

## 2021 Limited English Proficiency Population Estimate Methodology

### Purpose

Understanding what languages are spoken across the state, and the number of individuals who speak a given language, is critical to ensuring language access to important information for all Washingtonians. The Washington State Office of Financial Management (OFM) created the Limited English Proficiency Population Estimate to estimate populations across the state. Users are advised that these estimates are not actual counts of the Limited English Proficiency (LEP) population, nor are they intended to substitute for counts. Rather they were created to help meet the access and planning needs of state agencies in providing language access.

Providing language access is a requirement for all organizations and agencies that receive federal financial assistance. Therefore, this requirement applies to nearly all state agencies, local governments, and more. To meet this requirement, the Department of Justice provides the [Safe Harbor provision](#) that recommends translating vital information into every language that is used by 5% of the population or 1,000 people (whichever is less).

The Limited English Proficiency Population Estimate can be used to understand what languages at both the state and county levels meet the Safe Harbor Threshold. Some state agencies, including the Department of Health, use this list as a baseline for translating vital information. More recently, the [COVID-19 Language Access Plan](#) for the Washington State Coronavirus Response required that all cabinet agencies translate vital documents into the 36+ languages identified by the 2016 OFM analysis.

### Overview of Methodology

To estimate the number of people who use a specific language in the state, this analysis utilizes three sources of data:

1. From the [Office of the Superintendent of Public Instruction \(OSPI\)](#): Student's primary language (defined as a student's native or first language spoken) reported by households. The March 15, 2021, student records are used.
2. From the 2010 Census: Household size information for specific racial and ethnic groups.
3. From the [Department of Social and Health Services \(DSHS\)](#): Languages requested by clients for language assistance accommodation for DSHS services. To be consistent with the OSPI time period, client data from March 2021 are used.

Only language groups that either 1) represent at least 250 students, or 5% of the student population, from the OSPI student population estimates, or 2) had at least 500 DSHS clients request language accommodations in the estimate year, are included in this analysis. A 250-person threshold for inclusion is used for the OSPI data because the student population is limited to people aged 5-18 (with limited exceptions). As the student population is expanded to the total population (using the methodology described below), many language groups with 250 students meet the 1,000-person Safe Harbor threshold. DSHS clients requesting language assistance include the entire family unit, but only

for those individuals/families who qualify for benefits. A 500-person threshold is used to qualify for listing in recognition that even though the entire benefit family is represented in the data, there may be other household members and/or individuals speaking the language who do not qualify or have not applied for benefits.

**OSPI and Census data details**

Student level data were obtained from the Comprehensive Education Data and Research System's (CEDARS) database that is maintained by OSPI. The data represented here is the primary language of students enrolled in grades one through 12 as of March 15, 2021. Primary language refers to the self-identified first learned language spoken by the student.

The OSPI data only represents enrolled first through 12th grade students enrolled in schools. To estimate the total population in specific language groups, we multiplied the student population by the average household size of the race/ethnic group according to the 2010 Census. For Spanish speakers, we used the average household size at the county level. For all other language groups, we used the average household size at the state level. An example of this methodology is below:

# of students who speak a language (OSPI)	×	Household size of closest race/ethnic group (2010 Census)	=	Estimate of language speakers in WA
10,129 students speak Russian	×	2.37 (estimated household size based on white non-Hispanic population)	=	Approximately 24,006 Russian speakers in WA

The average household size of 2.51 across all language groups was not used because it is lower than the average household size of numerous race and ethnic groups. This would result in an underestimation of population size.

**DSHS data details**

If not already meeting the criteria for estimate using the OSPI data, a language is also included if at least 500 DSHS clients request translation for March of the estimate year. If the language was identified using both DSHS and OSPI data, we use the estimates from OSPI rather than the DSHS count. The count of clients for March 2021 is used for languages qualifying in this manner. The count of DSHS clients includes the entire benefit family so we use the DSHS client count rather than create an estimate of the LEP population.

DSHS clients include individuals accesses a variety of cash, food, and medical programs including: Temporary Assistance for Needy Families (TANF); State Family Assistance (SFA); Diversion Cash Assistance (DCA); Additional Requirements Emergent Needs (AREN); Refugee Cash Assistance (RCA);

Consolidated Emergency Assistance Program (CEAP); Disaster Cash Assistance Program (DCAP); State Supplemental Program (SSP); Aged, Blind, or Disabled Cash Assistance (ABD); A Referral to the Housing and Essential Needs (HEN) program administered by the Department of Commerce; Supplemental Nutrition Assistance Program (SNAP); Food Assistance Program (FAP) for Legal Immigrants; Working Family Support (WFS); Non-Immigrant Emergency Medicaid; SSI and SSDI Related Medicaid; Medicare Savings Program; Medical Care Services; Healthcare for Workers with Disabilities; Long Term Care; Refugee Medical; and both Federal and State Medical for Adults, Children, Pregnant Women, and Families. Clients are unduplicated across different programs.

Primary language comes from the language indicated by the client during eligibility determination. It is used for letters, notices, and other written communications.

### **Pairing language to detailed race/ethnicity household size**

To estimate total population from the OSPI student population counts, we need household size by language group. The Census data doesn't have average household size available by detailed languages but does provide household size data for race/ethnicity groups, such as Hispanic, Asian, Native Hawaiian, Other Pacific Islander, Native American, and Alaskan Native groups. Therefore, pairing languages to detailed race/ethnicity is needed. For many language groups the pairing is straightforward i.e., Bengali to Bangladeshi, Burmese to Burmese, Cambodian to Cambodian, etc. Other pairings are not as straightforward.

Detailed categories for White, Black/African American, and other race groups are not available in the 2010 Census, leading to pairing languages to broader racial groups that may be less reflective of the language group. The race and ethnicity coding from the 2010 Census can be found in the [2010 Census Summary File 1: 2010 Census of Population and Housing: Technical Documentation](#) (see Hispanic or Latino and Race under Appendix F). The Census Bureau used the [1997 OMB standards](#) to categorize race and ethnicity in the 2010 Census. We tried to be consistent with how the Census Bureau coded detailed race and ethnic groups in 2010. Languages from Europe (except Spanish) are paired with White alone, not Hispanic or Latino and languages from Africa (apart from North Africa) are paired with Black or African American.

According to a [Frequently Asked Questions](#) page from the Census Bureau's website, "White" is defined as "A Person having origins in any of the original peoples of Europe, the Middle East, or North Africa." The Census Bureau coded most Middle Eastern countries as White, so we paired languages spoken in those countries with White alone, not Hispanic or Latino household sizes. Arabic is spoken in many countries spanning from North Africa to the Middle East. Since most of those countries are categorized under White, we paired Arabic with White alone, not Hispanic or Latino. This is a departure from the 2014 and 2016 LEP estimates where Arabic was paired with Black or African American.

We made an exception to this rule for a few languages spoken in countries including and surrounding Pakistan. Those listing Pakistani origin were categorized by the 2010 Census under Asian. Since Asian racial groups had household sizes for detailed race available, the Pakistani household size was available for Washington state.

The languages paired with the Pakistan household size include:

- Dari, the official language in Afghanistan
- Farsi, the official language in Iran and Tajikistan
- Pashto, spoken in Pakistan and Afghanistan
- Urdu, official state language of Pakistan and officially recognized language in India

The rationale for using the Pakistani household size for these languages is that the language communities originating in countries/communities close to Pakistan and are likely to have more similar household sizes to those of Pakistani decent than to the White alone, not Hispanic, or Latino decent.

### **County estimates**

We used the same population thresholds and methodology for estimate inclusion for the county level estimates.

For all populations except Hispanic, the statewide estimate for the race/ethnic household size is applied. County-specific Hispanic household sizes are available and are used for Spanish. Based on feedback from members of the Governor’s Interagency Council on Health Disparities in 2014, county level estimates Chinese-Mandarin, Chinese-Taiwanese, and Chinese-Unspecified are combined into Chinese-Mandarin. Similarly, Chinese-Cantonese, Chinese-Fukienese, and Toishanese are combined into Chinese-Cantonese.

### **Limitations of this methodology**

OFM pairs languages identified in the OSPI data with race/ethnicity from Census data. This is an imprecise science.

The following limitations have been identified:

- The language to race/ethnicity pairings assume racial/ethnic homogeneity within the language group.
- Household size is available for some but not all race and Hispanic ethnicity groups. Specifically household size is not available for detailed ethnicities within the White, Black, or African American, and Other Race classifications. Due to data availability, some language groups have a more tailored estimates than others.
- Race and ethnic categories are very complicated. There is lots of diversity within every category. While the race and ethnicity categories are consistent with the [1997 OMB standards](#), many may disagree with these categorizations.

There may be over- or under-representation of language service need:

- Over-representation: Including the household of students with siblings more than once, language ability may vary across family members, using a household size that is larger than the actual size, and uncertainty regarding what the language ability threshold is for listing a non-English language in OSPI data.
- Under-representation: school does not register a student's language need, the student speaks English, but his/her parents do not, a sizable population speaking a language with no school age children and/or children attending public school, the household size applied to the language group is lower than the actual size, and/or a large language group does not have 500 people or more in the county/state receiving DSHS benefits.

Household size for the race/ethnic group is obtained from the 2010 Census because the relevant data from the 2020 Census is still unavailable.

We recognize that language and race/ethnicity pairings are not always precise and will not always represent how individual language speakers identify, but this methodology was determined to be the best approach with the limited data and resources available at this time. With each new version of these estimates, we will revisit the methodology to see if we can better reflect how individuals identify. **If you have any feedback on language and race/ethnicity pairings, please email the email address provided at the bottom of this document.**

### **Ongoing work**

As the Census Bureau continues to release the data needed to be able to switch from using the household size from the 2010 Census to the 2020 Census, we will need to re-evaluate the methodology used for these estimates. We could have more detailed race data available (including for White and Black or African American categories) that would allow for improved language to detailed race/ethnicity pairings. However, there is a great deal of uncertainty about the timing of the releases, the tables that will be available for the detailed race/ethnicity groups, and the quality of the data that will be released from the 2020 Census. Depending on when the next update of the LEP estimates is produced, we may only have the household size for aggregate race and Hispanic categories from the 2020 Census. About the only certainty that we have is that we will need to re-evaluate our methodology at the time of the next estimate update.

We welcome and appreciate feedback on the methodology and the race/ethnicity pairings used for the OSPI based language estimates (please email: [ofm.forecasting@ofm.wa.gov](mailto:ofm.forecasting@ofm.wa.gov)). We will incorporate your feedback into future estimates and/or revise the estimates as appropriate.