes were added in 2016, most responses that were previously coded as 623-French Creole were coded as Haitian beginning in 2016. This includes many responses of just “KREYOL” or “CREOLE,” if there was no evidence (i.e. other languages spoken in the household, place of birth, or ancestry) to show that what exactly the creole language referred to is. Much thought and analysis went into this decision, as a write-in of “CREOLE” could mean many things. However, all of our analyses over the years have shown that over 90% of people who wrote in “KREYOL” or “CREOLE” in the ACS were born in Haiti or had Haitian ancestry.

Comparable between 2015 and 2016, but requires the use of PUMS data (see Appendix B for PUMS categories):

- Yiddish was previously on its own row but has now been combined with Other West Germanic languages in B16001. Yiddish is still available in the PUMS.
- Bengali, Punjabi, Telugu, and Tamil are all new to the B16001 table. Previous years' data for these languages are available in the PUMS. The 2015 PUMS category "Panjabi" is comparable to the 2016 B16001 row and PUMS category "Punjabi." In the 2005-2015 B16001 tables, Bengali and Punjabi were included under "Other Indic languages," while Telugu and Tamil were included under "Other Asian languages."
- Scandinavian languages were previously on their own row but have now been combined with Other Indo-European languages. Individual Scandinavian languages are still available in the PUMS.
- Thai and Laotian were previously on their own rows but are now combined within Thai, Lao, or other Tai-Kadai languages. Thai and Lao estimates are individually available in the PUMS.

C16001 comparability between 2015 and 2016

Many rows in C16001 have received label changes, to clarify what some of the major languages in each category are. While the label may have changed, all of these rows are comparable with the previous years C16001 row, except:

- Arabic is new to C16001, and was previously only available in B16001 and PUMS.
- The "Other and Unspecified" row no longer contains Arabic.
- C16001 has added English proficiency categories for languages other than English (speak English very well or less than very well).

New table in 2016

A new household-level table, B16002, is available using 2016 1-year ACS data. In this table, more languages are available for households defined as limited English-speaking. This new B16002 is available using 1-year data until the 2016-2020 5-year data is released, because it sources HHLAN. A collapsed version of this table, C16002, is available in 2016 1-year and 2012-2016 5-year data. The format of C16002 is the same as B16002 from 2015 and before.

Updating the PUMS categories for LANP and HHLANP

A language code list is provided with the release of the 2016 PUMS.¹⁰ The collapsing scheme that creates PUMS categories out of the 1,334 microdata language codes for individual spoken language and household language was based on several considerations. PUMS Categories were chosen based on sample size, mutual intelligibility of languages, and precision of coding. These categories had to be large enough to publish for all Public Use Microdata Areas (PUMAs) without compromising privacy protections for ACS respondents, yet small enough to be useful for a data user that needs to know exactly what language is involved. New categories were added for 36 languages and language groups

¹⁰ The collapsed code lists for these topics can be found in the "Code Lists" section at <https://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html>

that previously had not been available. For the first time, American Community Survey data users could find official estimates of Somali, Oromo, Tigrinya, Yoruba, Igbo and other African languages that were previously grouped together, as well as many other languages from around the world that have grown or were recently added to the code list.

PUMS categories had to meet a threshold of 10,000 speakers nationwide, in both the 1-year 2016 ACS microdata and the 5-year 2012-2016 ACS microdata, in order to be included in the dataset.¹¹ In this way, we could keep the language categories and labels in the 1-year PUMS dataset exactly the same as the 5-year PUMS dataset. Only about 100 individual languages met this threshold. All of the other language codes had to be combined into categories large enough to meet the threshold. These categories also needed to provide utility to translators, by keeping mutually intelligible languages together. Wherever possible, we tried to carve out categories that kept languages together within a small, closely related language family. However, a few broader “Other” categories still remain that combine distantly related or unrelated languages. The main resource used to determine mutual intelligibility between languages was Ethnologue. Glottolog and other academic and cultural resources were consulted.

The PUMS categories were also chosen in part to avoid giving a false impression of precision. In cases where there were barriers to precise coding into individual ISO-639-3 languages (please see above, under “Coding challenges” for more discussion of precision of coding), the preference was to roll up related languages, especially when they fell into a macrolanguage. For example, the majority of Arabic speakers only write in “ARABIC,” so publishing the estimate of “Arabic” speakers (including all individual languages) is preferable than publishing data for different individual Arabic languages.

Creating the multiyear Public Use Microdata Sample (PUMS) LANP data

The PUMS multiyear data for 2012-2016 were created using a similar multi-year crosswalking process as the multi-year microdata. This resulted in the probabilistic allocation of the 2012 through 2015 portion of the 2012-2016 5-year PUMS LANP data from the old PUMS variable LANP12. The values created by this recode will also carry over into the 2013-2017, 2014-2018, and 2015-2019 5-year PUMS datasets.

The PUMS multiyear crosswalk was created from the specifications for the crosswalking the microdata (see Multiyear Microdata Crosswalk section on page 9). The steps taken were:

1. Start with specifications used to create the 2012-2015 portion of the 2012-2016 multiyear microdata. These contained conversion rates for the LAN variable in the microdata for the nation and states.
2. Convert all microdata codes (LAN or HHLAN) into their corresponding PUMS code (LANP, LANP12 or HHLANP).
 - a. Values of LAN for the years 2012 to 2015 were converted to the corresponding PUMS LANP12 value (i.e. 613-Luxembourgian was converted to 607-German)

¹¹ The 10,000 speaker minimum was based on the internal ACS weighted, edited microdata, which includes the entire ACS sample. Estimates derived from the ACS PUMS often do not match estimates from the ACS microdata because the PUMS includes a subset of the full ACS sample. For this reason, LANP estimates obtained using the PUMS dataset may show fewer than 10,000 speakers in some categories.

- b. Values of LAN for 2016 were converted to the corresponding LANP value (i.e. 1123-Luxembourgish (Itz) was converted to 1110-German)
- 3. Conversion rates were generated for the PUMS recode, by nation and state. Because there were many fewer categories, this resulted in many simpler and more one-to-one conversions than in the multiyear microdata allocation.

Table 5. Public Use Microdata Sample (PUMS) variable changes

Variable	Label	Changed in 2016?
LANX	Language other than English spoken at home	NO
LANP05	Language spoken at home for data collected prior to 2012	NO – not in dataset
LANP12	Language spoken at home for data collected in 2012 or later	YES – 2012 through 2015 data recoded into new LANP codes for 5-year data
LANP	Language spoken at home for data collected in 2016 or later	YES – new codes added
HHL	Household language	NO
HHLANP	Household language – detailed	YES – new variable added for 1-year PUMS only until 2020 data release
FLANP	Language spoken at home allocation flag	NO
FLANXP	Language other than English allocation flag	NO

Acknowledgements

The language data are primarily the result of the efforts of millions of households who chose to respond to the Census and the American Community Survey.

The ACS language redesign was led by Kurt Bauman. This report was prepared under the direction of Stephanie Galvin and David Waddington. Erik Vickstrom made many of the conceptual and statistical decisions for the language redesign, including choosing the standards and languages to use. The four analysts who cleaned and coded all the responses in the master file were Savet Hong, Erik Vickstrom, Christine Gambino, and Maribel Cervantes-Ortega. The language changes to the 2016 1-year and 2016 5-year datasets and data products could never have happened without the expertise and support from many other staff involved in data processing, statistical review, SAS programming, data product development, and customer outreach in the American Community Survey Office (ACSO) and Decennial Information Technology Division (DITD).

References

Gibson, Campbell and Kay Jung, 2006. Population Division Working Paper No. 81, U. S. Census Bureau, Washington, DC [PDF Version](#) [1.17M]

Simons, Gary F. and Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the World, Twentieth edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>. From 11/13/17.

American Community Survey Redesign of
Language-Spoken-at-Home Data, 2016

Christine P. Gambino
Social, Economic and Housing Statistics Division
U.S. Census Bureau

SEHSD Working Paper Number 2018-31

Appendix A: American Community Survey Language Code List Page 1

Appendix B: Public Use Microdata Sample (PUMS) LANP-LANP12
Categories and Comparability, 2012-2016 Page 15

APPENDIX A: AMERICAN COMMUNITY SURVEY LANGUAGE CODE LIST

INDO-EUROPEAN LANGUAGES (1053-1056, 1069-1073, 1110-1564)

Haitian (1053-1056)¹

1053 Guadeloupean Creole French (gcf)
 1054 Saint Lucian Creole French (acf)
 1055 Haitian (hat)

West Germanic languages (1110-1139, 1234)

German (1110-1124)

1110 German (deu)
 1111 Bavarian (bar)
 1112 Hutterite German (geh)
 1113 Low German (nds)
 1114 Plautdietsch (ptd)
 1115 Swabian (swg)
 1120 Swiss German (gsw)
 1121 Upper Saxon (sxu)
 1122 Limburgish (lim)
 1123 Luxembourgish (ltz)

 1125 Pennsylvania German (pdc)
 1130 Yiddish (macrolanguage) (yid)
 1131 Eastern Yiddish (ydd)
 1132 Dutch (nld)
 1133 Vlaams (vls)
 1134 Afrikaans (afr)
 1135 Northern Frisian (frr)
 1136 Western Frisian (fry)
 1234 Scots (sco)

Scandinavian languages (1140-1146)

1140 Swedish (swe)
 1141 Danish (dan)
 1142 Norwegian (nor)
 1143 Icelandic (isl)
 1144 Faroese (fao)

1147 Belgium N.E.C.

Romance languages (1155-1227)

Italian (1155-1169)

1155 Italian (ita)
 1156 Corsican (cos)
 1157 Friulian (fur)

1158 Ligurian (lij)
 1159 Lombard (lmo)
 1160 Neapolitan (nap)
 1161 Piemontese (pms)
 1162 Romagnol (rgn)
 1163 Sardinian (macrolanguage) (srd)
 1164 Sicilian (scn)
 1165 Venetian (vec)

French (1170-1175)

1170 French (fra)
 1172 Jèrriais (nrf)
 1174 Walloon (wln)
 1175 Cajun French (frc)

 1176 Occitan (post 1500) (oci)

Spanish (1200-1205)

1200 Spanish (spa)
 1201 Asturian (ast)
 1202 Ladino (lad)
 1205 Caló (rmq)

1206 Catalan (cat)

Portuguese (1069-1073, 1210-1217)

1069 Kabuverdianu (kea)¹
 1072 Upper Guinea Crioulo (pov)¹
 1210 Portuguese (por)
 1211 Galician (glg)

1218 Aromanian (aen)

1220 Romanian (ron)
 1221 Istro Romanian (ruo)
 1223 Ladin (lld)
 1225 Romansh (roh)

Celtic languages (1228-1233)

1228 Welsh (cym)
 1229 Breton (bre)
 1230 Cornish (cor)
 1231 Irish (gle)
 1232 Manx (glv)
 1233 Scottish Gaelic (gla)

1235 Modern Greek (1453-) (ell)

¹ Although they are linguistically Creole languages, some French-based Creoles (Guadeloupean Creole French, Saint Lucian Creole French, and Haitian) and Portuguese-based creoles (Kabuverdianu and Upper Guinea Crioulo) are included under Indo-European languages in standardized tabulations, in order to maintain comparability with ACS estimates prior to 2016.

Note: N.E.C. stands for Not Elsewhere Classified. Whenever possible, language names and three-letter codes follow the ISO-639-3 standard. Other codes have been added for common write-ins that cannot be identified at the ISO-639-3 level, but can be classified within a language family (i.e. Berber languages, Karen languages), or within a geographical region (i.e. Europe N.E.C., Nigeria N.E.C.). Languages within a macrolanguage are coded at the more detailed individual language level whenever possible, and at the macrolanguage level (i.e. Chinese, Arabic, Persian) when it is not possible to determine the individual language from the write-in answer. ACS language questions are only designed to capture spoken languages and do not give accurate estimates of signed languages. Signed languages are recorded within the internal unedited data; however, they are combined with the closest spoken language in the edited data (i.e., American Sign Language is combined with English).

1238 Pontic (pnt)
 1242 Albanian (macrolanguage) (sqi)
 1243 Arbëreshë Albanian (aae)
 1244 Gheg Albanian (alh)
 1245 Tosk Albanian (als)

Slavic languages (1250-1280)

1250 Russian (rus)
 1251 Belarusian (bel)
 1252 Rusyn (rue)
 1260 Ukrainian (ukr)
 1262 Czech (ces)
 1263 Slovak (slk)
 1265 Lower Sorbian (dsb)
 1266 Upper Sorbian (hsb)
 1268 Kashubian (csb)
 1269 Silesian (szl)
 1270 Polish (pol)
 1273 Bulgarian (bul)
 1274 Macedonian (mkd)
 1275 Serbo-Croatian (macrolanguage) (hbs)
 1276 Bosnian (bos)
 1277 Croatian (hrv)
 1278 Serbian (srp)
 1280 Slovenian (slv)

Baltic languages (1281-1287)

1281 Lithuanian (lit)
 1283 Standard Latvian (lvs)
 1284 Latgalian (ltg)

1288 Armenian (hye)

Indo-Iranian languages (1289-1543)

Persian (1289-1307)

1290 Persian (macrolanguage) (fas)
 1292 Dari (prs)
 1295 Iranian Persian/Farsi (pes)
 1300 Tajik (tgk)
 1302 Bukharic (bhh)
 1303 Hazaragi (haz)
 1304 Judeo-Persian (jpr)

 1308 Judeo-Tat (jdt)
 1309 Northern Luri (lcr)
 1310 Parsi (prp)
 1311 Zoroastrian Dari (gbz)
 1312 Dezfuli (def)
 1315 Kurdish (macrolanguage) (kur)
 1316 Central Kurdish (ckb)
 1317 Northern Kurdish (kmr)
 1320 Zaza (macrolanguage) (zza)
 1321 Baluchi (macrolanguage) (bal)
 1322 Eastern Balochi (bgp)
 1323 Southern Balochi (bcc)
 1325 Ossetian (oss)
 1327 Pushto (macrolanguage) (pus)

1328 Northern Pashto (pbu)
 1329 Southern Pashto (pbt)
 1331 Shughni (sgh)
 1333 Wakhi (wbl)

Indo-Aryan (Indic) languages (1340-1543)

1340 India N.E.C.

Hindi (1341-1359)

1341 Chhattisgarhi (hne)
 1342 Garhwali (gbm)
 1343 Haryanvi (bcg)
 1349 Awadhi (awa)
 1350 Hindi (hin)
 1351 Fiji Hindi (hif)

1360 Urdu (urd)
 1365 Deccan (dcc)
 1370 Memoni (mby)
 1375 Assamese (asm)

Bengali (1380-1387)

1380 Bengali (ben)
 1382 Chittagonian (ctg)
 1384 Rangpuri (rkt)
 1386 Sylheti (syl)

1388 Halbi (hlb)
 1400 Rohingya (rhg)

Punjabi (1420-1429)

1420 Panjabi (pan)
 1421 Lahnda (macrolanguage) (lah)
 1422 Western Panjabi (pnb)
 1423 Pahari-Potwari (phr)
 1424 Northern Hindko (hno)
 1425 Southern Hindko (hnd)
 1426 Saraiki (skr)

 1430 Goan Konkani (gom)
 1435 Konkani (individual language) (knn)
 1440 Marathi (mar)
 1442 Od (odk)
 1450 Gujarati (guj)
 1455 Saurashtra (saz)
 1460 Bhojpuri (bho)
 1463 Magahi (mag)
 1467 Maithili (mai)
 1470 Rajasthani (raj)
 1472 Bagri (bgq)
 1474 Dhatki (mki)
 1478 Lambadi (lmn)
 1480 Marwari (mwr)
 1481 Marwari (India) (rwr)
 1482 Dhundari (dhd)
 1486 Bodo Parja (bdv)
 1488 Desiya (dso)
 1490 Oriya (individual language) (ory)

1491	Sambalpuri (spv)
1495	Kashmiri (kas)
1497	Khovar (khw)
1499	Dangaura Tharu (thl)
1500	Nepali (individual language) (npi)
1501	Dotyali (dty)
1510	Kachhi (kfr)
1515	Sindhi (snd)
1518	Dogri (dgo)
1521	Kumaoni (kfy)
1525	Pakistan N.E.C.
1526	Dhivehi (div)
1530	Sinhala (sin)
1540	Romany (rom)
1541	Balkan Romani (rmn)
1542	Carpathian Romani (rmc)
1543	Vlax Romani (rmy)

ASIAN AND PACIFIC ISLAND LANGUAGES (1643-3798)

Turkic languages (1643-1684)

1643	Chuvash (chv)
1648	Kara-Kalpak (kaa)
1649	Kazakh (kaz)
1650	Kirghiz (kir)
1652	Uzbek (uzb)
1653	Northern Uzbek (uzn)
1654	Southern Uzbek (uzs)
1657	Karachay-Balkar (krc)
1658	Bashkir (bak)
1659	Tatar (tat)
1662	Crimean Tatar (crh)
1666	Uighur (uig)
1668	Azerbaijani (aze)
1669	North Azerbaijani (azj)
1670	South Azerbaijani (azb)
1675	Turkish (tur)
1676	Balkan Gagauz Turkish (bgx)
1677	Gagauz (gag)
1680	Turkmen (tuk)
1681	Yakut (sah)

Mongolic languages (1685-1698)

1685	Daur (dta)
1687	Bonan (peh)
1688	Tu (mjg)
1690	Mongolian (macrolanguage) (mon)
1691	Halh Mongolian (khk)
1692	Peripheral Mongolian (mvf)
1695	Buriat (bua)
1696	Russia Buriat (bxr)
1698	Kalmyk (xal)

Tungusic languages (1699-1710)

1699	Even (eve)
1702	Evenki (evn)
1703	Oroqen (orh)
1705	Nanai (gld)
1708	Manchu (mnc)

Dravidian languages (1711-1798)

1711	Northwestern Kolami (kfb)
1716	Kui (India) (kxu)
1720	Northern Gondi (gno)
1730	Telugu (tel)
1735	Badaga (bfq)
1737	Kannada (kan)
1750	Malayalam (mal)
1755	Paniya (pcg)
1760	Kodava (kfa)
1765	Tamil (tam)
1770	Tulu (tey)
1775	Brahui (brh)
1780	Kurukh (kru)

1799	Burushaski (bsk)
------	------------------

Austro-Asiatic languages (1800-1969)

1800	Korku (kfq)
1810	Asuri (asr)
1812	Munda (unx)
1820	Mahali (mjx)
1822	Santali (sat)
1824	Turi (trd)
1835	Juang (jun)
1838	Kharai (khr)
1850	Sora (srb)

Mon-Khmer languages (1855-1949)

1855	Khasi (kha)
1860	Jah Hut (jah)
1870	Central Mnong (cmo)
1873	Bahnar (bdq)
1876	Koho (kpm)
1880	Laven (lbo)
1890	Upper Ta'oih (tth)
1895	Western Bru (brv)
1900	Central Khmer (khm)
1905	Mon (mnw)
1920	Khuen (khf)
1922	Khmu (kjg)
1925	Lua' (prb)
1927	Mal (mlf)
1929	Phai (prt)
1931	Phong-Kniang (pnx)
1935	Eastern Lawa (lwl)
1940	Ruching Palaung (pce)

1950	Muong (mtq)
1960	Vietnamese (vie)

Sino-Tibetan languages (1970-2429)

	<i>Chinese (1970-2059)</i>
1970	Chinese (macrolanguage) (zho)
1975	Gan Chinese (gan)
1980	Xiang Chinese (hsn)
1985	Hakka Chinese (hak)
1990	Jinyu Chinese (cyj)
2000	Mandarin Chinese (cmn)
2010	Min Bei Chinese (mnp)
2020	Min Dong Chinese (cdo)
2030	Min Nan Chinese (nan)
2040	Wu Chinese (wuu)
2050	Yue Chinese (yue)

Tibeto-Burman languages (2060-2429)

2060	Kulung (Nepal) (kle)
2065	Gamale Kham (kgj)
2075	Nepal Bhasa (new)
2080	Balti (bft)
2085	Dzongkha (dzo)
2090	Sikkimese (sip)
2092	Tshangla (tsj)
2095	Sherpa (xsr)
2100	Tibetan (bod)
2105	Khams Tibetan (khg)
2115	Gurung (gvr)
2120	Eastern Tamang (taj)
2135	Lepcha (lep)
2140	Adi (adi)
2145	Mising (mrg)
2150	Rawang (raw)

Burmese (2160-2179)

2160	Burmese (mya)
2165	Marma (rmz)
2170	Rakhine (rki)

2175	Zaiwa (atb)
2185	Lahu (lhu)
2190	Lisu (lis)
2195	Sichuan Yi (iii)
2200	Phunoi (pho)
2210	Ao Naga (njo)
2215	Tangkhu Naga (India) (nmf)

Kuki-Chin languages (2220-2299)

2220	Falam Chin (cfm)
2225	Haka Chin (cnh)
2230	Hmar (hmr)
2235	Lushai (lus)
2245	Daai Chin (dao)
2250	Kaang Chin (ckn)
2255	Matu Chin (hlt)
2265	Siyin Chin (csy)
2270	Tedim Chin (ctd)
2275	Zou (zom)
2285	Zotung Chin (czt)

2290 Zyphe Chin (zyp)

2300	Manipuri (mni)
2310	Toto (txo)
2320	Bodo (India) (brx)
2325	Garó (grt)
2340	Kachin (kac)

Karen languages (2350-2429)

2350	Karen languages (kar)
2355	Bwe Karen (bwe)
2360	Geko Karen (ghk)
2365	Western Kayah (kyu)
2370	Kayan (pdu)
2375	Mobwa Karen (jkm)
2380	Pa'O (ppa)
2385	Pwo Eastern Karen (kjp)
2390	S'gaw Karen (ksw)

Tai-Kadai languages (2430-2524)*Thai (2430-2474)*

2430	Thai (tha)
2435	Northeastern Thai (tts)
2440	Phu Thai (pht)
2445	Lü (khb)
2450	Nyaw (nyw)
2455	Shan (shn)
2460	Tai Dam (blt)
2465	Tai Daeng (tyr)

2475	Lao (lao)
2485	Bouyei (pcc)
2490	Zhuang (macrolanguage) (zha)
2500	Nung (Vietnam) (nut)
2505	Tày (tyz)
2515	Southern Dong (kmc)

Hmong-Mien languages (2525-2559)

2525 Iu Mien (ium)

Hmong (2535-2559)

2535	Hmong (macrolanguage) (hmn)
2536	Hmong Daw (mww)
2537	Hmong Njua (hnj)

Japonic languages (2560-2568)

2560	Japanese (jpn)
2563	Central Okinawan (ryu)
2566	Northern Amami-Oshima (ryn)

2569	Ainu (Japan) (ain)
2575	Korean (kor)

Siberian languages (2577-2589)

2577	Koryak (kpy)
2579	Itelmen (itl)

2582 Ket (ket)
 2585 Gilyak (niv)
 2588 Northern Yukaghir (ykg)

Austronesian languages (2590-3798)

2590 Buginese (bug)
 2600 Balinese (ban)
 2610 Malagasy (macrolanguage) (mlg)
 2611 Plateau Malagasy (plt)
 2630 Javanese (jav)
 2640 Biatah Bidayuh (bth)
 2650 Madurese (mad)
 2660 Achinese (ace)
 2670 Eastern Cham (cjm)
 2675 Western Cham (cja)
 2680 Chru (cje)
 2690 Jarai (jra)
 2700 Rade (rad)
 2710 Malay (macrolanguage) (msa)
 2715 Malay (individual language) (zlm)
 2716 Brunei (kxd)
 2730 Kubu (kvb)
 2740 Minangkabau (min)
 2750 Musi (mui)
 2760 Negeri Sembilan Malay (zmi)
 2770 Indonesian (ind)
 2780 Iban (iba)
 2790 Tetum (tet)
 2800 Central Melanau (mel)
 2810 Central Dusun (dtp)
 2820 Coastal Kadazan (kzj)
 2830 Kelabit (kzi)
 2840 Mainstream Kenyah (xkl)
 2850 Sundanese (sun)

Philippine languages (2870-3219)

2870 Pampanga (pam)
 2880 Bolinao (smk)
 2890 Botolan Sambal (sbl)
 2900 Sambal (xsb)

Tagalog (2910-2939)

2910 Filipino (fil)
 2920 Tagalog (tgl)

2940 Mandaya (mry)
 2950 Cebuano (ceb)
 2960 Hiligaynon (hil)
 2970 Masbatenyo (msb)
 2980 Waray (Philippines) (war)
 2990 Aklanon (akl)
 3000 Cuyonon (cyo)
 3010 Kinaray-A (krj)
 3020 Butuanon (btw)
 3030 Tausug (tsg)
 3040 Surigaonon (sgd)
 3050 Bikol (macrolanguage) (bik)
 3055 Central Bikol (bcl)

3060 Gorontalo (gor)
 3070 Obo Manobo (obo)
 3080 Ibaloi (ibl)
 3090 Pangasinan (pag)
 3100 Batad Ifugao (ifb)
 3110 Tuwali Ifugao (ifk)
 3120 Bontok (macrolanguage) (bnc)
 3125 Central Bontok (lbk)
 3130 Kankanaey (kne)
 3150 Iloko (ilo)
 3160 Ibanag (ibg)
 3170 Itawit (itv)
 3180 Ivatan (ivv)
 3190 Tondano (tdn)

Micronesian languages (3220-3419)

3220 Chamorro (cha)
 3250 Gilbertese (gil)
 3260 Kosraean (kos)
 3270 Marshallese (mah)
 3280 Carolinian (cal)
 3290 Mokilese (mkj)
 3300 Pingelapese (pif)
 3310 Mortlockese (mrl)
 3320 Nauru (nau)
 3330 Palauan (pau)
 3340 Pohnpeian (pon)
 3350 Chuukese (chk)
 3360 Puluwatese (puw)
 3370 Ulithian (uli)
 3380 Woleaian (woe)
 3390 Satawalese (stw)
 3400 Sonsorol (sov)
 3410 Yapese (yap)

Polynesian languages (3420-3798)

3420 Samoan (smo)
 3430 Kapingamarangi (kpg)
 3440 Nukuoro (nkr)
 3470 Tuvalu (tvl)
 3480 Tokelau (tkl)
 3490 Wallisian (wls)
 3500 Tonga (Tonga Islands) (ton)
 3510 Niuean (niu)
 3520 North Marquesan (mrq)
 3530 Rapanui (rap)
 3540 Tahitian (tah)
 3550 Rarotongan (rar)
 3560 Maori (mri)
 3570 Hawaiian (haw)
 3600 Eastern Malayo-Polynesian languages N.E.C.
 3610 Fijian (fij)
 3620 Rotuman (rtm)
 3630 Labo (mwi)
 3640 Hano (lml)
 3650 Mota (mtt)
 3660 Kwamera (tnk)
 3670 Lau (llu)

3680 Amara (aie)
 3690 Arifama-Miniafia (aai)
 3700 Kilivila (kij)
 3710 Motu (meu)
 3720 Kuanua (ksd)
 3730 Niwer Mil (hrc)
 3740 Penchal (pek)
 3750 Kola (kvv)
 3755 Sula (szn)
 3760 Boano (bpw)

OTHER AND UNSPECIFIED LANGUAGES (1000-1052, 1057-1063, 1074-1109, 1565-1642, 3799-9499, 9600-9999)

Creole languages (1000-1052, 1057-1063, 1074-1109)

English-based Creole languages (1000-1032, 1043-1050)

1000 Jamaican Creole English (jam)
 1001 Bajan (bjs)
 1003 Antigua and Barbuda Creole English (aig)
 1004 Bahamas Creole English (bah)
 1005 Belize Kriol English (bjz)
 1006 Grenadian Creole English (gcl)
 1007 Guyanese Creole English (gyn)
 1008 Tobagonian Creole English (tgh)
 1009 Trinidadian Creole English (trf)
 1010 Vincentian Creole English (svc)
 1012 Virgin Islands Creole English (vic)
 1015 Eastern Maroon Creole (djk)
 1016 Saramaccan (srm)
 1017 Sranan Tongo (srn)
 1020 Sea Island Creole (gul)
 1022 Hawai'i Creole English (hwc)
 1025 Krio (kri)
 1027 Cameroon Pidgin (wes)
 1028 Nigerian Pidgin (pcm)
 1030 Liberian English (lir)
 1043 Bislama (bis)
 1045 Pijin (pis)
 1047 Tok Pisin (tpi)

1033 Flaaitaal (fly)
 1034 Nubi (kcn)
 1038 Kituba (Democratic Republic of Congo) (ktu)
 1040 Sango (sag)
 1051 Hiri Motu (mho)
 1057 Louisiana Creole French (lou)
 1059 Morisyen (mfe)
 1061 R union Creole French (rcf)
 1064 Chavacano (cbk)
 1066 Papiamentu (pap)
 1074 Macanese (mzs)
 1077 Naga Pidgin (nag)
 1080 Manado Malay (xmm)
 1081 North Moluccan Malay (max)

1088 Chinook jargon (chn)
 1107 Pidgin N.E.C.
 1109 Creole N.E.C.

Uralic languages (1565-1602)

1565 Finnish (fin)
 1566 Karelian (krl)
 1570 Standard Estonian (ekk)
 1571 V oro (vro)
 1572 Liv (liv)
 1573 Votic (vot)
 1575 Inari Sami (smn)
 1576 Skolt Sami (sms)
 1577 Ume Sami (sju)
 1578 Lule Sami (smj)
 1579 Northern Sami (sme)
 1580 Pite Sami (sje)
 1582 Hungarian (hun)
 1584 Mari (mhg)
 1585 Eastern Mari (mhr)
 1587 Erzya (myv)
 1588 Moksha (mdf)
 1589 Komi-Permyak (koi)
 1590 Komi-Zyrian (kpv)
 1591 Udmurt (udm)
 1594 Samoyedic languages (syd)
 1595 Nenets (yrk)
 1596 Nganasan (nio)
 1597 Kamas (xas)
 1599 Khanty (kca)
 1600 Mansi (mns)

Caucasian languages (1603-1642)

1603 Georgian (kat)
 1605 Svan (sva)
 1606 Laz (lzz)
 1607 Mingrelian (xmf)
 1609 North Caucasian languages (ccn)
 1610 Avaric (ava)
 1612 Lak (lbe)
 1615 Lezghian (lez)
 1616 Tabassaran (tab)
 1620 Chechen (che)
 1621 Ingush (inh)
 1625 Dargwa (dar)
 1630 Abaza (abq)
 1631 Abkhazian (abk)
 1635 Adyghe (ady)
 1636 Kabardian (kbd)

3799 Papua New Guinea N.E.C.
 3800 Kuman (kdi)
 3810 Wahgi (wgi)
 3830 Orocolo (oro)
 3850 Enga (enq)
 3870 Southern Kiwai (kjd)
 3890 Mala (Papua New Guinea) (ped)
 3910 Waube (kop)

3930 Marind (mrz)
 3950 Laua (luf)
 3970 Wiru (wiu)
 4150 Tabla (tnm)
 4170 Qaqet (byx)
 4190 Gresi (grs)
 4200 Nimboran (nir)
 4240 Rapoisi (kyx)
 4270 Tanggu (tgu)
 4320 Sinagen (siu)
 4440 Eastern Arrernte (aer)
 4450 Yan-nhangu (jay)

Afro-Asiatic languages (4500-4899)

Semitic languages (4500-4644)

Arabic (4500-4544)

4500 Arabic (macrolanguage) (ara)
 4502 Algerian Arabic (arq)
 4504 Chadian Arabic (shu)
 4506 Cypriot Arabic (acy)
 4508 Egyptian Arabic (arz)
 4510 Gulf Arabic (afb)
 4512 Mesopotamian Arabic (acm)
 4514 Moroccan Arabic (ary)
 4516 North Levantine Arabic (apc)
 4518 Omani Arabic (acx)
 4520 Sanaani Arabic (ayn)
 4522 South Levantine Arabic (ajp)
 4524 Standard Arabic (arb)
 4526 Sudanese Arabic (apd)
 4528 Tunisian Arabic (aeb)
 4530 Hassaniyya (mey)
 4535 Maltese (mlt)
 4540 Judeo-Arabic (jrb)

4545 Hebrew (heb)
 4555 Syriac (macrolanguage) (syr)
 4560 Assyrian Neo-Aramaic (aii)
 4565 Chaldean Neo-Aramaic (cld)
 4570 Hulaulá (huy)
 4575 Turoyo (tru)
 4590 Amharic (amh)
 4600 Harari (har)
 4610 Silt'e (stv)
 4620 Sebat Bet Gurage (sgw)
 4630 Tigre (tig)
 4640 Tigrinya (tir)

4645 Coptic (cop)

Berber languages (4660-4719)

4660 Berber languages (ber)
 4661 Siwi (siz)
 4670 Central Atlas Tamazight (tzm)
 4675 Tachelhit (shi)
 4680 Kabyle (kab)

4685 Nafusi (jbn)
 4690 Tachawit (shy)
 4695 Tarifit (rif)
 4700 Tamashek (macrolanguage) (tmh)
 4701 Tamasheq (taq)
 4710 Zenaga (zen)

Chadic languages (4720- 4789)

4720 Bura-Pabir (bwr)
 4725 Huba (hbb)
 4735 Bata (bta)
 4740 Gidar (gid)
 4750 Hausa (hau)
 4760 Tangale (tan)
 4765 Cakfem-Mushere (cky)

Cushitic languages (4790-4899)

4790 Bilin (byn)
 4800 Afar (aar)
 4805 Saho (ssy)
 4815 Hadiyya (hdy)
 4825 Sidamo (sid)
 4830 Oromo (macrolanguage) (orm)
 4832 Borana-Arsi-Guji Oromo (gax)
 4834 Eastern Oromo (hae)
 4836 West Central Oromo (gaz)
 4840 Somali (som)
 4845 Garre (gex)
 4850 Maay (ymm)
 4855 Beja (bej)

Omotic languages (4860-4899)

4860 Aari (aiw)
 4870 Kafa (kbr)
 4880 Wolaytta (wal)

Nilo-Saharan languages (4900-5119)

4900 Mangbetu (mdj)
 4905 Lugbara (lgb)
 4910 Moru (mgd)
 4915 Ma'di (mhi)
 4919 Laka (Chad) (lap)
 4920 Ngambay (sba)
 4925 Sar (mwm)
 4930 Sara Kaba Náà (kwv)
 4935 Ama (Sudan) (nyi)
 4940 Didinga (did)
 4950 Kalenjin (macrolanguage) (kln)
 4953 Kipsigis (sgc)
 4955 Nandi (niq)
 4960 Pökoot (pko)
 4965 Dinka (macrolanguage) (din)
 4967 Southwestern Dinka (dik)
 4970 Nuer (nus)
 4975 Adhola (adh)
 4978 Acoli (ach)
 4980 Lango (Uganda) (laj)
 4985 Alur (alz)

- 4990 Luo (Kenya and Tanzania) (luo)
 4995 Anuak (anu)
 5000 Luwo (lwo)
 5005 Shilluk (shk)
 5010 Mabaan (mfz)
 5015 Bari (bfa)
 5020 Kakwa (keo)
 5030 Otuho (lot)
 5035 Masai (mas)
 5040 Teso (teo)
 5045 Karamojong (kdj)
 5048 Turkana (tuv)
 5055 Nobiin (fia)
 5060 Kanuri (macrolanguage) (kau)
 5061 Central Kanuri (knc)
 5063 Manga Kanuri (kby)
 5070 Tedaga (tuq)
 5075 Zaghawa (zag)
 5080 Amdang (amj)
 5085 Fur (fvr)
 5090 Katcha-Kadugli-Miri (xtc)
 5095 Kunama (kun)
 5100 Maba (Chad) (mde)
 5105 Masalit (mls)
 5110 Kibet (kie)
 5115 Koyraboro Senni Songhay (ses)
 5117 Zarma (dje)
- 5120 Korana (kqz)
 5125 Gana (gnk)
 5127 Nama (Namibia) (ylo)
 5130 Hadza (hts)
 5135 Sandawe (sad)
- Niger-Congo languages (5140-6701)*
- Bantoid languages (5140-5829)*
 5140 Bantu languages (bnt)
 5145 Samba Daka (ccg)
- Swahili (5150-5159)*
 5150 Swahili (macrolanguage) (swa)
 5151 Swahili (individual language) (swh)
 5152 Congo Swahili (swc)
- 5160 Ekajuk (eka)
 5170 Nkem-Nkum (isi)
 5180 Ejagham (etu)
 5185 Bankal (jjr)
 5190 Mama (mma)
 5195 Kenyang (ken)
 5200 Mbe (mfo)
 5205 Venda (ven)
 5210 Bembe (bmb)
 5220 Nyanga (nyj)
 5225 Mochi (old)
 5230 Kikuyu (kik)
- 5235 Meru (mer)
 5240 Embu (ebu)
 5245 Kamba (Kenya) (kam)
 5248 Giryama (nyf)
 5250 Sagalla (tga)
 5255 Sukuma (suk)
 5260 Nyamwezi (nym)
 5265 Hehe (heh)
 5270 Pangwa (pbr)
 5275 Zigula (ziw)
 5278 Mushungulu (xma)
 5280 Kongo (macrolanguage) (kon)
 5281 Koongo (kng)
 5283 Laari (ldi)
 5285 Kimbundu (kmb)
 5290 Haya (hay)
 5300 Luyia (macrolanguage) (luy)
 5301 Bukusu (bxk)
 5302 East Nyala (nle)
 5303 Idakho-Isukha-Tiriki (ida)
 5307 Logooli (rag)
 5310 Saamia (ism)
 5315 Gusii (guz)
 5320 Masaaba (myx)
 5330 Nyoro (nyo)
 5335 Nyankole (nyn)
 5340 Chiga (cgg)
 5345 Ganda (lug)
 5350 Soga (xog)
 5355 Gwere (gwr)
 5360 Tooro (ttj)
 5365 Kinyarwanda (kin)
 5370 Rundi (run)
 5375 Fuliuru (flr)
 5380 Nyemba (nba)
 5385 Luvale (lue)
 5390 Lozi (loz)
 5400 Kaonde (kqn)
 5405 Luba-Lulua (lua)
 5407 Luba-Katanga (lub)
 5410 Lunda (lun)
 5415 Bemba (Zambia) (bem)
 5420 Mambwe-Lungu (mgr)
 5425 Lamba (lam)
 5430 Ila (ilb)
 5435 Lenje (leh)
 5437 Tonga (Zambia) (toi)
 5445 Nyanja (nya)
 5455 Tonga (Nyasa) (tog)
 5460 Nsenga (nse)
 5462 Sena (seh)
 5465 Tumbuka (tum)
 5470 Makhwa (vmw)
 5475 Matumbi (mgw)
 5480 Yao (yao)
 5485 Makonde (kde)
 5490 Herero (her)
 5495 Umbundu (umb)

5500	Kuanyama (kua)	5765	Meta' (mgo)
5505	Xhosa (xho)	5770	Ngie (ngj)
5510	Zulu (zul)	5775	Ngwo (ngn)
5515	Swati (ssw)	5780	Babanki (bbk)
5520	North Ndebele (nde)	5785	Kom (Cameroon) (bkm)
5525	Shona (sna)	5790	Oku (oku)
5530	Ndau (ndc)	5795	Lamnso' (lns)
5535	Kalanga (kck)	5800	Kenswei Nsei (ndb)
5540	Tswana (tsn)	5805	Vengo (bav)
5545	Northern Sotho (nso)	5810	Aghem (agq)
5546	Southern Sotho (sot)	5815	Isu (Menchum Division) (isu)
5550	Tsonga (tso)	5820	Njebi (nzb)
5555	Bafia (ksf)		
5560	Basa (Cameroon) (bas)		<i>Mande languages (5830-5939)</i>
5565	Nyokon (nvo)	5830	Soninke (snk)
5570	Nomaande (lem)	5835	Southern Bobo Madaré (bwq)
5575	Bube (bvb)	5840	Kuranko (knk)
5580	Benga (bng)	5845	Mandingo (macrolanguage) (man)
5585	Mokpwe (bri)	5850	Mandinka (mnk)
5590	Duala (dua)	5853	Eastern Maninkakan (emk)
5595	Ewondo (ewo)	5855	Western Maninkakan (mlq)
5600	Eton (Cameroon) (eto)	5860	Bambara (bam)
5605	Bulu (Cameroon) (bum)	5865	Dyula (dyu)
5610	Fang (Equatorial Guinea) (fan)	5870	Mahou (mxx)
5615	Oroko (bdu)	5880	Kono (Sierra Leone) (kno)
5620	Mbo (zms)	5885	Vai (vai)
5625	Akoose (bss)	5890	Susu (sus)
5630	Bakaka (bqz)	5895	Yalunka (yal)
5633	Punu (puu)	5900	Kpelle (macrolanguage) (kpe)
5635	Ibali Teke (tek)	5901	Liberia Kpelle (xpe)
5640	Bangala (bxg)	5905	Bandi (bza)
5645	Lingala (lin)	5910	Mende (Sierra Leone) (men)
5650	Mabaale (mmz)	5915	Loma (Liberia) (lom)
5655	Bangi (bni)	5917	Toma (tod)
5660	Ntomba (nto)	5920	Bissa (bib)
5665	Mongo (lol)	5925	Dan (dnj)
5670	Ligenza (lgz)	5930	Mano (mev)
5675	Bwa (bww)	5935	Gagu (ggu)
5677	Tetela (tll)		
5680	Abon (abo)		<i>Fulah-Wolof languages (5940-5959)</i>
5683	Batu (btu)	5940	Fulah (macrolanguage) (ful)
5685	Tiv (tiv)	5942	Adamawa Fulfulde (fub)
5695	Bamileke languages (bai)	5945	Pulaar (fuc)
5696	Fe'fe' (fmp)	5946	Pular (fuf)
5700	Mengaka (xmg)	5950	Wolof (wol)
5705	Ngiemboon (nnh)		
5710	Ngwe (nwe)	5960	Serer (srr)
5715	Yemba (ybb)	5965	Wamey (cou)
5720	Bafut (bfd)	5970	Balanta-Kentohe (ble)
5721	Bambili-Bambui (baw)	5975	Jola-Fonyi (dyo)
5725	Mendankwe-Nkwen (mfd)	5980	Mandjak (mfv)
5730	Ngemba (nge)	5985	Gola (gol)
5735	Pinyin (pny)	5990	Krim (krm)
5740	Limbum (lmp)	5995	Southern Kisi (kss)
5745	Bamali (bbq)	6000	Timne (tem)
5750	Bamun (bax)	6005	West-Central Limba (lia)
5755	Medumba (byv)		
5760	Mungaka (mhk)		

Gur languages (6015-6119)

6015	Baatonum (bba)
6020	Nafaanra (nfr)
6025	Cebaara Sénoufo (sef)
6026	Djimini Sénoufo (dyi)
6035	Ntcham (bud)
6040	Gourmanchéma (gux)
6045	Buli (Ghana) (bwu)
6050	Dagbani (dag)
6055	Kusaal (kus)
6060	Mampruli (maw)
6065	Mossi (mos)
6070	Farefare (gur)
6075	Southern Dagaare (dga)
6080	Wali (Ghana) (wlx)
6085	Kaansa (gna)
6090	Kasem (xsm)
6095	Kabiyè (kbp)
6098	Tem (kdh)
6100	Paasaal (sig)
6105	Vagla (vag)
6110	Wara (wbf)

*Kwa languages (6120-6249)**Akan (6120-6134)*

6120	Akan (macrolanguage) (aka)
6125	Fanti (fat)
6130	Twi (twi)

6135	Anyin (any)
6140	Baoulé (bci)
6145	Sehwi (sfw)
6150	Nzima (nzi)
6155	Gonja (gjn)
6160	Krache (kye)
6165	Nawuri (naw)
6170	Awutu (afu)
6175	Larteh (lar)
6180	Logba (lgq)
6185	Sekpele (lip)
6190	Adiukrou (adj)
6195	Attié (ati)
6200	Adangme (ada)
6205	Ga (gaa)
6210	Adangbe (adq)
6215	Ikposo (kpo)
6220	Aja (Benin) (ajg)
6225	Gun (guw)
6230	Ewe (ewe)
6235	Fon (fon)
6240	Gen (gej)
6245	Waci Gbe (wci)

Ijoid languages (6250-6284)

6250	Southeast Ijo (ijs)
6255	Ibani (iby)
6260	Kalabari (ijn)
6265	Kirike (okr)

6270	Izon (ijc)
6275	Biseni (ije)
6280	Arigidi (aqg)

Defoid languages (6285-6299)

6285	Igala (igl)
6290	Yoruba (yor)
6293	Ifè (ife)
6295	Isekiri (its)

Edoid languages (6300-6356)

6300	Bini (bin)
6305	Esan (ish)
6310	Emai-Iuleha-Ora (ema)
6315	Epie (epi)
6320	Ivbie North-Okpela-Arhe (atg)
6325	Uneme (une)
6330	Yekhee (ets)
6335	Iyayu (iya)
6340	Isoko (iso)
6345	Okpe (Southwestern Edo) (oke)
6350	Urhobo (urh)
6355	Uvbie (evh)

Idomoid languages (6357-6369)

6357	Idoma (idu)
6360	Igede (ige)
6363	Yala (yba)
6365	Ekpeye (ekp)

Igboid languages (6370-6389)

6370	Igbo (ibo)
6375	Ika (ikk)
6378	Ikwere (ikw)
6381	Ogbah (ogc)
6384	Ukwuani-Aboh-Ndoni (ukw)

6390	Jibu (jib)
6395	Kutep (kub)
6400	Bauchi (bsf)
6405	Bina (Nigeria) (byj)
6408	Tumi (kku)
6410	Ebira (igb)
6415	Gade (ged)
6420	Gbagyi (gbr)
6425	Nupe-Nupe-Tako (nup)
6435	Adara (kad)
6440	Berom (bom)
6445	Eggon (ego)
6450	Jju (kaj)
6455	Ukaan (kcf)

Kru languages (6460-6499)

6460	Wané (hwa)
6465	Kuwaa (blh)
6470	Bassa (bsq)
6475	Grebo (macrolanguage) (grb)
6476	Northern Grebo (gbo)

6480 Klao (klu)
 6485 Eastern Krahn (kqo)
 6486 Western Krahn (krw)
 6490 Sapo (krn)

6500 Efik (efi)
 6505 Anaang (anw)
 6510 Ibibio (ibb)
 6515 Ekit (eke)
 6520 Ibino (ibn)
 6525 Obolo (ann)
 6530 Oro (orx)
 6535 Eleme (elm)
 6540 Gokana (gkn)
 6545 Khana (ogo)
 6550 Tee (tkq)
 6555 Abua (abn)
 6560 Agwagwune (yay)
 6563 Kohumono (bcs)
 6565 Legbo (agb)
 6570 Lokaa (yaz)
 6575 Olulumo-Ikom (iko)
 6580 Bete-Bendi (btt)
 6585 Bokyi (bky)
 6590 Baka (Cameroon) (bkc)
 6595 Mayogo (mdm)
 6600 Gbaya (gba)
 6601 Northwest Gbaya (gya)
 6605 Zande (individual language) (zne)
 6610 Day (dai)
 6615 Mbum (mdd)
 6620 Samba Leko (ndi)
 6625 Teme (tdo)
 6630 Kaan (ldl)
 6635 Tula (tul)
 6640 Jamsay Dogon (djm)
 6650 Moro (mor)

6702 Benin N.E.C.
 6707 Cameroon N.E.C.
 6716 Eritrea N.E.C.
 6717 Ethiopia N.E.C.
 6719 Gambia N.E.C.
 6720 Ghana N.E.C.
 6721 Guinea N.E.C.
 6723 Kenya N.E.C.
 6725 Liberia N.E.C.
 6728 Mali N.E.C.
 6733 Mozambique N.E.C.
 6734 Namibia N.E.C.
 6735 Niger N.E.C.
 6736 Nigeria N.E.C.
 6738 Senegal N.E.C.
 6740 Sierra Leone N.E.C.
 6745 Sudan N.E.C.
 6747 Tanzania N.E.C.
 6750 Uganda N.E.C.
 6795 Africa N.E.C.

Native North American languages (6800-7099)

Aleut languages (6800-6814)

6800 Aleut (ale)
 6801 Inupiaq (macrolanguage) (ipk)
 6802 North Alaskan Inupiatun (esi)
 6803 Northwest Alaska Inupiatun (esk)
 6804 Eastern Canadian Inuktitut (ike)
 6806 Greenlandic (kal)
 6808 Pacific Gulf Yupik (ems)
 6809 Central Siberian Yupik (ess)
 6810 Central Yupik (esu)

Algonquian languages (6815-6848)

6815 Algonquin (alq)
 6817 Mohegan-Pequot (xpq)
 6818 Narragansett (xnt)
 6819 Powhatan (pim)
 6820 Wampanoag (wam)
 6821 Lumbee (lmz)
 6822 Arapaho (arp)
 6823 Gros Ventre (ats)
 6824 Siksika (bla)
 6825 Cheyenne (chy)
 6827 Cree (cre)
 6828 Michif (crg)
 6829 Delaware (del)
 6830 Munsee (umu)
 6831 Unami (unm)
 6832 Nanticoke (nnt)
 6833 Meskwaki (sac)
 6834 Kickapoo (kic)
 6835 Menominee (mez)
 6837 Miami (mia)
 6838 Mi'kmaq (mic)
 6839 Ojibwa (oji)
 6840 Chippewa (ciw)
 6841 Ottawa (otw)
 6842 Severn Ojibwa (ojs)
 6843 Potawatomi (pot)
 6844 Malecite-Passamaquoddy (pqm)
 6845 Eastern Abnaki (aaq)
 6846 Western Abnaki (abe)
 6848 Shawnee (sjw)

6849 Wiyot (wiy)
 6850 Yurok (yur)
 6856 Kutenai (kut)
 6857 Quileute (qui)
 6859 Kwakiutl (kwk)
 6861 Ditidaht (dtd)
 6862 Nuu-chah-nulth (nuk)
 6863 Makah (myh)

Salish languages (6866-6891)

6866 Quinault (qun)
 6867 Lower Chehalis (cea)

6868 Upper Chehalis (cjh)
 6869 Cowlitz (cow)
 6872 Clallam (clm)
 6873 Coeur d'Alene (crd)
 6874 Columbia-Wenatchi (col)
 6875 Okanagan (oka)
 6876 Kalispel-Pend d'Oreille (fla)
 6877 Spokane (spo)
 6878 Shuswap (shs)
 6879 Halkomelem (hur)
 6880 Straits Salish (str)
 6881 Nooksack (nok)
 6883 Lushootseed (lut)
 6884 Skagit (ska)
 6885 Snohomish (sno)
 6886 Southern Puget Sound Salish (slh)
 6888 Tillamook (til)
 6890 Twana (twa)

6892 Haida (macrolanguage) (hai)
 6893 Northern Haida (hdn)
 6894 Southern Haida (hax)

Eyak-Athabaskan languages (6895-6935)

6895 Athapascan languages (ath)
 6897 Eyak (eya)
 6898 Tlingit (tli)
 6900 Slave (Athapascan) (den)
 6902 Chilcotin (clc)
 6903 Ahtena (aht)
 6904 Tanaina (tfn)
 6905 Degexit'an (ing)
 6906 Koyukon (koy)
 6907 Holikachuk (hoi)
 6908 Upper Kuskokwim (kuu)
 6909 Lower Tanana (taa)
 6910 Tanacross (tcb)
 6911 Upper Tanana (tau)
 6912 Northern Tutchone (ttm)
 6913 Southern Tutchone (tce)
 6914 Gwich'in (gwi)
 6915 Han (haa)
 6916 Chetco (ctc)
 6917 Coquille (coq)
 6918 Galice (gce)
 6919 Tolowa (tol)
 6920 Tututni (tuu)
 6921 Hupa (hup)
 6922 Kato (ktw)
 6923 Mattole (mvb)
 6924 Wailaki (wlk)
 6927 Jicarilla Apache (apj)
 6928 Lipan Apache (apl)
 6929 Mescalero-Chiricahua Apache (apm)
 6930 Western Apache (apw)
 6931 Kiowa Apache (apk)
 6933 Navajo (nav)

Kiowa-Tanoan languages (6936-6941)

6936 Kiowa (kio)
 6937 Northern Tiwa (twf)
 6938 Southern Tiwa (tix)
 6939 Tewa (USA) (tew)
 6940 Jemez (tow)

Penutian languages (6942-6976)

6942 Northeast Maidu (nmu)
 6943 Northwest Maidu (mjd)
 6944 Nisenan (nsz)
 6946 Northern Ohlone (cst)
 6948 Coast Miwok (csi)
 6949 Plains Miwok (pmw)
 6950 Central Sierra Miwok (csm)
 6951 Northern Sierra Miwok (nsq)
 6952 Southern Sierra Miwok (skd)
 6955 Nomlaki (nol)
 6956 Patwin (pwi)
 6957 Wintu (wnw)
 6960 Yokuts (yok)
 6961 Kalapuya (kyl)
 6962 Siuslaw (sis)
 6963 Klamath-Modoc (kla)
 6965 Nez Perce (nez)
 6966 Tenino (tqn)
 6967 Umatilla (uma)
 6968 Walla Walla (waa)
 6969 Yakama (yak)
 6971 Chinook (chh)
 6972 Wasco-Wishram (wac)
 6974 Gitksan (git)
 6975 Nisga'a (ncg)
 6976 Tsimshian (tsi)

Hokan languages (6977-7009)

6978 Achumawi (acv)
 6979 Atsugewi (atw)
 6980 Karok (kyh)
 6981 Central Pomo (poo)
 6982 Kashaya (kju)
 6983 Southern Pomo (peq)
 6984 Northern Pomo (pej)
 6985 Eastern Pomo (peb)
 6986 Northeastern Pomo (pef)
 6987 Southeastern Pomo (pom)
 6988 Shasta (sht)
 6989 Washo (was)
 6990 Maricopa (mrc)
 6991 Mohave (mov)
 6992 Quechan (yum)
 6994 Kumiai (dih)
 6995 Cocopa (coc)
 6997 Havasupai-Walapai-Yavapai (yuf)
 7000 Barbareño (boi)
 7001 Ineseño (inz)
 7002 Purisimeño (puy)
 7003 Ventureño (veo)

7004 Cruzeño (crz)
 7005 Obispeño (obi)
 7007 Chimariko (cid)
 7008 Esselen (esq)
 7009 Salinan (sln)

7010 Tonkawa (tqw)
 7011 Yuchi (yuc)

Siouan languages (7013-7027)

7013 Crow (cro)
 7014 Hidatsa (hid)
 7015 Mandan (mhq)
 7016 Catawba (chc)
 7017 Assiniboine (asb)
 7018 Dakota (dak)
 7019 Lakota (lkt)
 7020 Stoney (sto)
 7022 Iowa-Oto (iow)
 7023 Ho-Chunk (win)
 7024 Kansa (ksk)
 7025 Omaha-Ponca (oma)
 7026 Osage (osa)
 7027 Quapaw (qua)

Muskogean languages (7028-7033)

7028 Alabama (akz)
 7029 Koasati (cku)
 7030 Mikasuki (mik)
 7031 Creek (mus)
 7032 Choctaw (cho)
 7033 Chickasaw (cic)

7034 Chitimacha (ctm)
 7036 Yuki (yuk)
 7037 Wappo (wao)
 7039 Eastern Keres (kee)
 7040 Western Keres (kjq)

Iroquoian languages (7041-7050)

7042 Iroquoian languages (iro)
 7043 Mohawk (moh)
 7044 Oneida (one)
 7045 Onondaga (ono)
 7046 Cayuga (cay)
 7047 Seneca (see)
 7048 Wyandot (wya)
 7049 Tuscarora (tus)
 7050 Cherokee (chr)

Caddoan languages (7051-7055)

7052 Caddo (cad)
 7053 Arikara (ari)
 7054 Pawnee (paw)
 7055 Wichita (wic)

7059 Zuni (zun)

Uto-Aztecan languages (7060-7124)

7060 Comanche (com)
 7062 Shoshoni (shh)
 7064 Panamint (par)
 7066 Mono (USA) (mnr)
 7068 Northern Paiute (pao)
 7069 Ute-Southern Paiute (ute)
 7070 Kawaiisu (xaw)
 7072 Cahuilla (chl)
 7074 Cupeño (cup)
 7076 Luiseno (lui)
 7078 Serrano (ser)
 7080 Hopi (hop)
 7082 Tübatulabal (tub)
 7084 Tohono O'odham (ood)
 7100 Southeastern Tepehuan (stp)
 7103 Eastern Huasteca Nahuatl (nhe)
 7105 Michoacán Nahuatl (ncl)
 7109 El Nayar Cora (crn)
 7111 Huichol (hch)
 7115 Central Tarahumara (tar)
 7120 Mayo (mfy)
 7123 Yaqui (yaq)

Central and South American languages (7100-7999)

7125 Matagalpa (mtn)
 7126 Mayangna (yan)
 7127 Mískito (miq)

Mayan languages (7130-7181)

7130 Huastec (hus)
 7135 Chol (ctu)
 7136 Chortí (caa)
 7137 Tzeltal (tzh)
 7138 Tzotzil (tzo)
 7143 Aguacateco (agu)
 7144 Ixil (ixl)
 7150 Mam (mam)
 7152 Tektiteko (ttc)
 7155 Kaqchikel (cak)
 7156 Tz'utujil
 7160 K'iche' (quc)
 7161 Kekchí (kek)
 7162 Poqomam (poc)
 7163 Poqomchi' (poh)
 7166 Chuj (cac)
 7170 Popti' (jac)
 7171 Q'anjob'al (kjb)
 7172 Western Kanjobal (knj)
 7175 Itzá (itz)
 7177 Mopán Maya (mop)
 7179 Lacandon (lac)
 7180 Yucateco (yua)

7182	Purepecha (tsz)	<i>Other and unspecified languages (8000-9999)</i>
7185	Huilliche (huh)	8000 Basque (eus)
7186	Mapudungun (arn)	8010 Esperanto (epo)
7188	Highland Totonac (tos)	8020 Shelta (sth)
		8030 Latin (lat)
	<i>Otomanguean languages (7200-7399)</i>	8050 American Indian N.E.C.
7200	Guerrero Amuzgo (amu)	8055 Caribbean N.E.C.
7210	Copala Triqui (trc)	8060 Eastern Europe N.E.C.
7220	Alacatlazala Mixtec (mim)	8061 Europe N.E.C.
7230	Mixtepec Mixtec (mix)	8089 Specified not listed
7240	Silacayoapan Mixtec (mks)	9999 Unspecified
7250	Mixtec N.E.C.	
7260	Tepeuxila Cuicatec (cux)	
7270	Huautla Mazatec (mau)	
7280	Ixcatec (ixc)	
7290	San Juan Atzingo Popoloca (poe)	
7300	Zapotec (zap)	
7310	Western Tlacolula Valley Zapotec (twb)	
7320	Zoogocho Zapotec (zpq)	
7330	Central Mazahua (maz)	
7340	Mezquital Otomi (ote)	
7350	Ojitlán Chinantec (chj)	
7400	Quechua (que)	
7500	Aymara (aym)	
7501	Central Aymara (ayr)	
7505	Galibi Carib (car)	
7515	Wayana (way)	
7525	Waiwai (waw)	
7535	Guahibo (guh)	
7540	Asháninka (cni)	
7550	Arawak (arw)	
7555	Garifuna (cab)	
7560	Taino (tnq)	
7565	Wayuu (guc)	
7575	Wapishana (wap)	
7585	Yucuna (ycn)	
7600	Northern Emberá (emp)	
7635	Boruca (brn)	
7640	Ngäbere (gym)	
7645	Arhuaco (arh)	
7648	Cogui (kog)	
7650	San Blas Kuna (cuk)	
7655	Cofán (con)	
7660	Paumarí (pad)	
7665	Mondé (mnd)	
7670	Tuparí (tpr)	
7680	Guarani (grn)	
7682	Paraguayan Guaraní (gug)	
7690	Sirionó (srq)	
7700	Xingú Asuriní (asn)	
7710	Cocama-Cocamilla (cod)	
7715	Nhengatu (yrl)	
7735	Tehuelche (teh)	
7800	Argentina N.E.C.	
7810	Guatemala N.E.C.	

Appendix B. Public Use Microdata Sample (PUMS) LANP-LANP12 Categories and Comparability, 2012-2016

PUMS (Public Use Microdata Sample) categories are created for use with Public Use files, and must meet a disclosure avoidance threshold of 10,000 speakers nationwide. For this reason, similar individual languages are often collapsed together within a larger PUMS category in order for the data to be publishable. The 2012-2015 PUMS codes and labels are displayed in columns A and B, with corresponding detailed language codes that make up each PUMS code shown in columns C and D. The 2016 PUMS codes and labels are displayed in columns E and F, with corresponding detailed language codes that make up each PUMS code shown in columns G and H.

Columns A through D (2012-2015 ACS PUMS Code List-Language) refer to codes for the LANP variable in the 2012, 2013, 2014 and 2015 1-year PUMS datasets, and LANP12 in the 2012-2015 portion of 5-year PUMS datasets from 2008-2012 to 2012-2015. Columns E through H (2016 ACS PUMS Code List-Language) refer to codes for LANP in the 2016 1-year PUMS and 2012-2016 5-year PUMS, and HHLANP in the 2016 1-year PUMS.

2012-2015 ACS PUMS Code List-Language				2016 ACS PUMS Code List-Language			
PUMS LANP12 Code	PUMS Label	Language Microdata Code	Corresponding Detailed Language	2016 PUMS LANP/HHLANP Code	2016 PUMS Label	2016 Language Microdata Code	Corresponding Detailed Language and three-letter ISO-639-3 code
601	Jamaican Creole	601	Jamaican Creole	1000	Jamaican Creole English	1000	Jamaican Creole English (jam)
						1005	Belize Kriol English (bzi)
602	Krio	602	Krio	1025	Other English-based Creole languages	1001	Bajan (bjs)
						1003	Antigua and Barbuda Creole English (aig)
						1004	Bahamas Creole English (bah)
						1006	Grenadian Creole English (gcl)
						1007	Guanese Creole English (gyn)
						1008	Tobagonian Creole English (tgh)
						1009	Trinidadian Creole English (trf)
						1010	Vincentian Creole English (svc)
						1012	Virgin Islands Creole English (vic)
						1015	Eastern Maroon Creole (djk)
						1016	Saramaccan (srm)
						1017	Sranan Tongo (srn)
						1020	Sea Island Creole English (gul)
						1022	Hawai'i Creole English (hwc)
						1025	Krio (kri)
						1027	Cameroon Pidgin (wes)
						1028	Nigerian Pidgin (pcm)
						1030	Liberian English (lir)
						1043	Bislama (bis)
						1045	Pijin (pis)
						1047	Tok Pisin (tpi)
				1055	Haitian	1053	Guadeloupean Creole French (gcf)
						1054	Saint Lucian Creole French (acf)
						1055	Haitian (hat)
				1069	Kabuverdianu	1069	Kabuverdianu (kea)
						1072	Upper Guinea Crioulo (pov)
607	German	607	German	1110	German	1110	German (deu)
		613	Luxembourgian			1111	Bavarian (bar)
						1112	Hutterite German (geh)
						1113	Low German (nds)
						1114	Plautdietsch (pdt)
						1115	Swabian (swg)
						1121	Upper Saxon (sxu)
						1122	Limburgish (lim)
						1123	Luxembourgish (ltz)
				1120	Swiss German	1120	Swiss German (gsw)
608	Pennsylvania Dutch	608	Pennsylvania Dutch	1125	Pennsylvania German	1125	Pennsylvania German (pdc)
609	Yiddish	609	Yiddish	1130	Yiddish	1130	Yiddish (yid)
						1131	Eastern Yiddish (ydd)
610	Dutch	610	Dutch	1132	Dutch	1132	Dutch (nld)
						1133	Vlaams (vls)
611	Afrikaans	611	Afrikaans	1134	Afrikaans	1033	Flaaitaal (flv)
						1134	Afrikaans (afr)
614	Swedish	614	Swedish	1140	Swedish	1140	Swedish (swe)
615	Danish	615	Danish	1141	Danish	1141	Danish (dan)
616	Norwegian	616	Norwegian	1142	Norwegian	1142	Norwegian (nor)
619	Italian	619	Italian	1155	Italian	1155	Italian (ita)
						1156	Corsican (cos)
						1157	Friulian (fur)
						1158	Ligurian (lij)
						1159	Lombard (lmo)
						1160	Neapolitan (nap)
						1161	Piemontese (pms)
						1162	Romagnol (rgn)
						1163	Sardinian (srd)
						1164	Sicilian (scn)
						1165	Venetian (vec)
620	French	620	French	1170	French	1170	French (fra)
		621	Provençal			1172	Jèrriais (nrf)
						1174	Walloon (wln)
622	Patois	622	Patois				
623	French Creole	623	French Creole				
624	Cajun	624	Cajun	1175	Cajun French	1175	Cajun French (frc)
625	Spanish	625	Spanish	1200	Spanish	1200	Spanish (spa)
		627	Ladino			1201	Asturian (ast)
		628	Pachuco			1202	Ladino (lad)
						1205	Caló (rmg)
629	Portuguese	629	Portuguese	1210	Portuguese	1210	Portuguese (por)
		630	Papia Mentae			1211	Galician (glg)
631	Romanian	631	Romanian	1220	Romanian	1220	Romanian (ron)

						1221	Istro Romanian (ruo)
635	Irish Gaelic	635	Irish Gaelic	1231	Irish	1231	Irish (gle)
637	Greek	637	Greek	1235	Greek	1235	Modern Greek (1453-) (ell)
						1238	Pontic (pnt)
638	Albanian	638	Albanian	1242	Albanian	1242	Albanian (sqi)
						1243	Arbëreshë Albanian (aae)
						1244	Gheg Albanian (ain)
						1245	Tosk Albanian (als)
639	Russian	639	Russian	1250	Russian	1250	Russian (rus)
641	Ukrainian	641	Ukrainian	1260	Ukrainian	1252	Rusyn (rue)
						1260	Ukrainian (ukr)
642	Czech	642	Czech	1262	Czech	1262	Czech (ces)
646	Slovak	646	Slovak	1263	Slovak	1263	Slovak (slk)
645	Polish	645	Polish	1270	Polish	1270	Polish (pol)
647	Bulgarian	647	Bulgarian	1273	Bulgarian	1273	Bulgarian (bul)
648	Macedonian	648	Macedonian	1274	Macedonian	1274	Macedonian (mkd)
649	Serbocroatian	649	Serbocroatian	1275	Serbocroatian	1275	Serbocroatian (hbs)
						1276	Bosnian (bos)
650	Croatian	650	Croatian	1277	Croatian	1277	Croatian (hrv)
651	Serbian	651	Serbian	1278	Serbian	1278	Serbian (srp)
653	Lithuanian	653	Lithuanian	1281	Lithuanian	1281	Lithuanian (lit)
654	Latvian	654	Latvian	1283	Latvian	1283	Standard Latvian (lvs)
						1284	Latgalian (ltg)
655	Armenian	655	Armenian	1288	Armenian	1288	Armenian (hye)
656	Persian	656	Persian	1290	Farsi	1290	Persian (fas)
						1295	Iranian Persian (pes)
						1304	Judeo-Persian (jpr)
						1292	Dari (prs)
						1303	Hazaragi (haz)
658	Kurdish	658	Kurdish	1315	Kurdish	1315	Kurdish (kur)
						1316	Central Kurdish (ckb)
						1317	Northern Kurdish (kmr)
657	Pashto	657	Pashto	1327	Pashto	1327	Pushto (pus)
						1328	Northern Pashto (pbu)
						1329	Southern Pashto (pbt)
662	India N.E.C.	662	India N.E.C.	1340	India N.E.C.	1340	India N.E.C.
663	Hindi	663	Hindi	1350	Hindi	1341	Chhattisgarhi (hne)
						1342	Garhwali (gbm)
						1343	Haryanvi (bgc)
						1349	Awadhi (awa)
						1350	Hindi (hin)
						1351	Fiji Hindi (hif)
671	Urdu	671	Urdu	1360	Urdu	1360	Urdu (urd)
664	Bengali	664	Bengali	1380	Bengali	1380	Bengali (ben)
						1382	Chittagonian (ctg)
						1384	Rangpuri (rkt)
						1386	Sylheti (syl)
665	Panjabi	665	Panjabi	1420	Punjabi	1420	Paniabi (pan)
						1421	Lahnda (lah)
						1422	Western Panjabi (pnb)
						1423	Pahari-Potwari (phr)
						1424	Northern Hindko (hno)
						1425	Southern Hindko (hnd)
						1426	Saraiki (skr)
						1430	Goan Konkani (gom)
						1435	Konkani (knn)
666	Marathi	666	Marathi	1440	Marathi	1440	Marathi (mar)
667	Gujarati	667	Gujarati	1450	Gujarati	1450	Gujarati (gui)
						1455	Saurashtra (saz)
674	Nepali	674	Nepali	1500	Nepali	1500	Nepali (nep)
						1501	Dotyali (dty)
675	Sindhi	675	Sindhi				
676	Pakistan N.E.C.	676	Pakistan N.E.C.	1525	Pakistan N.E.C.	1525	Pakistan N.E.C.
677	Sinhalese	677	Sinhalese	1530	Sinhala	1530	Sinhala (sin)
985	Other Indo-European languages	603	Hawaiian Pidgin				
		604	Pidgin				
		605	Gullah				
		606	Saramacca				
		612	Frisian				
		617	Icelandic				
		618	Faroese				
		626	Catalonian				
		632	Rhaeto-Romanic				
		633	Welsh				
		634	Breton				
		636	Scottic Gaelic				
		640	Belorussian				
		643	Kashubian				
		644	Lusatian				
		652	Slovene				
		659	Balochi				
		660	Tadzhik				
		661	Ossete				
		668	Bihari				
		669	Rajasthani				
		670	Oriya				
		672	Assamese				
		673	Kashmiri				
		678	Romany				
				1540	Other Indo-Iranian languages	1300	Tajik (tgk)
						1302	Bukharic (bhh)
						1308	Judeo-Tat (jdt)
						1309	Northern Luri (lrc)
						1310	Parsi (prp)
						1311	Zoroastrian Dari (qzb)
						1312	Dezfuli (def)

					1320 Zaza (zza)	
					1321 Baluchi (bal)	
					1322 Eastern Balochi (bqp)	
					1323 Southern Balochi (bcc)	
					1325 Ossetian (oss)	
					1331 Shughni (sqh)	
					1333 Wakhi (wbl)	
					1365 Deccan (dcc)	
					1370 Memoni (mby)	
					1375 Assamese (asm)	
					1388 Halbi (hbl)	
					1400 Rohingya (rhg)	
					1442 Od (odk)	
					1460 Bhojpuri (bho)	
					1463 Maqahi (maq)	
					1467 Maithili (mai)	
					1470 Rajasthani (raj)	
					1472 Bagri (bqg)	
					1474 Dhatki (mki)	
					1478 Lambadi (lmn)	
					1480 Marwari (mwr)	
					1481 Marwari (India) (rwr)	
					1482 Dhundari (dhd)	
					1486 Bodo Paria (bdv)	
					1488 Desiya (dso)	
					1490 Oriya (ory)	
					1491 Sambalpuri (spv)	
					1495 Kashmiri (kas)	
					1497 Khowar (khw)	
					1499 Dangaura Tharu (thl)	
					1510 Kachhi (kfr)	
					1515 Sindhi (snd)	
					1518 Dogri (dgo)	
					1521 Kumaoni (kyf)	
					1526 Dhivehi (div)	
					1540 Romany (rom)	
					1541 Balkan Romani (rmn)	
					1542 Carpathian Romani (rmc)	
					1543 Vlax Romani (rmy)	
					1135 Northern Frisian (frr)	
					1136 Western Frisian (fry)	
					1143 Icelandic (isl)	
					1144 Faroese (fao)	
					1147 Belgium N.E.C.	
					1176 Occitan (post 1500) (oci)	
					1206 Catalan (cat)	
					1218 Aromanian (aen)	
					1223 Ladin (lld)	
					1225 Romansh (roh)	
					1228 Welsh (cym)	
					1229 Breton (bre)	
					1230 Cornish (cor)	
					1232 Manx (glv)	
					1233 Scottish Gaelic (gla)	
					1234 Scots (sco)	
					1251 Belarusian (bel)	
					1265 Lower Sorbian (dsb)	
					1266 Upper Sorbian (hsb)	
					1268 Kashubian (csb)	
					1269 Silesian (szl)	
					1280 Slovenian (slv)	
					1565 Finnish (fin)	
					1582 Hungarian (hun)	
679	Finnish	679	Finnish	1565	Finnish	
682	Hungarian	682	Hungarian	1582	Hungarian	
689	Uighur	689	Uighur			
691	Turkish	691	Turkish	1675	Turkish	1675 Turkish (tur)
						1676 Balkan Gagauz Turkish (bgx)
						1677 Gagauz (gaq)
694	Mongolian	694	Mongolian	1690	Mongolian	1690 Mongolian (mon)
						1691 Halh Mongolian (khk)
						1692 Peripheral Mongolian (mvf)
701	Telugu	701	Telugu	1730	Telugu	1730 Telugu (tel)
702	Kannada	702	Kannada	1737	Kannada	1735 Badaga (bfq)
						1737 Kannada (kan)
703	Malayalam	703	Malayalam	1750	Malayalam	1750 Malayalam (mal)
						1755 Paniya (pcg)
704	Tamil	704	Tamil	1765	Tamil	1765 Tamil (tam)
726	Mon-Khmer, Cambodian	726	Mon-Khmer, Cambodian	1900	Khmer	1900 Central Khmer (khm)
728	Vietnamese	728	Vietnamese	1960	Vietnamese	1960 Vietnamese (vie)
708	Chinese	708	Chinese	1970	Chinese	1970 Chinese (zho)
						1975 Gan Chinese (gan)
						1980 Xiang Chinese (hsn)
						1985 Hakka Chinese (hak)
						1990 Jinyu Chinese (cjy)
						2010 Min Bei Chinese (mnp)
						2020 Min Dong Chinese (cdo)
						2040 Wu Chinese (wuu)
712	Mandarin	712	Mandarin	2000	Mandarin	2000 Mandarin Chinese (cmn)
714	Formosan	714	Formosan	2030	Min Nan Chinese	2030 Min Nan Chinese (nan)
711	Cantonese	711	Cantonese	2050	Cantonese	2050 Yue Chinese (yue)
						2095 Sherpa (xsr)
						2100 Tibetan (bod)
						2105 Khams Tibetan (khq)
717	Burmese	717	Burmese	2160	Burmese	2160 Burmese (mya)
						2165 Marma (rmz)

			2270	Chin languages	2170 Rakhine (rki) 2220 Falam Chin (cfm) 2225 Haka Chin (cnh) 2230 Hmar (hmr) 2235 Lushai (lus) 2245 Daai Chin (dao) 2250 Kaang Chin (ckn) 2255 Matu Chin (hlt) 2265 Siyin Chin (csy) 2270 Tedim Chin (ctd) 2275 Zou (zom) 2285 Zotung Chin (czt) 2290 Zyphe Chin (zyp)
			2350	Karen languages	2350 Karen languages (kar) 2355 Bwe Karen (bwe) 2360 Geko Karen (ghk) 2365 Western Kayah (kyu) 2370 Kayan (pdu) 2375 Mobwa Karen (ikm) 2380 Pa'O (ppa) 2385 Pwo Eastern Karen (kip) 2390 S'gaw Karen (ksw)
720	Thai	720	Thai	2430	Thai (tha) 2435 Northeastern Thai (tts) 2440 Phu Thai (pht) 2445 Lü (khh) 2450 Nyaw (myw) 2455 Shan (shn) 2460 Tai Dam (blt) 2465 Tai Daeng (tyr)
725	Laotian	725	Laotian	2475	Lao (lao)
721	Mien	721	Mien	2525	Iu Mien (ium)
722	Hmong	722	Hmong	2535	Hmong (hmn) 2536 Hmong Daw (mww) 2537 Hmong Niua (hni)
723	Japanese	723	Japanese	2560	Japanese (jpn) 2563 Central Okinawan (ryu) 2566 Northern Amami-Oshima (ryn)
724	Korean	724	Korean	2575	Korean (kor)
739	Malay	739	Malay	2715	Malay (ms) 1080 Manado Malay (xmm) 1081 North Moluccan Malay (max) 2710 Malay (macrolanguage) (msa) 2715 Malay (individual language) (zlm) 2716 Brunei (kxd) 2730 Kubu (kvb) 2740 Minangkabau (min) 2750 Musi (mui) 2760 Negeri Sembilan Malay (zmi) 2780 Iban (iba)
732	Indonesian	732	Indonesian	2770	Indonesian (ind)
986	Other Asian languages	684	Chuvash		
		685	Karakalpak		
		686	Kazakh		
		687	Kirghiz		
		688	Karachay		
		690	Azerbaijani		
		692	Turkmen		
		693	Yakut		
		695	Tungus		
		698	Dravidian		
		699	Brahui		
		700	Gondi		
		705	Kurukh		
		706	Munda		
		707	Burushaski		
		709	Hakka		
		710	Kan, Hsianq		
		713	Fuchow		
		715	Wu		
		716	Tibetan		
		718	Karen		
		719	Kachin		
		727	Paleo-siberian		
		729	Muong		
		730	Buginese		
		731	Moluccan		
		733	Achinese		
			2850	Other languages of Asia	1643 Chuvash (chv) 1648 Kara-Kalpak (kaa) 1649 Kazakh (kaz) 1650 Kirghiz (kir) 1652 Uzbek (uzb) 1653 Northern Uzbek (uzn) 1654 Southern Uzbek (uzs) 1657 Karachay-Balkar (krc) 1658 Bashkir (bak) 1659 Tatar (tat) 1662 Crimean Tatar (crh) 1666 Uighur (uig) 1668 Azerbaijani (aze) 1669 North Azerbaijani (azi) 1670 South Azerbaijani (azb) 1680 Turkmen (tuk) 1681 Yakut (sah) 1685 Daur (dta) 1687 Bonan (peh)

1688	Tu (mig)
1695	Buriat (bua)
1696	Russia Buriat (bxr)
1698	Kalmyk (xal)
1699	Even (eve)
1702	Evenki (evn)
1703	Oroqen (orh)
1705	Nanai (gld)
1708	Manchu (mnc)
1711	Northwestern Kolami (kfb)
1716	Kui (India) (kxu)
1720	Northern Gondi (gno)
1760	Kodava (kfa)
1770	Tulu (tcy)
1775	Brahui (brh)
1780	Kurukh (kru)
1799	Burushaski (bsk)
1800	Korku (kfq)
1810	Asuri (asr)
1812	Munda (unx)
1820	Mahali (mix)
1822	Santali (sat)
1824	Turi (trd)
1835	Juang (jun)
1838	Kharia (khr)
1850	Sora (srb)
1855	Khasi (kha)
1860	Jah Hut (jah)
1870	Central Mnong (cmo)
1873	Bahnar (bdg)
1876	Koho (kpm)
1880	Laven (lbo)
1890	Upper Ta'oih (tth)
1895	Western Bru (brv)
1905	Mon (mnw)
1920	Khuen (khf)
1922	Khmu (kig)
1925	Lua' (prb)
1927	Mal (mlf)
1929	Phai (prt)
1931	Phong-Kniang (pnx)
1935	Eastern Lawa (lwl)
1940	Ruching Palaung (pce)
1950	Muong (mtg)
2060	Kulung (Nepal) (kle)
2065	Gamale Kham (kgj)
2075	Nepal Bhasa (new)
2080	Balti (bft)
2085	Dzongkha (dzo)
2090	Sikkimese (sip)
2092	Tshanqla (tsi)
2115	Gurung (gvr)
2120	Eastern Tamang (taj)
2135	Lepcha (lep)
2140	Adi (adi)
2145	Mising (mrq)
2150	Rawang (raw)
2175	Zaiwa (atb)
2185	Lahu (lhu)
2190	Lisu (lis)
2195	Sichuan Yi (iii)
2200	Phunoi (pho)
2210	Ao Naga (nio)
2215	Tangkhul Naga (India) (nmf)
2300	Manipuri (mni)
2310	Toto (txo)
2320	Bodo (India) (brx)
2325	Garó (grt)
2340	Kachin (kac)
2485	Bouyei (pcc)
2490	Zhuang (zha)
2500	Nung (Vietnam) (nut)
2505	Tày (tyz)
2515	Southern Dong (kmc)
2569	Ainu (Japan) (ain)
2577	Koryak (kpv)
2579	Itelmen (itl)
2582	Ket (ket)
2585	Gilyak (niv)
2588	Northern Yukaghir (vkq)
2590	Buginese (bug)
2600	Balinese (ban)
2630	Javanese (jav)
2640	Biatah Bidayuh (bth)
2650	Madurese (mad)
2660	Achinese (ace)
2670	Eastern Cham (cim)
2675	Western Cham (cja)
2680	Chru (cje)
2690	Jarai (ira)
2700	Rade (rad)
2790	Tetum (tet)
2800	Central Melanau (mel)
2810	Central Dusun (dtp)
2820	Coastal Kadazan (kzi)
2830	Kelabit (kzi)

					2840	Mainstream Kenyah (xkl)
					2850	Sundanese (sun)
742	Tagalog	742	Tagalog	2910	Filipino	Filipino (fil)
				2920	Tagalog	Tagalog (tgl)
743	Bisayan	743	Bisayan	2950	Cebuano	Cebuano (ceb)
744	Sebuano	744	Sebuano			
746	Ilocano	746	Ilocano	3150	Ilocano	Iloko (ilo)
				3190	Other Philippine languages	2870 Pampanga (pam)
						2880 Bolinao (smk)
						2890 Botolan Sambal (sbl)
						2900 Sambal (xsb)
						2940 Mandaya (mry)
						2960 Hiligaynon (hil)
						2970 Masbatenyo (msb)
						2980 Waray (Philippines) (war)
						2990 Aklanon (akl)
						3000 Cuyonon (cyo)
						3010 Kinaray-A (kri)
						3020 Butuanon (btw)
						3030 Tausug (tsq)
						3040 Suriqaonon (sgd)
						3050 Bikol (bik)
						3055 Central Bikol (bcl)
						3060 Gorontalo (gor)
						3070 Obo Manobo (obo)
						3080 Ibaloi (ibl)
						3090 Pangasinan (pag)
						3100 Batad Ifugao (ifb)
						3110 Tuwali Ifugao (ifk)
						3120 Bontok (bnc)
						3125 Central Bontok (lbc)
						3130 Kankanaey (kne)
						3160 Ibanag (ibq)
						3170 Itawit (itv)
						3180 Ivatan (ivv)
						3190 Tondano (tdn)
750	Micronesian	750	Micronesian			
752	Chamorro	752	Chamorro	3220	Chamorro	Chamorro (cha)
				3270	Marshallese	Marshallese (mah)
761	Trukese	761	Trukese	3350	Chuukese	Chuukese (chk)
767	Samoan	767	Samoan	3420	Samoan	Samoan (smo)
768	Tongan	768	Tongan	3500	Tongan	Tonga (Tonga Islands) (ton)
776	Hawaiian	776	Hawaiian	3570	Hawaiian	Hawaiian (haw)
988	Other Pacific Island languages	734	Balinese			
		735	Cham			
		736	Javanese			
		737	Madurese			
		738	Malaqasy			
		740	Minangkabau			
		741	Sundanese			
		745	Pangasinan			
		747	Bikol			
		748	Pampangan			
		749	Gorontalo			
		751	Carolinian			
		753	Gilbertese			
		754	Kusaiean			
		755	Marshallese			
		756	Mokilese			
		757	Mortlockese			
		758	Nauruan			
		759	Palau			
		760	Ponapean			
		762	Ulithian			
		763	Woleai-Ulithi			
		764	Yapese			
		765	Melanesian			
		766	Polynesian			
		769	Niuean			
		770	Tokelauan			
		771	Fijian			
		772	Marquesan			
		773	Rarotongan			
		774	Maori			
		775	Nukuoro			
				3600	Other Eastern Malayo-Polynesian languages	3250 Gilbertese (gil)
						3260 Kosraean (kos)
						3280 Carolinian (cal)
						3290 Mokilese (mkj)
						3300 Pingelapese (pif)
						3310 Mortlockese (mrl)
						3320 Nauru (nau)
						3330 Palauan (pau)
						3340 Pohnpeian (pon)
						3360 Puluwatese (puw)
						3370 Ulithian (uli)
						3380 Woleaian (woe)
						3390 Satawalese (stw)
						3400 Sonsorol (sov)
						3410 Yapese (vap)
						3430 Kapingamarangi (kpg)
						3440 Nukuoro (nkr)
						3470 Tuvalu (tvl)
						3480 Tokelau (tkl)
						3490 Wallisian (wls)
						3510 Niuean (niu)

					3520 North Marquesan (mrq)
					3530 Rapanui (rap)
					3540 Tahitian (tah)
					3550 Rarotongan (rar)
					3560 Maori (mri)
					3600 Eastern Malayo-Polynesian languages N.E.C.
					3610 Fijian (fii)
					3620 Rotuman (rtm)
					3630 Labo (mwi)
					3640 Hano (lml)
					3650 Mota (mtt)
					3660 Kwamera (tnk)
					3670 Lau (llu)
					3680 Amara (aie)
					3690 Arifama-Miniafia (aai)
					3700 Kilivila (kij)
					3710 Motu (meu)
					3720 Kuanua (ksd)
					3730 Niwer Mil (hrc)
					3740 Penchal (pek)
					3750 Kola (kvv)
					3755 Sula (szn)
					3760 Boano (bpw)
777	Arabic	777	Arabic	4500	Arabic (ara)
					4502 Algerian Arabic (arg)
					4504 Chadian Arabic (shu)
					4506 Cypriot Arabic (acy)
					4508 Egyptian Arabic (arz)
					4510 Gulf Arabic (afb)
					4512 Mesopotamian Arabic (acm)
					4514 Moroccan Arabic (ary)
					4516 North Levantine Arabic (apc)
					4518 Omani Arabic (acx)
					4520 Sanaani Arabic (ayn)
					4522 South Levantine Arabic (aip)
					4524 Standard Arabic (arb)
					4526 Sudanese Arabic (apd)
					4528 Tunisian Arabic (aeb)
					4530 Hassaniyya (mey)
					4535 Maltese (mlt)
					4540 Judeo-Arabic (irb)
778	Hebrew	778	Hebrew	4545	Hebrew (heb)
779	Syriac	779	Syriac	4560	Assyrian Neo-Aramaic (syr)
				4565	Assyrian Neo-Aramaic (aii)
				4565	Chaldean Neo-Aramaic (cld)
780	Amharic	780	Amharic	4590	Amharic (amh)
				4640	Tigrinya (tir)
783	Cushite	783	Cushite	4830	Oromo (orm)
					4832 Borana-Arsi-Guji Oromo (qax)
					4834 Eastern Oromo (hae)
					4836 West Central Oromo (qaz)
					4840 Somali (som)
				4880	Other Afro-Asiatic languages
					4570 Hulaulá (huy)
					4575 Turoyo (tru)
					4600 Harari (har)
					4610 Silt'e (stv)
					4620 Sebat Bet Gurage (sqw)
					4630 Tigre (tiq)
					4645 Coptic (cop)
					4660 Berber languages (ber)
					4661 Siwi (siz)
					4670 Central Atlas Tamazight (tzm)
					4675 Tachelhit (shi)
					4680 Kabyle (kab)
					4685 Nafusi (jbn)
					4690 Tachawit (shy)
					4695 Tarifit (rif)
					4700 Tamashek (tmh)
					4701 Tamasheg (tag)
					4710 Zenaga (zen)
					4720 Bura-Pabir (bwr)
					4725 Huba (hbb)
					4735 Bata (bta)
					4740 Gidar (qid)
					4750 Hausa (hau)
					4760 Tangale (tan)
					4765 Cakfem-Mushere (cky)
					4790 Biliin (byn)
					4800 Afar (aar)
					4805 Saho (ssy)
					4815 Hadiyya (hdy)
					4825 Sidamo (sid)
					4845 Garre (gex)
					4850 Maay (ymm)
					4855 Beja (bej)
					4860 Ari (aiw)
					4870 Kafa (kbr)
					4880 Wolaytta (wal)
					6717 Ethiopia N.E.C.
					6716 Eritrea N.E.C.
				4900	Nilo-Saharan languages
					4900 Mangbetu (mdj)
					4905 Lugbara (lqq)
					4910 Moru (mqd)

						4915	Ma'di (mhi)
						4919	Laka (Chad) (lap)
						4920	Nqambay (sba)
						4925	Sar (mwm)
						4930	Sara Kaba Naa (kwv)
						4935	Ama (Sudan) (nyi)
						4940	Didinga (did)
						4950	Kalenjin (kln)
						4953	Kipsigis (sqc)
						4955	Nandi (niq)
						4960	Pokoot (pko)
						4965	Dinka (din)
						4967	Southwestern Dinka (dik)
						4970	Nuer (nus)
						4975	Adhola (adh)
						4978	Acoli (ach)
						4980	Lang'o (Uganda) (lai)
						4985	Alur (alz)
						4990	Luo (Kenya and Tanzania) (luo)
						4995	Anuak (anu)
						5000	Luwo (lwo)
						5005	Shilluk (shk)
						5010	Mabaan (mfz)
						5015	Bari (bfa)
						5020	Kakwa (keo)
						5030	Otuho (lot)
						5035	Masai (mas)
						5040	Teso (teo)
						5045	Karamojong (kdj)
						5048	Turkana (tuv)
						5055	Nobiin (fia)
						5060	Kanuri (kau)
						5061	Central Kanuri (knc)
						5063	Manga Kanuri (kby)
						5070	Tedaga (tuq)
						5075	Zaghawa (zag)
						5080	Amdang (ami)
						5085	Fur (fvr)
						5090	Katcha-Kadugli-Miri (xtc)
						5095	Kunama (kun)
						5100	Maba (Chad) (mde)
						5105	Masalit (mls)
						5110	Kibet (kie)
						5115	Koyraboro Senni Songhay (ses)
						5117	Zarma (dje)
791	Swahili	791	Swahili	5150	Swahili	5150	Swahili (macrolanguage) (swa)
						5151	Swahili (individual language) (swh)
						5152	Congo Swahili (swc)
792	Bantu	792	Bantu	5345	Ganda	5345	Ganda (lug)
						5350	Soqa (xog)
						5355	Gwere (gwr)
				5525	Shona	5525	Shona (sna)
						5530	Ndau (ndc)
						5535	Kalanga (kck)
				5645	Other Bantu languages	5140	Bantu languages (bnt)
						5145	Samba Daka (ccg)
						5160	Ekajuk (eka)
						5170	Nkem-Nkum (isi)
						5180	Ejagham (etu)
						5185	Bankal (jir)
						5190	Mama (mma)
						5195	Kenyang (ken)
						5200	Mbe (mfo)
						5205	Venda (ven)
						5210	Bembe (bmb)
						5220	Nyanga (nyj)
						5225	Mochi (old)
						5230	Kikuyu (kik)
						5235	Meru (mer)
						5240	Embu (ebu)
						5245	Kamba (Kenya) (kam)
						5248	Giryama (nyf)
						5250	Sagalla (tga)
						5255	Sukuma (suk)
						5260	Nyamwezi (nym)
						5265	Hehe (heh)
						5270	Pangwa (pbr)
						5275	Zizula (ziw)
						5278	Mushungulu (xma)
						5280	Kongo (kon)
						5281	Koongo (kng)
						5283	Laari (ldi)
						5285	Kimbundu (kmb)
						5290	Haya (hay)
						5300	Luvia (luy)
						5301	Bukusu (bxk)
						5302	East Nyala (nle)
						5303	Idakho-Isukha-Tiriki (ida)
						5307	Logooli (raq)
						5310	Saamia (lsm)
						5315	Gusii (guz)
						5320	Masaaba (myx)
						5330	Nvoro (nvo)
						5335	Nyankole (nyn)
						5340	Chiga (cgg)
						5360	Tooro (tti)

5365	Kinyarwanda (kin)
5370	Rundi (run)
5375	Fuliru (flr)
5380	Nyemba (nba)
5385	Luvale (lue)
5390	Lozi (loz)
5400	Kaonde (kqn)
5405	Luba-Lulua (lua)
5407	Luba-Katanga (lub)
5410	Lunda (lun)
5415	Bemba (Zambia) (bem)
5420	Mambwe-Lungu (mgr)
5425	Lamba (lam)
5430	Ila (ilb)
5435	Lenje (leh)
5437	Tonga (Zambia) (toi)
5445	Nyanja (nya)
5455	Tonga (Nyasa) (tog)
5460	Nsenga (nse)
5462	Sena (seh)
5465	Tumbuka (tum)
5470	Makhuwa (vmw)
5475	Matumbi (mgw)
5480	Yao (yao)
5485	Makonde (kde)
5490	Herero (her)
5495	Umbundu (umb)
5500	Kuanyama (kua)
5505	Xhosa (xho)
5510	Zulu (zul)
5515	Swati (ssw)
5520	North Ndebele (nde)
5540	Tswana (tsn)
5545	Northern Sotho (nso)
5546	Southern Sotho (sot)
5550	Tsonga (tso)
5555	Bafia (ksf)
5560	Basa (Cameroon) (bas)
5565	Nvokon (nvo)
5570	Nomaande (lem)
5575	Bube (bvb)
5580	Benga (bnq)
5585	Mokpwe (bri)
5590	Duala (dua)
5595	Ewondo (ewo)
5600	Eton (Cameroon) (eto)
5605	Bulu (Cameroon) (bum)
5610	Fang (Equatorial Guinea) (fan)
5615	Oroko (bdu)
5620	Mbo (zms)
5625	Akoose (bss)
5630	Bakaka (baq)
5633	Punu (puu)
5635	Ibali Teke (tek)
5640	Bangala (bxg)
5645	Lingala (lin)
5650	Mabaale (mmz)
5655	Bangi (bni)
5660	Ntomba (nto)
5665	Mongo (lol)
5670	Ligenza (lgz)
5675	Bwa (bww)
5677	Tetela (ttl)
5680	Abon (abo)
5683	Batu (btu)
5685	Tiv (tiv)
5695	Bamileke languages (bai)
5696	Fe'fe' (fmp)
5700	Mengaka (xmg)
5705	Ngiemboon (nnh)
5710	Nqwe (nwe)
5715	Yemba (ybb)
5720	Bafut (bfd)
5721	Bambili-Bambui (baw)
5725	Mendankwe-Nkwen (mfd)
5730	Ngemba (nge)
5735	Pinyin (pny)
5740	Limbang (lmp)
5745	Bamali (bbq)
5750	Bamun (bax)
5755	Medumba (bvv)
5760	Mungaka (mhk)
5765	Meta' (mgo)
5770	Ngie (ngi)
5775	Nqwo (ngn)
5780	Babanki (bbk)
5785	Kom (Cameroon) (bkm)
5790	Oku (oku)
5795	Lamns'o' (lms)
5800	Kenswei Nsei (ndb)
5805	Vengo (bav)
5810	Aghem (agg)
5815	Isu (Menchum Division) (isu)
5820	Njebi (nzb)
6701	Angola N.E.C.
6718	Gabon N.E.C.

						6727	Malawi N.E.C.
						6733	Mozambique N.E.C.
						6751	Zambia N.E.C.
793	Mande	793	Mande	5845	Manding languages	5845	Mandingo (man)
						5840	Kuranko (knk)
						5850	Mandinka (mnk)
						5853	Eastern Maninkakan (emk)
						5855	Western Maninkakan (mlq)
						5860	Bambara (bam)
						5865	Dvula (dvu)
						5870	Mahou (mxx)
						5880	Kono (Sierra Leone) (kno)
						5885	Vai (vai)
				5900	Other Mande languages	5830	Soninke (snk)
						5835	Southern Bobo Madaré (bwq)
						5890	Susu (sus)
						5895	Yalunka (yal)
						5900	Kpelle (kpe)
						5901	Liberia Kpelle (xpe)
						5905	Bandi (bza)
						5910	Mende (Sierra Leone) (men)
						5915	Loma (Liberia) (lom)
						5917	Toma (tod)
						5920	Bissa (bib)
						5925	Dan (dni)
						5930	Mano (mev)
						5935	Gagu (ggu)
794	Fulani	794	Fulani	5940	Fulah	5940	Fulah (ful)
						5942	Adamawa Fulfulde (fub)
						5945	Pulaar (fuc)
						5946	Pular (fuf)
				5950	Wolof	5950	Wolof (wol)
796	Kru, Ibo, Yoruba	796	Kru, Ibo, Yoruba	6120	Akan (incl. Twi)	6120	Akan (aka)
						6125	Fanti (fat)
						6130	Twi (twi)
				6205	Ga	6200	Adangme (ada)
						6205	Ga (gaa)
				6230	Gbe languages	6220	Aja (Benin) (ajg)
						6225	Gun (quw)
						6230	Ewe (ewe)
						6235	Fon (fon)
						6240	Gen (gei)
						6245	Waci Gbe (wci)
				6290	Yoruba	6290	Yoruba (yor)
						6293	Ifè (ife)
						6295	Isekiri (its)
				6300	Edoid languages	6300	Bini (bin)
						6305	Esan (ish)
						6310	Emai-Iuleha-Ora (ema)
						6315	Epie (epi)
						6320	Ivbie North-Okpela-Arhe (atg)
						6325	Uneme (une)
						6330	Yekhee (ets)
						6335	Iyavu (iya)
						6340	Isoko (iso)
						6345	Okpe (Southwestern Edo) (oke)
						6350	Urhobo (urh)
						6355	Uyvie (evh)
				6370	Igbo	6370	Igbo (ibo)
						6375	Ika (ikk)
						6378	Ikwere (ikw)
						6381	Ogbah (ogc)
						6384	Ukwuani-Aboh-Ndoni (ukw)
				6500	Other Niger-Congo languages	5960	Serer (srr)
						5965	Wamey (cou)
						5970	Balanta-Kentohe (ble)
						5975	Jola-Fonyi (dvo)
						5980	Mandjak (mfv)
						5985	Gola (qol)
						5990	Krim (krm)
						5995	Southern Kisi (kss)
						6000	Timne (tem)
						6005	West-Central Limba (lia)
						6015	Baatonum (bba)
						6020	Nafaanra (nfr)
						6025	Cebaara Sênoufo (sef)
						6026	Diimini Sênoufo (dvi)
						6035	Ntcham (bud)
						6040	Gourmanchéma (gux)
						6045	Buli (Ghana) (bwu)
						6050	Dagbani (daq)
						6055	Kusaal (kus)
						6060	Mampruli (maw)
						6065	Mossi (mos)
						6070	Farefare (gur)
						6075	Southern Daqaare (dqa)
						6080	Wali (Ghana) (wlx)
						6085	Kaansa (qna)
						6090	Kasem (xsm)
						6095	Kabiyè (kbp)
						6098	Tem (kdh)
						6100	Paasaal (sig)
						6105	Vaqla (vaq)
						6110	Wara (wbf)

					6135 Anyin (any)
					6140 Baoulé (bci)
					6145 Sehwi (sfw)
					6150 Nzima (nzi)
					6155 Gonja (gin)
					6160 Krache (kve)
					6165 Nawuri (naw)
					6170 Awutu (afu)
					6175 Larteh (lar)
					6180 Logba (lqq)
					6185 Sekpele (lip)
					6190 Adioukrou (adj)
					6195 Attié (ati)
					6210 Adangbe (adq)
					6215 Ikposo (kpo)
					6250 Southeast Ijo (ijs)
					6255 Ibani (iby)
					6260 Kalabari (iin)
					6265 Kirike (okr)
					6270 Izon (ijc)
					6275 Biseni (ije)
					6280 Arigidi (aqg)
					6285 Igala (igl)
					6357 Idoma (idu)
					6360 Igede (ige)
					6363 Yala (yba)
					6365 Ekpeye (ekp)
					6390 Jibu (jib)
					6395 Kutep (kub)
					6400 Bauchi (bsf)
					6405 Bina (Nigeria) (bvi)
					6408 Tumi (kku)
					6410 Ebira (iqb)
					6415 Gade (qed)
					6420 Gbagyi (qbr)
					6425 Nupe-Nupe-Tako (nup)
					6435 Adara (kad)
					6440 Berom (born)
					6445 Eggon (ego)
					6450 Jiu (kaj)
					6455 Ukaan (kcf)
					6460 Wané (hwa)
					6465 Kuwaa (blh)
					6470 Bassa (bsq)
					6475 Grebo (grb)
					6476 Northern Grebo (gbo)
					6480 Klao (klu)
					6485 Eastern Krahn (kqo)
					6486 Western Krahn (krw)
					6490 Sapo (krn)
					6500 Efik (efi)
					6505 Anaang (anw)
					6510 Ibibio (ibb)
					6515 Ekit (eke)
					6520 Ibino (ibn)
					6525 Obolo (ann)
					6530 Oro (orx)
					6535 Eleme (elm)
					6540 Gokana (qkn)
					6545 Khana (ogo)
					6550 Tee (tka)
					6555 Abua (abn)
					6560 Aqwaqwune (yay)
					6563 Kohumono (bcs)
					6565 Legbo (aqb)
					6570 Lokaa (yaz)
					6575 Olulumo-Ikom (iko)
					6580 Bete-Bendi (btt)
					6585 Bokyi (bky)
					6590 Baka (Cameroon) (bkc)
					6595 Mayogo (mdm)
					6600 Gbaya (gba)
					6601 Northwest Gbaya (qya)
					6605 Zande (zne)
					6610 Day (dai)
					6615 Mbum (mdd)
					6620 Samba Leko (ndi)
					6625 Teme (tdo)
					6630 Kaan (ldl)
					6635 Tula (tul)
					6640 Jamsay Dogon (dim)
					6650 Moro (mor)
					6702 Benin N.E.C.
					6707 Cameroon N.E.C.
					6719 Gambia N.E.C.
					6720 Ghana N.E.C.
					6721 Guinea N.E.C.
					6725 Liberia N.E.C.
					6738 Senegal N.E.C.
					6740 Sierra Leone N.E.C.
799	African	799	African N.E.C.		
989	Other Specified African languages	781	Berber		
		782	Chadic		
		784	Sudanic		
		785	Nilotic		
		786	Nilo-Hamitic		

		787 Nubian			
		788 Saharan			
		789 Nilo-Saharan			
		790 Khoisan			
		795 Gur			
		797 Efik			
		798 Mbum (and Related)			
				6795 Other languages of Africa	5120 Korana (kqz) 5125 //Gana (gnk) 5127 Nama (Namibia) (ylo) 5130 Hadza (hts) 5135 Sandawe (sad) 6723 Kenya N.E.C. 6728 Mali N.E.C. 6734 Namibia N.E.C. 6735 Niger N.E.C. 6736 Nigeria N.E.C. 6745 Sudan N.E.C. 6747 Tanzania N.E.C. 6750 Uganda N.E.C. 6795 Africa N.E.C.
990	Aleut-Eskimo languages	800 Aleut		6800 Aleut languages	6800 Aleut (ale) 6801 Inupiaq (ipk) 6802 North Alaskan Inupiatun (esi) 6803 Northwest Alaska Inupiatun (esk) 6804 Eastern Canadian Inuktitut (ike) 6806 Greenlandic (kal) 6808 Pacific Gulf Yupik (ems) 6809 Central Siberian Yupik (ess) 6810 Central Yupik (esu)
		801 Pacific Gulf Yupik			
		802 Eskimo			
		803 Inupik			
		804 St. Lawrence Island Yupik			
		805 Yupik			
806	Other Algonquian languages	820 Ottawa			
		806 Algonquin			
		807 Arapaho			
		808 Atsina			
		809 Blackfoot			
		810 Cheyenne			
		811 Cree			
		816 French Cree			
		812 Delaware			
		813 Fox			
		814 Kickapoo			
		815 Menomini			
		817 Miami			
		818 Micmac			
		824 Potawatomi			
		821 Passamaquoddy			
		822 Penobscot			
		823 Abnaki			
		825 Shawnee			
		826 Wiyot			
		827 Yurok			
819	Ojibwa	819 Ojibwa		6839 Ojibwa	6839 Ojibwa (oji) 6840 Chippewa (ciw) 6841 Ottawa (otw) 6842 Severn Ojibwa (ois)
862	Apache	862 Apache		6930 Apache languages	6927 Jicarilla Apache (api) 6928 Lipan Apache (apl) 6929 Mescalero-Chiricahua Apache (apm) 6930 Western Apache (apw) 6931 Kiowa Apache (apk)
864	Navajo	864 Navajo		6933 Navajo	6933 Navajo (nav)
				6936 Kiowa-Tanoan languages	6936 Kiowa (kio) 6937 Northern Tiwa (twf) 6938 Southern Tiwa (tix) 6939 Tewa (USA) (tew) 6940 Jemez (tow)
907	Dakota	907 Dakota		7019 Dakota languages	7017 Assiniboine (asb) 7018 Dakota (dak) 7019 Lakota (lkt) 7020 Stoney (sto) 7028 Alabama (akz) 7029 Koasati (cku) 7030 Mikasuki (mik) 7031 Creek (mus) 7032 Choctaw (cho) 7033 Chickasaw (cic)
				7032 Muskogean languages	7039 Eastern Keres (kee) 7040 Western Keres (kiq)
924	Keres	924 Keres		7039 Keres	7050 Cherokee (chr)
933	Cherokee	933 Cherokee		7050 Cherokee	7059 Zuni (zun)
964	Zuni	964 Zuni		7059 Zuni	7060 Comanche (com) 7062 Shoshoni (shh) 7064 Panamint (par) 7066 Mono (USA) (mnr) 7068 Northern Paiute (pao) 7069 Ute-Southern Paiute (ute) 7070 Kawaiisu (xaw) 7072 Cahuilla (chl) 7074 Cupeño (cup) 7076 Luiseno (lui) 7078 Serrano (ser) 7080 Hopi (hop) 7082 Tübatulabal (tub) 7084 Tohono O'odham (ood) 7100 Southeastern Tepehuan (stp)
				7060 Uto-Aztecan languages	

			7103 Eastern Huasteca Nahuatl (nhe)
			7105 Michoacán Nahuatl (ncl)
			7109 El Nayar Cora (crn)
			7111 Huichol (hch)
			7115 Central Tarahumara (tar)
			7120 Mayo (mfy)
			7123 Yaqui (yaq)
993	Other Specified North American Indian languages	828 Kutenai	
		829 Makah	
		830 Kwakiutl	
		831 Nootka	
		833 Lower Chehalis	
		834 Upper Chehalis	
		835 Clallam	
		836 Coeur d'Alene	
		837 Columbia	
		838 Cowlitz	
		839 Salish	
		840 Nootsack	
		841 Okanogan	
		842 Puget Sound Salish	
		843 Quinault	
		844 Tillamook	
		845 Twana	
		846 Haida	
		847 Athapascan	
		848 Ahtena	
		849 Han	
		850 Ingalit	
		851 Koyukon	
		852 Kuchin	
		853 Upper Kuskokwim	
		854 Tanaina	
		855 Tanana	
		856 Tanacross	
		857 Upper Tanana	
		858 Tutchone	
		859 Chasta Costa	
		860 Hupa	
		861 Other Athapascan-Eyak	
		863 Kiowa	
		865 Eyak	
		866 Tlingit	
		867 Mountain Maidu	
		868 Northwest Maidu	
		869 Southern Maidu	
		870 Coast Miwok	
		871 Plains Miwok	
		872 Sierra Miwok	
		873 Nomlaki	
		874 Patwin	
		875 Wintun	
		876 Foothill North Yokuts	
		877 Tachi	
		878 Santiam	
		879 Siuslaw	
		880 Klamath	
		881 Nez Perce	
		882 Sahaptian	
		883 Upper Chinook	
		884 Tsimshian	
		885 Achumawi	
		886 Atsugewi	
		887 Karok	
		888 Pomo	
		889 Shastan	
		890 Washo	
		891 Up River Yuman	
		892 Cocomaricopa	
		893 Mohave	
		894 Yuma	
		895 Diegueno	
		896 Delta River Yuman	
		897 Upland Yuman	
		898 Havasupai	
		899 Walapai	
		900 Yavapai	
		901 Chumash	
		902 Tonkawa	
		903 Yuchi	
		904 Crow	
		905 Hidatsa	
		906 Mandan	
		908 Chiwere	
		909 Winnebago	
		910 Kansa	
		911 Omaha	
		912 Osage	
		913 Ponca	
		914 Quapaw	
		915 Alabama	
		916 Choctaw	
		917 Mikasuki	
		918 Hichita	
		919 Koasati	

	920	Muskogee	
	921	Chetemacha	
	922	Yuki	
	923	Wappo	
	925	Iroquois	
	926	Mohawk	
	927	Oneida	
	928	Onondaga	
	929	Cayuga	
	930	Seneca	
	931	Tuscarora	
	932	Wyandot	
	934	Arikara	
	935	Caddo	
	936	Pawnee	
	937	Wichita	
	938	Comanche	
	939	Mono	
	940	Paiute	
	941	Northern Paiute	
	942	Southern Paiute	
	943	Chemehuevi	
	944	Kawaiisu	
	945	Ute	
	946	Shoshoni	
	947	Panamint	
	948	Hopi	
	949	Cahuilla	
	950	Cupeno	
	951	Luiseno	
	952	Serrano	
	953	Tubatulabal	
	954	Pima	
	955	Yaqui	
	959	Picuris	
	960	Tiwa	
	961	Sandia	
	962	Tewa	
	963	Towa	
	965	Chinook Jargon	
	969	Tarascan	
	978	Chiricahua	
	979	San Carlos	
	980	Kiowa-Apache	
	981	Kalispel	
	982	Spokane	
	966	American Indian	
	7124	Other Native North American languages	6815 Algonquin (alg)
			6817 Mohegan-Pequot (xpg)
			6818 Narragansett (xnt)
			6819 Powhatan (pim)
			6820 Wampanoag (wam)
			6821 Lumbee (lmz)
			6822 Arapaho (arp)
			6823 Gros Ventre (ats)
			6824 Siksika (bla)
			6825 Cheyenne (chy)
			6827 Cree (cre)
			6828 Michif (crg)
			6829 Delaware (del)
			6830 Munsee (umu)
			6831 Unami (unm)
			6832 Nanticoke (nnt)
			6833 Meskwaki (sac)
			6834 Kickapoo (kic)
			6835 Menominee (mez)
			6837 Miami (mia)
			6838 Mi'kmaq (mic)
			6843 Potawatomi (pot)
			6844 Malecite-Passamaquoddy (pqm)
			6845 Eastern Abnaki (aaq)
			6846 Western Abnaki (abe)
			6848 Shawnee (sjw)
			6849 Wiyot (wiy)
			6850 Yurok (yur)
			6856 Kutenai (kut)
			6857 Quileute (qui)
			6859 Kwakiutl (kwk)
			6861 Ditidaht (dtd)
			6862 Nuu-chah-nulth (nuk)
			6863 Makah (myh)
			6866 Quinalt (qun)
			6867 Lower Chehalis (cea)
			6868 Upper Chehalis (cjh)
			6869 Cowlitz (cow)
			6872 Clallam (clm)
			6873 Coeur d'Alene (crd)
			6874 Columbia-Wenatchi (col)
			6875 Okanagan (oka)
			6876 Kalispel-Pend d'Oreille (fla)
			6877 Spokane (spo)
			6878 Shuswap (shs)
			6879 Halkomelem (hur)
			6880 Straits Salish (str)
			6881 Nooksack (nok)
			6883 Lushootseed (lut)

6884	Skaqit (ska)
6885	Snohomish (sno)
6886	Southern Puget Sound Salish (slh)
6888	Tillamook (til)
6890	Twana (twa)
6892	Haida (hai)
6893	Northern Haida (hdn)
6894	Southern Haida (hax)
6895	Athapascan languages (ath)
6897	Evak (eva)
6898	Tlingit (tli)
6900	Slave (Athapascan) (den)
6902	Chilcotin (clc)
6903	Ahtena (aht)
6904	Tanaina (tfn)
6905	Degexit'an (ing)
6906	Koyukon (koy)
6907	Holikachuk (hoi)
6908	Upper Kuskokwim (kuu)
6909	Lower Tanana (taa)
6910	Tanacross (tcb)
6911	Upper Tanana (tau)
6912	Northern Tutchone (tm)
6913	Southern Tutchone (tce)
6914	Gwich'in (gwi)
6915	Han (haa)
6916	Chetco (ctc)
6917	Coquille (cog)
6918	Galice (gce)
6919	Tolowa (tol)
6920	Tututni (tuu)
6921	Hupa (hup)
6922	Kato (ktw)
6923	Mattole (mvb)
6924	Wailaki (wlk)
6942	Northeast Maidu (nmu)
6943	Northwest Maidu (mjd)
6944	Nisenan (nsz)
6946	Northern Ohlone (cst)
6948	Coast Miwok (csi)
6949	Plains Miwok (pmw)
6950	Central Sierra Miwok (csm)
6951	Northern Sierra Miwok (nsq)
6952	Southern Sierra Miwok (skd)
6955	Nomlaki (nol)
6956	Patwin (pwi)
6957	Wintu (wnw)
6960	Yokuts (yok)
6961	Kalapuya (kyl)
6962	Siuslaw (sis)
6963	Klamath-Modoc (kla)
6965	Nez Perce (nez)
6966	Tenino (tgn)
6967	Umatilla (uma)
6968	Walla Walla (waa)
6969	Yakama (yak)
6971	Chinook (chh)
6972	Wasco-Wishram (wac)
6974	Gitsan (git)
6975	Nisga'a (ncg)
6976	Tsimshian (tsi)
6978	Achumawi (acv)
6979	Atsugewi (atw)
6980	Karok (kyh)
6981	Central Pomo (poo)
6982	Kashaya (kju)
6983	Southern Pomo (peq)
6984	Northern Pomo (pei)
6985	Eastern Pomo (peb)
6986	Northeastern Pomo (pef)
6987	Southeastern Pomo (pom)
6988	Shasta (sht)
6989	Washo (was)
6990	Maricopa (mrc)
6991	Mohave (mov)
6992	Quechan (yum)
6994	Kumiai (dih)
6995	Cocopa (coc)
6997	Havasupai-Walapai-Yavapai (yuf)
7000	Barbareño (boi)
7001	Ineseño (inz)
7002	Purisimeño (puy)
7003	Ventureño (veo)
7004	Cruzeño (crz)
7005	Obispeño (obi)
7007	Chimariko (cid)
7008	Esselen (esq)
7009	Salinan (sln)
7010	Tonkawa (tgw)
7011	Yuchi (yuc)
7013	Crow (cro)
7014	Hidatsa (hid)
7015	Mandan (mhq)
7016	Catawba (chc)
7022	Iowa-Oto (iow)
7023	Ho-Chunk (win)

				7024	Kansa (ksk)
				7025	Omaha-Ponca (oma)
				7026	Osage (osa)
				7027	Quapaw (qua)
				7034	Chitimacha (ctm)
				7036	Yuki (yuk)
				7037	Wappo (wao)
				7042	Iroquoian languages (iro)
				7043	Mohawk (moh)
				7044	Oneida (one)
				7045	Onondaga (ono)
				7046	Cayuga (cay)
				7047	Seneca (see)
				7048	Wyandot (wya)
				7049	Tuscarora (tus)
				7052	Caddo (cad)
				7053	Arikara (ari)
				7054	Pawnee (paw)
				7055	Wichita (wic)
				8050	American Indian N.E.C.
992	South/Central America Indian languages	956	Aztecán		
		957	Sonoran, N.E.C.		
		967	Misumalpan		
		968	Mayan languages		
		970	Mapuche		
		971	Oto - Manguen		
		972	Quechua		
		973	Aymara		
		974	Arawakian		
		975	Chibchan		
		976	Tupi-Guarani		
				7125	Matagalpa (mtn)
				7126	Mayangna (yan)
				7127	Miskito (mig)
				7130	Huastec (hus)
				7135	Chol (ctu)
				7136	Chortí (caa)
				7137	Tzeltal (tzh)
				7138	Tzotzil (tzo)
				7143	Aquacateco (aqu)
				7144	Ixil (ixl)
				7150	Mam (mam)
				7152	Tektiteko (ttc)
				7155	Kaqchikel (cak)
				7156	Tz'utujil
				7160	K'iche' (quc)
				7161	Kekchi (kek)
				7162	Poqomam (poc)
				7163	Poqomchi' (poh)
				7166	Chuj (cac)
				7170	Popti' (jac)
				7171	Q'anjob'al (kjb)
				7172	Western Kanjobal (kni)
				7175	Itzá (itz)
				7177	Mopán Maya (mop)
				7179	Lacandon (lac)
				7180	Yucateco (yua)
				7182	Purepecha (tsz)
				7185	Huilliche (huh)
				7186	Mapudungun (arn)
				7188	Highland Totonac (tos)
				7200	Guerrero Amuzgo (amu)
				7210	Copala Triqui (trc)
				7220	Alacatlazala Mixtec (mim)
				7230	Mixtepec Mixtec (mix)
				7240	Silacayoapan Mixtec (mks)
				7250	Mixtec N.E.C.
				7260	Tepeuxila Cuicatec (cux)
				7270	Huaulla Mazatec (mau)
				7280	Ixcatec (ixc)
				7290	San Juan Atzingo Popoloca (poe)
				7300	Zapotec (zap)
				7310	Western Tlaxolula Valley Zapotec (twb)
				7320	Zooqocho Zapotec (zpq)
				7330	Central Mazahua (maz)
				7340	Mezquital Otomi (ote)
				7350	Ojitán Chinantec (chi)
				7400	Quechua (que)
				7500	Aymara (aym)
				7501	Central Aymara (ayr)
				7505	Galibi Carib (car)
				7515	Wavana (wav)
				7525	Waiwai (waw)
				7535	Guahibo (guh)
				7540	Asháninka (cni)
				7550	Arawak (arw)
				7555	Garifuna (cab)
				7560	Taino (tnq)
				7565	Wayuu (quc)
				7575	Wapishana (wap)
				7585	Yucuna (ycn)
				7600	Northern Emberá (emp)
				7635	Boruca (brn)
				7640	Ngäbere (qym)
				7645	Arhuaco (arh)
				7648	Coqui (koq)



4270	Tanngu (tgu)
4320	Sinagen (siu)
4440	Eastern Arrernte (aer)
4450	Yan-nhangu (jau)
8000	Basque (eus)
8010	Esperanto (epo)
8020	Shelta (sth)
8030	Latin (lat)
8055	Caribbean N.E.C.
8060	Eastern Europe N.E.C.
8061	Europe N.E.C.
8089	Specified not listed
9999	Unspecified

N.E.C. stands for not elsewhere classified.