# Note for Language Spoken at Home from the 2016 American Community Survey

The American Community Survey (ACS) collects data on whether or not people five years old or older speak a language other than English at home. If a respondent indicates speaking a language other than English, the ACS asks what language the person speaks and how well the person speaks English. Figure 1 shows how the question appears on the ACS questionnaire.

**Figure 1: Presentation of language questions on 2016 ACS paper questionnaire**



## 2016 ACS Language coding update

Beginning with 2016 1-year and 2012-2016 5-year American Community Survey (ACS) data, coding of specific languages reflects languages in accordance with the International Organization for Standardization's ISO-639-3 standard.[1] To the extent possible, languages in the code list are those considered individual spoken languages. A few languages reflect common write-ins that cannot be classified as individual languages (i.e., country names or language families).

## Changes to tables B16001, C16001, and B16002

The rows presented in American Community Survey tables B16001, C16001, and B16002 have been updated to reflect changes in the number of people speaking different languages. Languages and language categories that have grown have been added to these tables, while some that have decreased are no longer displayed individually but instead included in an aggregated form. Some categories are the same except for an updated label. Guidance for comparing estimates from the 2015 and 2016 language tables is included below.

---

[1] Please see http://www-01.sil.org/iso639-3/scope.asp for more information.

In tabulations, languages have sometimes been combined to create a category that reflects a major language family or geographical area instead of an individual spoken language. Our chief reason for aggregating languages together is out of concern for the privacy of respondents, especially in standardized data products that are designed to be available for small towns and rural areas where there may be only a few people speaking a given language. We also want to avoid presenting data with a small sample size in order to ensure data quality. The goal of the 2016 table redesign was to publish language data that are as useful as possible, working within these constraints.

Example languages have been added to the labels of residual "other" categories. For example, "Other Slavic languages" is now "Ukrainian and other Slavic languages." These example languages are among the largest within their respective "other" categories, but not large enough to provide data for in our standardized tables. Data for these example languages, and other language categories with at least 10,000 speakers nationwide, are available in the 2016 1-year and 2012-2016 5-year Public Use Microdata Sample (PUMS) data files.[2]

## B16001 and C16001 redesign

In American FactFinder, Table B16001 provides the most detail for individual languages, tabulated by English-speaking ability. 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.

With the exception of Navajo and Other Native North American languages, each language and language category shown in B16001 had 200,000 speakers or more nationwide in 2016. Each language and "other" category in C16001 had one million speakers or more nationwide in 2016.

Categories in 2016:

**Four and Forty-Two Group Classifications of Languages Spoken at Home with Examples**

| Four Group Classification | Forty-Two Group Classification | Examples |
|---|---|---|
| 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 | Navajo | Navajo |
| | Other Native languages of North America | Apache languages, Cherokee, Lakota, Tohono O'odham, Yupik 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 |
| | Other and unspecified languages | Hungarian, Jamaican Creole English, Unspecified |

## B16001 comparability

Languages and groups that can be compared between 2015 and 2016:

| 2015 | 2016 |
|---|---|
| Spanish and Spanish Creoles | Spanish |
| Italian | Italian |
| Portuguese and Portuguese Creoles | Portuguese |
| German | German |
| Yiddish | Use PUMS data[2] |
| Add together "Yiddish" and "Other West Germanic languages" rows | Yiddish, Pennsylvania Dutch or other West Germanic languages |
| Scandinavian languages | Use PUMS data[2] |
| 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 |
| Use PUMS data[2] | Punjabi |
| Use PUMS data[2] | Bengali |
| Other Indic languages | Add together "Punjabi," "Bengali," and "Nepali, Marathi, and other Indic languages" rows |
| Use PUMS data[2] | Telugu |
| Use PUMS data[2] | Tamil |
| Use PUMS data[2] | 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 data[2] |
| Laotian | Use PUMS data[2] |
| Tagalog | Tagalog (incl. Filipino) |
| Other Pacific Island Languages | Ilocano, Samoan, Hawaiian, or other Austronesian languages |
| Arabic | Arabic |
| Hebrew | Hebrew |
| Navajo | Navajo |
| Other Native languages of North America | Other Native languages of North America |

Compare with caution between 2015 and 2016:

- French (incl. Cajun): In 2015, the code for Patois was grouped with French in table B16001. Beginning in 2016, "Patois" is usually coded as Jamaican Creole English unless a more appropriate code is indicated. Data for Jamaican Creole English are available in the PUMS files.
- French Creole: In 2015, all French-based creole languages were coded as "French Creole." In the United States, the majority of French Creole speakers speak Haitian. In 2016, Haitian and some French-based creoles of the Caribbean that are generally mutually intelligible are categorized as Haitian. These include Guadeloupe Creole French and Saint Lucian Creole French. French-based creole languages that are not intelligible with Haitian are grouped in "Other and unspecified languages."

Comparable between 2015 and 2016, but requires the use of PUMS data:

- Yiddish has been combined with Other West Germanic languages in the B16001 table. 2016 data for Yiddish are available in the PUMS files.
- Bengali, Punjabi, Telugu, and Tamil are all new to the B16001 table. Previous years' data for these languages are available in the PUMS files.
- Scandinavian languages has been combined with Other Indo-European languages in the 2016 B16001 table. 2016 data for individual Scandinavian languages are available in the PUMS files.
- Thai and Laotian were previously on their own rows but are now combined within "Thai, Lao, or other Tai-Kadai languages." 2016 data for Thai and Lao are individually available in the PUMS.
- Malayalam, Kannada, or other Dravidian languages is a new category for the 2016 B16001 table. All Dravidian languages were included under "Other Asian languages" in the 2015 B16001 table. Individual-language data for Malayalam or Kannada are available in the PUMS files.

Not comparable between 2015 and 2016:

- Other Indo-European languages
- Other languages of Asia
- Amharic, Somali, or other Afro-Asiatic languages
- Yoruba, Twi, Igbo, or other languages of Western Africa
- Swahili or other languages of Central, Eastern, and Southern Africa
- Other and unspecified languages

## *C16001 comparability*

Many rows in C16001 have received label changes. However, languages in C16001 are comparable with the previous year's C16001, except:
- Arabic is new to C16001, and was previously only available in B16001 and PUMS datasets.
- The "Other and Unspecified" row no longer contains Arabic.
- In 2016, C16001 also displays English proficiency categories. Each language category now contains three rows – total speakers, population who speak English "very well," and population who speak English "less than very well."

## *Additional geographical restrictions to B16001 for 5-year estimates*

Geographical restrictions have been applied to Table B16001 - LANGUAGE SPOKEN AT HOME BY ABILITY TO SPEAK ENGLISH FOR THE POPULATION 5 YEARS AND OVER for the 5-year data estimates. These restrictions are in place to protect data privacy for the speakers of smaller languages. Geographic areas published for the 5-year B16001 table include: Nation (010), States (040), Metropolitan Statistical Area - Metropolitan Divisions (314), Combined Statistical Areas (330), Congressional Districts (500), and Public Use Microdata Sample Areas (PUMAs) (795). For more information on these geographical delineations, see the Metropolitan Statistical Area Reference Files. County and tract-level data are no longer available for table B16001. For specific language data for these smaller geographies, please use table C16001. Additional languages are also available in the Public Use Microdata Sample (PUMS), at the State and Public Use Microdata Sample Area (PUMA) levels of geography.

## *Expanded household-level data*

A new household language table, B16002, is available using 2016 ACS 1-year data. This table shows more detailed data for households defined as limited English-speaking.[3] The 2016 B16002 table uses the same categories as table C16001, but at a household level.

A collapsed household language 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. The 2016 table C16002 is directly comparable with the 2015 household language table B16002. In addition to this new detailed household language data, more new household language data are available in the 1-year 2016 PUMS, using the new variable HHLANP, which contains the same languages as listed for individuals (LANP).

---

[3] A "limited English speaking household" is one in which no member 14 years old and over (1) speaks only English or (2) speaks a non-English language and speaks English "very well." In other words, all members 14 years old and over have at least some difficulty with English. By definition, English-only households cannot belong to this group. Previous Census Bureau data products referred to these households as "linguistically isolated" and "Household where no one age 14 and over speaks English only or speaks English 'very well.'"