



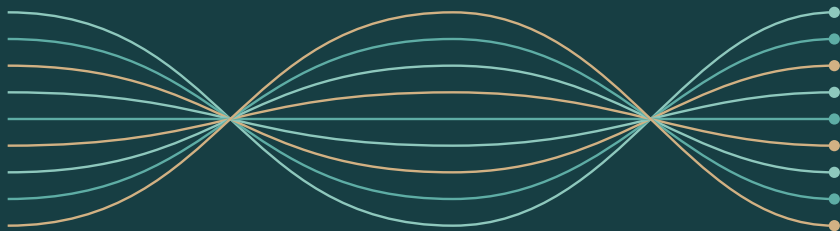
SECOND
LANGUAGE
PROJECT

LANGUAGE • CONNECTION • COMMUNITY

A COMMUNITY RESEARCH GUIDE

Languages of the Bay Area

Home, learning, work
and community access



NINE COUNTIES. MANY WAYS TO CONNECT.

Start with curiosity.

Languages connect people to home, learning, work and community. Each setting invites a different question.

01 What do people use at home?

Explore the regional patterns, then look closer at counties, schools and families.

02 What do people want to learn?

Consider existing language skills, learning goals and the situations in which people want to participate.

03 What helps people take part?

Ask about communication preferences and the support people find useful.

Languages of the Bay Area

Home, learning, work and community access

A guide to what the evidence shows, what it does not measure, and how to ask better questions.

Nine counties. Many language repertoires. Different needs in different settings.

Prepared for the Second Language Project

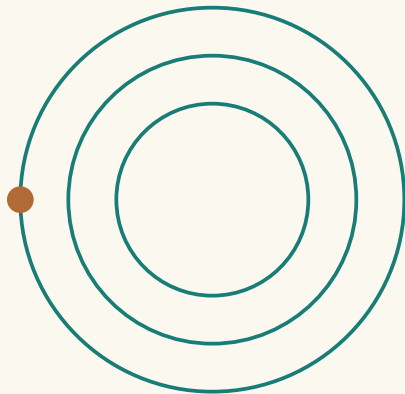
22 October 2026

English edition • 40 pages • public aggregate data

2020-2024 American Community Survey

2025-26 California English learner records

District documents and selected research literature



Start with the question

About 3.18 million Bay Area residents age 5 and older report a language other than English at home: 43.9% of that population. This includes people with many different levels of English ability. [C01](#), [C02](#), [C03](#)

The most useful headline is linguistic diversity. A ranking alone cannot decide which language a person prefers, which language they learned first, or what support they need for a particular conversation.

Read at your level	Where to go
Quick orientation	Pages 2-3 and 6-10
Primary and additional languages	Pages 4-5 and 12-19
Work and language access	Pages 20-23
Survey, methods and evidence	Pages 24-31
Sources and detailed tables	Pages 33-39

An evidence-based research edition, not a peer-reviewed journal article. Findings, interpretations and proposed practices are labeled separately. The online reader adds filters, source links and downloadable data.

Three perspectives, three denominators

Perspective	Evidence and its boundary
Home	ACS: residents age 5+, reported home-language category and English-speaking ability.
Work	A San Francisco staffing case and measurement research. No region-wide workplace-language ranking established.
Community	School English learner records, district plans and service-access guidance. Each describes a different population.

Sources: C01, C07, D02

School counts are attached to enrollment records and school locations. ACS counts concern residents. A child may cross a district boundary to attend school. Neither population can be substituted for the other.

Workplace language use would require a defined population of workers, jobs and locations. Bilingual position requirements show institutional demand; they do not count everyone who uses a language at work.

Interpretation: use the three perspectives together to design questions and services. Do not add their counts or treat them as independent estimates of the same population.

What does “primary language” mean?

Possible meaning	What to ask
Most-used at home	Which language(s) do you use most often at home during a stated period?
First acquired	Which language(s) did you learn earliest?
Strongest language	Strongest for which skill, task and setting?
Preferred for a service	Which language and format would help for this interaction?
School administrative primary language	What does the district instrument and coding rule mean by the field?

Sources: C04, AL03, AL04

These answers can differ within one person. This report uses “primary home language” only for an instrument that actually asks about most-frequent home use. It allows ties rather than forcing one language.

Research on bilingual profiles treats history, current use and proficiency as separate dimensions. Earliest acquisition does not establish current dominance. The same person can be comfortable speaking a language and less comfortable reading it. AL03, AL04, AL06

Proposed reporting rule: retain the source’s original term, then state the operational definition alongside it. Do not silently rename an ACS category “primary language.”

What the ACS actually asks

The ACS asks whether a person sometimes or always speaks a language other than English at home, which language, and how well the person speaks English. If several non-English languages are used, instructions select the one used most often; if that cannot be determined, the first learned. This does not establish whether that language is used more often than English.

C04

Standard coding retains one non-English language. Research on multiple write-ins demonstrates why the published field cannot describe a person's entire repertoire. A Census conference poster explores this limitation; it is research, not a peer-reviewed regional estimate. C06

"English only" means only English reported at home. It does not mean the person knows no other language. "Non-English at home" does not mean English is absent or that the person needs an interpreter. C04, C05

Also distinguish people from households. ACS household language is assigned by an ordering rule using household members' responses. It is not a direct vote on the household's dominant language. This guide ranks people, not households. C04

A regional snapshot

2020-2024 estimate	Residents age 5+
All home-language categories	7,234,466
English only at home	4,055,544
A non-English language at home	3,178,922
Share reporting a non-English language	43.94%

Sources: [C01](#), [C02](#), [C03](#)

The region is Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano and Sonoma counties. We reconstruct it from 62 whole 2020 Public Use Microdata Areas (PUMAs), all verified to nest within those counties. County results in this guide are derived from the same PUMAs. [C03](#)

The five-year estimate describes a period, not the population on one day in 2024 or 2026. Counts are survey estimates, not a complete enumeration. The printed precision supports reproducibility; it should not imply exact head counts.

Use “about 44% report a non-English language at home” in public-facing copy. Always retain the age-5+ universe, the 2020-2024 period and a source link.

Top 20 reported categories

Rank	Language category	Estimate	Share
1	English only	4,055,544	56.06%
2	Spanish	1,235,891	17.08%
3	Chinese (incl. Mandarin, Cantonese)	591,975	8.18%
4	Tagalog (incl. Filipino)	235,935	3.26%
5	Vietnamese	172,101	2.38%
6	Hindi	122,462	1.69%
7	Korean	64,070	0.89%
8	Persian (incl. Farsi, Dari)	54,134	0.75%
9	Russian	51,979	0.72%
10	Telugu	46,628	0.64%
11	French (incl. Cajun)	45,138	0.62%
12	Tamil	44,169	0.61%
13	Punjabi	41,299	0.57%
14	Arabic	37,552	0.52%
15	Ilocano, Samoan, Hawaiian, or other Austronesian languages	35,408	0.49%
16	Japanese	34,748	0.48%
17	Nepali, Marathi, or other Indic languages	34,561	0.48%
18	Portuguese	33,380	0.46%
19	Other languages of Asia	28,636	0.40%
20	Malayalam, Kannada, or other Dravidian languages	26,999	0.37%

Sources: C01, C02, C03

Includes English. Source category labels include grouped languages; this is not a ranking of 20 separately identified languages. Point-estimate ranks may overlap in uncertainty. Full categories and margins of error are in the explorer.

What a ranking leaves out

Chinese combines Mandarin, Cantonese and other Chinese varieties. Persian includes Farsi and Dari. Several categories combine distinct languages, such as Ilocano, Samoan, Hawaiian and other Austronesian languages. Their combined rank cannot be assigned to any one member. [C02](#)

The dataset separates Hindi, Telugu, Tamil and Punjabi, but combines Nepali and Marathi with other Indic languages, and Malayalam and Kannada with other Dravidian languages. Category design changes the apparent league table. Preserve these labels when comparing sources. [C02](#)

A lower regional count does not imply a less important access need. A language may be concentrated in one neighborhood, one school or one service audience. A small language community can be missed by a top-20 cutoff.

ACS does not provide a separate count of American Sign Language users in this table. It also does not enumerate languages known but unused at home. Absence from this ranking is not evidence of absence from the community.

[C05](#)

Comparable statistics need a consistent geography, period and category system. Do not combine a nine-county regional total with a differently defined metro or a figure from another year.

The regional average hides local patterns

County	Age 5+	Non-English home
Alameda	1,563,092	47.2%
Contra Costa	1,103,088	37.5%
Marin	246,644	23.9%
Napa	128,840	34.2%
San Francisco	796,324	43.6%
San Mateo	704,214	45.7%
Santa Clara	1,803,289	55.2%
Solano	426,437	31.5%
Sonoma	462,538	27.0%

Sources: C01, C03

The share ranges from about 24% in Marin to about 55% in Santa Clara. These are shares of each county’s age-5+ population, not each county’s contribution to the region.

Language-access planning should examine local concentrations as well as regional totals. The explorer offers all nine county aggregates and 62 PUMAs. PUMAs are statistical areas; they are not school districts or necessarily city boundaries.

Derived counties: the values are sums of complete PUMAs, not published B16001 county cells. Do not silently relabel a PUMA as a city.

Three useful local comparisons

Area and category	Age-5+ estimate
San Francisco: Chinese	139,507
San Francisco: Spanish	88,555
Santa Clara: Vietnamese	119,090
Alameda: Chinese	145,977
Alameda: Spanish	260,285

Sources: C01, C02, C03

In San Francisco, the Chinese category exceeds Spanish among reported non-English home languages. Santa Clara accounts for about 69% of the region’s Vietnamese category. These geographic patterns do not establish an individual’s preferred communication language.

Chinese is a combined category here. District records can distinguish Cantonese and Mandarin, but their school populations differ from the resident ACS population. More specific labels do not make the two datasets directly interchangeable.

Use local evidence to choose outreach partners and questions. Then ask participants which spoken language, written language or script, and communication format they want. A population table cannot supply those individual answers.

English ability is a separate measure

Among residents reporting a non-English home language, 1,224,228 report speaking English less than “very well”: 38.5% of that group, or 16.9% of all residents age 5+. This is a self- or household-reported survey category. [C01](#), [C04](#)

The published split separates “very well” from the combined “well,” “not well,” and “not at all” responses. It concerns speaking English, not reading, writing, translation skill or a tested level of proficiency. [C04](#)

For reproducibility, the data store names the field `english_less_than_very_well`. The reader labels its denominator explicitly. We do not convert it into a count of people who always require an interpreter.

Interpretation: complex health, school, legal or civic conversations can create different communication needs from ordinary conversation. Ask about the task and preferences; do not assign need from ethnicity, surname or home language.

Nor does this split measure “English as a second language.” ACS does not establish language acquisition order. Someone reporting English very well may still be learning another language. [C04](#), [AL03](#)

Schools ask different questions

School language evidence can describe a student’s background, English learner status, instructional needs or a family’s communication preferences. These are related but distinct fields. Home Language Surveys help identify students for English-language assessment; they do not by themselves test proficiency. [D01](#), [D02](#)

CALPADS primary language is a stable administrative designation informed by early and home language information. It is not an annually updated measure of current dominance. Castro Valley and San Leandro surveys distinguish first spoken language, most frequent child home language, adult-child language and adults’ home language. [D04](#), [D01](#), [D05](#)

Population or field	Do not equate with
English learner (EL)	Every student using a non-English home language
Initially fluent English proficient	English-only home background
Reclassified fluent English proficient	No continuing multilingual resources or interests
Parent communication preference	Student’s administrative primary-language field

Sources: [D01](#)

A district-wide ranking needs all acquisition statuses, the coding rule and enrollment denominator. Family communication preferences need separate fields.

Current data in this report: CDE’s 2025-26 EL-by-language file. Historical context: Castro Valley’s 2021-2024 plan. These are not a time series of the same measure.

Castro Valley: read the labels carefully

The 2021-2024 district plan’s language analysis makes the importance of home-language evidence clear. Its 2020-21 acquisition-status chart lists English Only 68.12%, EL 9.66%, initially fluent 6.18% and reclassified fluent 16.05%. [D01](#)

A nearby “Ten Most Commonly Spoken Languages in CVUSD” chart lists Cantonese 917, Spanish 686 and Mandarin 312. An adjacent chart refers to English learners but includes English. The population labels are not sufficiently clear to treat those figures as a current all-student or EL ranking. They are preserved as historical evidence, with this ambiguity. [D01](#)

2025-26 EL records only	Students
Spanish; Castilian	425
Yue Chinese (Cantonese)	281
Mandarin Chinese (Putonghua, Guoyu)	152
Vietnamese	40
Russian	34
Arabic	28

Sources: [D02](#)

Current EL total: 1,166. Separately, the August 2026 district newsletter reports about 9,750 students and 60 home languages. Its October newsletter announces a new multilingual plan; the underlying plan was not recovered. None supplies a verified current all-family language ranking. [D06](#), [D07](#)

San Leandro: an EL population profile

2025-26 EL records only	Students
Spanish; Castilian	1,274
Yue Chinese (Cantonese)	231
Uncoded languages (Other non-English languages)	112
Mandarin Chinese (Putonghua, Guoyu)	92
Arabic	90
Vietnamese	65
Persian (Farsi)	22
Philippine languages	18

Sources: [D02](#)

The CDE file yields 1,992 English learners in San Leandro Unified. Spanish accounts for 1,274 of these records (about 64%). Cantonese, Mandarin, Arabic and Vietnamese also appear prominently.

“Uncoded languages” is a published residual category. It is not one language. Do not distribute it among named languages without further evidence, or assume its students share the same communication needs.

The district is a useful local perspective for SLP. It is not identical to the city’s resident population, and EL enrollment is not an all-family language survey. A family’s communication preference requires its own field.

SLUSD’s plan describes Spanish and Mandarin dual-language immersion for English learners and English-dominant students. This demonstrates a learning pathway, not a count of all additional-language speakers. [D08](#)

A regional school-language lens

CDE language category	EL students
Spanish; Castilian	120,488
Mandarin Chinese (Putonghua, Guoyu)	6,036
Vietnamese	5,384
Yue Chinese (Cantonese)	4,579
Arabic	3,073
Persian (Farsi)	2,991
Uncoded languages (Other non-English languages)	2,451
Philippine languages	1,822
Russian	1,666
Panjabi; Punjabi	1,498
Japanese	1,365
Portuguese	1,289

Sources: D02, D03

The nine-county aggregation totals 166,614 English learners across 172 district entries. Spanish represents about 72.3% of EL records. These proportions use EL enrollment, not total enrollment or the resident population.

Language codes are more detailed than ACS in some places and less detailed in others. For example, the school file separates Mandarin and Cantonese, and also includes “Philippine languages” and other residual categories. Keep the published categories intact.

The full machine-readable store includes county, district and language aggregates. Very small published counts are retained in the research archive; this report emphasizes aggregates, not identification of individual students.

Let families answer the access question

Proposed practice: record the student’s language background separately from each caregiver’s communication preferences. One household may contain people who need different languages or formats.

Ask about	Useful distinction
Speaking or signing	Language preferred for a meeting or phone call
Written information	Language, script and reading comfort
Interpretation	Whether help is requested for this interaction
Child’s repertoire	Languages understood, spoken or signed, read and written
Updates	When preferences were last confirmed

SLUSD’s 2022-23 LCAP planned parent preferred-language settings in Aeries. This is historical evidence of a separate communication field, not a current preference distribution. Offer oral and accessible responses; keep nonresponse distinct from “English preferred.” [D09](#)

Recommendation, not a district finding: report preferences with their response rate and collection method. Do not generalize a parent survey to all families if only a self-selected group answered.

Preserve year, respondent, grade range, population and collection purpose. CDE updated reclassification guidance in September 2026; historical district procedures should not be treated as current rules. [D10](#)

“Second language” is only a starting point

For this guide, an additional language is one a person learns or develops alongside their earliest language or languages. This editorial term includes people learning a third or fourth language. It does not impose a single acquisition sequence on everyone.

The CEFR Companion Volume describes an interconnected repertoire with uneven and changing strengths. A learner can have different levels of listening, speaking, reading, writing and interaction. A single fluent/not-fluent label can hide useful capabilities. [AL02](#)

LEAP-Q research examines language history and self-reported experience alongside proficiency. Its original validation involved particular adult samples; a local abbreviated survey does not automatically inherit that validation. [AL03](#), [AL06](#)

There is no defensible count of all Bay Area “second-language speakers” in B16001. The table neither inventories every language a person knows nor establishes when each was learned.

SLP can separately describe existing language resources, desired learning, and access needs. An interest in learning English does not mean a person can already participate fully in an English-only service.

Recognize what learners already bring

Heritage-language development connects learning with family and intergenerational life. UCLA’s National Heritage Language Resource Center emphasizes existing learner resources and goals such as literacy and formal-register development. Heritage learners are not necessarily beginners. [AL05](#)

Illustrative profile	Possible learning goal
Uses a family language orally	Read messages, stories or formal documents
Understands more than they speak	Build confidence in conversation
English at home, studying Vietnamese	Connect with neighbors or colleagues
Several languages across settings	Develop one specific skill or professional register

These profiles are illustrations, not measured Bay Area population groups. The Bilingual Language Profile distinguishes history, use, proficiency and attitudes, including different settings and skills. This supports asking about actual activities rather than assuming ability from identity. [AL04](#)

Program recommendation: welcome beginners and people developing a family language. Ask participants how they describe their own language relationship; do not assign a heritage identity to them.

From population evidence to learning choices

Spanish is the largest reported non-English home-language category across the nine counties. Practical Spanish is therefore a reasonable program-development priority to investigate. Population size alone does not establish enrollment demand, teacher availability or the best learning format. C01

Vietnamese, Tagalog/Filipino, Japanese, Korean, Hindi and Persian/Farsi are all visible in the regional evidence. Chinese varieties and several South Asian languages also deserve attention. The exact list should remain open to local voices and staffing capacity.

Decision input	How to obtain it
Community presence	ACS and school data, with geography and date
Expressed learning goals	Voluntary participant needs survey
Practical feasibility	Instructor, partner and venue conversations
Access and inclusion	Language preferences, format needs and barriers

Before selecting a schedule or syllabus, pair population evidence with a direct assessment of local learning interests, available teachers and practical participation needs.

Language at work needs its own evidence

A resident can use one language at home, another with coworkers, and another with customers. Work may also involve reading, writing, interpretation or signing. The ACS home-language measure does not answer these workplace questions. [C04](#)

This review did not establish a representative nine-county ranking of languages used at work. Job advertisements, bilingual pay assignments and service requests are useful institutional indicators, but none counts all workers' language use.

Statistics Canada provides a useful design comparison: it distinguishes languages used most often at work from other regular use and allows ties. Its definitions concern work tasks, not merely the physical location of a person working from home. These are methodological examples, not Bay Area statistics. [AL07](#), [AL08](#)

A future workplace study must choose its population: residents who work, jobs located in the Bay Area, or workers served by an organization. Specify main versus additional jobs, remote work, reference period and customer-facing tasks.

A concrete workplace case: San Francisco

The San Francisco Budget and Legislative Analyst’s April 2025 report lists 1,157 positions with language special conditions in Exhibit 6. Spanish and Chinese account for most of these designated positions. The report’s staffing discussion concerns FY2022-23. [C07](#)

Designated language	Positions
Spanish	560
Chinese (includes varieties)	495
Vietnamese	32
Russian	29
Filipino	27
American Sign Language	8
Japanese	4
Khmer	1
Korean	1

Sources: [C07](#)

This is a case study of institutional staffing demand. A position is not a person, a filled job, a service encounter or an estimate of all workplace speakers. The report also discusses a broader certified bilingual employee pool; do not mix those denominators.

Practical use: compare language service requests with staffing capacity within the same institution and period. Do not use this table as the Bay Area’s workplace-language ranking.

Language access begins with participation

For a listening session, school meeting or neighborhood program, the planning question is practical: can people understand, express themselves, ask questions and follow up in a way that works for them?

Before	During and after
Ask language and format preferences	Provide requested support and explain how to use it
Use local evidence to plan outreach	Invite corrections to language labels and terminology
Prepare clear, readable materials	Check understanding without testing participants
Arrange appropriate interpreting	Share accessible follow-up and ways to respond

Recommendations should be tested with actual participants. Translated publicity may improve awareness, but it does not by itself make the event accessible. Interpretation, written translation and accessible communication formats address different parts of participation.

The MTC’s 2023 Title VI report provides a regional transportation planning example. Its audience, service context and period differ from this guide’s resident ranking. It belongs in the source library as institutional context, not as a substitute population estimate. [C10](#)

Include language, modality and context

Language assistance and language learning serve different purposes. A class develops skills over time. Interpretation or translation supports participation now. One should not be treated as proof that the other is unnecessary.

Effective communication can involve sign-language interpretation, captioning, accessible electronic text, large print or other aids. DOJ's ADA guidance emphasizes the person's usual method of communication and the nature, length and complexity of the interaction. Do not assume every Deaf person uses ASL or that all communication access is spoken-language translation. [C08](#)

For complex or consequential exchanges, the guide recommends qualified support appropriate to the task. Casual conversational ability does not demonstrate professional interpretation or translation competence. [C08](#)

Policy status, checked 22 October 2026: DOJ's July 2025 notice says [lep.gov](#) operations are suspended pending review following Executive Order 14224. Older federal guidance must not be presented as unchanged. This guide does not determine sector-specific legal duties. [C09](#)

A better local survey: home and preference

The following are original proposed fields, not a validated questionnaire. Pilot wording with community reviewers and make participation optional. Offer oral or accessible completion as well as written forms.

Field	Proposed question concept
Home use	All languages regularly used at home in the past month
Most-used home language	One or more equally most-used languages; allow uncertainty
Earliest languages	Languages learned earliest; allow simultaneous acquisition
Speaking or signing preference	Preferred language for this particular activity
Written preference	Preferred written language or script and format
Respondent	Self-report or a clearly identified proxy response

Use a consistent person-level unit. If asking about a household, explicitly allow differences between members. Record “not answered” separately from “none” or “English only.”

Do not retrospectively relabel ACS responses with these proposed definitions. The new survey would measure different concepts and would need its own recruitment and validation documentation. [C04](#), [AL08](#)

Add skills, work and learning goals

Dimension	What to collect
Language skills	Self-rated comfort understanding, speaking/signing, reading and writing, by language and task
Work tasks	Languages used with coworkers, customers and documents; most-used and other regular use
Learning goals	Language(s) to develop, activities, level and reasons
Participation barriers	Timing, cost, transport, digital access and preferred format
Provenance	Survey version, date, recruitment channel and response method

The proposed dimensions draw on research instruments that separate language history, use and proficiency. Their inclusion here is a design synthesis. We are not reproducing or validating the LEAP-Q or Bilingual Language Profile. [AL03](#), [AL04](#), [AL06](#)

A convenience sample can describe its respondents’ interests. It cannot establish region-wide prevalence without a defensible sample design and weighting. Report who was invited, who responded, and whose voices may be missing.

For a work module, define the job and reference period. A remote worker’s job language may differ from their household conversation language. [AL07](#), [AL08](#)

Rules for responsible analysis

Rule	Reason
Keep multiple answers	Language repertoires are not mutually exclusive categories
Publish denominators	Residents, EL students, survey respondents and positions differ
Keep missing values visible	Nonresponse is not a language preference
Separate observed and inferred	A program recommendation is not a population finding
Limit small cross-tabulations	Combinations can expose individuals even without names
Retain original labels	Translation or recoding can obscure meaning

For multiple-response survey questions, percentages may sum above 100%. Say so explicitly. For ACS B16001, the reported categories are mutually exclusive and sum to the age-5+ total. These are different data structures.

This publication uses only public aggregate data. A future local study should establish consent, access controls and retention practices before collecting personal information. It should avoid unnecessary names, exact employers or detailed demographic combinations.

Suggested quality check: have a second reader independently trace a sample of findings from printed page to source, variable, geography and calculation. Reproducibility and community review serve complementary purposes.

How the regional estimates were built

We use the Census Bureau’s table-based 2020-2024 ACS five-year B16001 file and its variable dictionary. Published PUMA rows were selected using the official 2020 PUMA-to-county-subdivision relationship file. No population allocation or proportional splitting was applied. [C01](#), [C02](#), [C03](#)

All 62 selected PUMAs fall entirely within one of the nine target counties. The script verifies complete inclusion and absence of outside-county areas, then sums the same disjoint areas for each county and the region.

Step	Reproducible check
Geography	62 PUMAs; 9 counties; 1 region
Categories	English only plus 42 non-English categories
Arithmetic	43 categories sum to every geography total
English split	Very well + less than very well = category estimate
Output	72 geographies × 43 = 3,096 estimate records

Source fields are preserved by B16001 variable ID. Source identifiers, file hashes and calculation methods support independent checking. The public reader contains derived aggregates and source links.

Uncertainty is part of the finding

For a sum of geographic estimates, this guide approximates the 90% count margin of error as the square root of the sum of squared component margins of error. This conventional approximation does not model covariance between estimates. [C12](#)

$$\text{MOE}(\text{sum}) \approx \sqrt{(\text{MOE}_1^2 + \text{MOE}_2^2 + \dots + \text{MOE}_n^2)}$$

Approximate interval: estimate $\pm$ MOE; display a lower bound no smaller than zero.

The explorer retains count MOEs for each category. Percentages use the corresponding age-5+ point-estimate total. We do not compute share confidence intervals, significance tests or rank confidence intervals. Do not use count MOEs as if they were percentage-point MOEs.

Adjacent ranks can be uncertain. “Rank 10” means tenth by the point estimate, not statistically greater than every category below it. Small-area estimates can be particularly imprecise.

The five-year period improves coverage but combines observations across years. A change comparison needs compatible categories and geographies, and should account for overlapping survey periods. This edition does not estimate growth rates.

How school records were aggregated

The CDE 2025-26 school-by-language EL file contains 59,111 statewide rows. Filtering the nine counties selects 15,630 rows, covering 1,764 school or office codes and 172 district entries. These records aggregate to 166,614 EL students. [D02](#), [D03](#)

The validation checks nonnegative integer values, duplicate school-language keys, grade sums against TOTAL_EL, and agreement across county, district and language totals. No duplicate school-language rows or arithmetic failures were found.

The source includes 44 rows explicitly labeled District Office with a zero school code, totaling 166 EL students. We retain them as published entries; they are small enrollment records rather than district-wide subtotals. These records are kept distinct from school-level entries in the analysis.

Limits: administrative coding, enrollment coverage and classifications differ from ACS survey measurement. This edition does not estimate all-student primary language or parent communication preference from the EL file.

For the current all-student language picture, obtain a district export across acquisition statuses and documentation of its Home Language Survey coding. Preserve school year and the district's enrollment population.

What is supported, and at what level?

Statement or question	Evidence status
About 44% use a non-English home language	Derived ACS statistic; age 5+, 2020-2024
Top 20 categories including English	Reproducible point-estimate order, with grouped labels
A primary-home-language top 20	Not established by ACS or available district extracts
Current CVUSD and SLUSD EL counts	CDE administrative aggregation, 2025-26
A Bay Area second-language count	Not measured in the sources reviewed
A workplace-language top 20	Not established; staffing case only
Program and survey suggestions	Report recommendations, not tested findings

Sources: [C01](#), [D02](#), [C07](#)

This is a targeted evidence review, not a systematic review. Sources were selected for direct relevance, primary authority and accessible definitions. We prioritized official data, original research and institutional methods over secondary rankings.

Citation IDs connect paragraphs to sources. The data store keeps finding text and citations, so readers can inspect the distinction between evidence and interpretation.

What would strengthen the next edition?

First, obtain current district-wide language and caregiver-preference data. Castro Valley's historical chart needs a clarified denominator; current EL records alone cannot answer the primary-language question. An October 2026 announcement records approval of a new multilingual plan, which we could not retrieve. D01, D07

Second, conduct a carefully piloted local needs assessment. Measure most-used home languages, full repertoires, communication preferences and learning interests separately. Document recruitment so readers can judge how far the findings generalize.

Third, collect comparable institutional service indicators: requests, completed assistance, unmet requests and participant feedback, using the same period and audience. Staffing counts can then be interpreted alongside actual service demand.

Sources not used as evidence: an inaccessible full-text heritage-language article and an unavailable ACTFL PDF are recorded as optional leads in the research memo. No conclusion depends on them. District documents not obtained are listed in the research notes.

Independent methodological review and community feedback can refine terminology, source interpretation and proposed survey questions. Check the edition date when comparing findings, since later evidence may change the interpretation.

Using this evidence in practice

Use this guide to identify questions, plan outreach and start conversations with residents. Regional patterns can help you decide where to look more closely; local experience helps explain what support people want.

Before deciding	Additional evidence to seek
Service languages and formats	Ask residents about spoken, written and accessibility preferences.
Staffing, classes and budgets	Measure local demand, workload, participation and costs.
City or neighborhood priorities	Use supported estimates for the relevant area and population.
Why a disparity occurs	Combine further analysis with community research.
Whether a program works	Evaluate outcomes with a suitable comparison or study design.

Until more local evidence is available, avoid treating regional rankings as a service plan or a home-language category as an individual preference. Later analyses may refine the picture; document the population, period and uncertainty behind each decision.

Use population evidence alongside listening. Findings can guide questions and outreach, but they do not establish causes or demonstrate a program’s effects.

Sources and reading notes

C01. 2020-2024 ACS B16001 table-based summary file

U.S. Census Bureau. 2026-01-29 file listing. B16001 E001-E128 and M001-M128; 795P200 PUMA rows

www2.census.gov • [linked original](#)

C02. B16001 variable definitions

U.S. Census Bureau. 2024 dataset. Labels for estimates and margins of error

api.census.gov • [linked original](#)

C03. 2020 PUMA to county subdivision relationship file

U.S. Census Bureau. 2020 geography. California PUMA and county-subdivision GEOIDs

www2.census.gov • [linked original](#)

C04. 2024 ACS Subject Definitions

U.S. Census Bureau. 2024. Language Spoken at Home and Ability to Speak English

www2.census.gov • [linked original](#)

C05. Frequently Asked Questions About Language Use

U.S. Census Bureau. 2021-12-16 page revision. Language questions; ASL; household language

www.census.gov • [linked original](#)

C06. Examining Multilingualism in the United States Using ACS Language Write-ins

Daniela Mejía, U.S. Census Bureau. 2024-04. Background and data; methods; research disclaimer

www.census.gov • [linked original](#)

C12. Calculating Margins of Error the ACS Way

U.S. Census Bureau. 2020. Derived estimates and uncertainty

www.census.gov • [linked original](#)

Sources and reading notes

C07. Language Access Services

San Francisco Budget and Legislative Analyst. 2025-04-14. Exhibit 6; staffing and language-access analysis

sfbos.archive.sf.gov • [linked original](#)

C08. ADA Requirements: Effective Communication

U.S. Department of Justice. 2020-02-28 update. Auxiliary aids; qualified interpreters; context and communication methods

www.ada.gov • [linked original](#)

C09. Limited English Proficiency notice

U.S. Department of Justice. 2025-07-17 notice. EO14224 and suspended operations notice

www.lep.gov • [linked original](#)

C10. MTC Title VI Triennial Report 2023

Metropolitan Transportation Commission. 2023. Language assistance plan and county language context

mtc.ca.gov • [linked original](#)

C11. DataQuest

California Department of Education. live data portal. English learner and enrollment reports

www.cde.ca.gov • [linked original](#)

D01. 2021–2024 Plan for English Language Learners

Castro Valley Unified School District. 2021–2024 plan; chart observations 2020–21. Printed pp.9–11 charts (PDF pp.10–12); printed pp.19–20 identification and HLS (PDF pp.20–21)

files.smartsites.parentsquare.com • [linked original](#)

D02. English Learners by Grade & Language — elsch25

California Department of Education. 2025–26.

dq.cde.ca.gov • [linked original](#)

Sources and reading notes

AL02. Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume

Council of Europe. 2020-04. Section 2.3, printed p.30; Section 2.7, printed pp.38–40
rm.coe.int • [linked original](#)

AL03. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals

Viorica Marian, Henrike K. Blumenfeld, Margarita Kaushanskaya. 2007-08. Abstract, printed p.940; Questionnaire appendix
[10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067)) • [linked original](#)

AL04. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism

David Birdsong, Libby M. Gertken, Mark Amengual. 2012. Overview; BLP Components modules 1–4
blp.coerll.utexas.edu • [linked original](#)

AL05. About the Center

National Heritage Language Resource Center, UCLA. Undated. Mission
nhlrc.ucla.edu • [linked original](#)

AL06. The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later

Margarita Kaushanskaya, Henrike K. Blumenfeld, Viorica Marian. 2020-11. Abstract; Best-practice discussion
[10.1017/S1366728919000038](https://doi.org/10.1017/S1366728919000038) • [linked original](#)

AL07. Language used most often at work of person 15 years or over

Statistics Canada. 2022-08-16. Definition; Usage
www23.statcan.gc.ca • [linked original](#)

AL08. Dictionary, Census of Population, 2021

Statistics Canada. 2023-11-15. Language spoken most often at home: printed pp.309–310; PDF indexes322–323; Language used most often at work: printed pp.310–312; PDF indexes323–325; Mother tongue: printed pp.312–313; PDF indexes325–326
www12.statcan.gc.ca • [linked original](#)

Sources and reading notes

D03. File Structure: ELs by Grade & Language

California Department of Education. 2026-04-10. File structure table
www.cde.ca.gov • [linked original](#)

D04. Primary Language — CALPADS User Manual

CALPADS. accessed 2026-10-22. Primary Language definition, three paragraphs
documentation.calpads.org • [linked original](#)

D05. EL Master Plan: Chapter 1 — Initial Identification, Assessment & Placement

San Leandro Unified School District. accessed 2026-10-22. Identification: four HLS questions; Assessment: initial ELPAC, IFEP, EL, RFEP
www.slusd.us • [linked original](#)

D06. CVUSD Community Connection Newsletter — August 15, 2026

Castro Valley Unified School District. 2026-08-15. From the Board Meeting / opening-school update
cv.k12.ca.us • [linked original](#)

D07. CVUSD Community Connection Newsletter — October 16, 2026

Castro Valley Unified School District. 2026-10-16. Strengthening Support for Multilingual Learners
www.cv.k12.ca.us • [linked original](#)

D08. EL Master Plan: Chapter 2

San Leandro Unified School District. accessed 2026-10-22. Dual Language Immersion program section
www.slusd.us • [linked original](#)

D09. LCAP and LCFF 2022–23

San Leandro Unified School District. 2022–23. Action 2.20 Multilingual Family Engagement Webinars and Workshops
www.slusd.us • [linked original](#)

D10. Survey, Assessments, Programs, & Reclassification

California Department of Education. 2026-09-27. Update notice; Home Language Survey section
www.cde.ca.gov • [linked original](#)

All language categories

Rank	Category	Estimate	Share
1	English only	4,055,544	56.06%
2	Spanish	1,235,891	17.08%
3	Chinese (incl. Mandarin, Cantonese)	591,975	8.18%
4	Tagalog (incl. Filipino)	235,935	3.26%
5	Vietnamese	172,101	2.38%
6	Hindi	122,462	1.69%
7	Korean	64,070	0.89%
8	Persian (incl. Farsi, Dari)	54,134	0.75%
9	Russian	51,979	0.72%
10	Telugu	46,628	0.64%
11	French (incl. Cajun)	45,138	0.62%
12	Tamil	44,169	0.61%
13	Punjabi	41,299	0.57%
14	Arabic	37,552	0.52%
15	Ilocano, Samoan, Hawaiian, or other Austronesian languages	35,408	0.49%
16	Japanese	34,748	0.48%
17	Nepali, Marathi, or other Indic languages	34,561	0.48%
18	Portuguese	33,380	0.46%
19	Other languages of Asia	28,636	0.40%
20	Malayalam, Kannada, or other Dravidian languages	26,999	0.37%
21	German	24,541	0.34%
22	Gujarati	22,973	0.32%

Sources: C01, C02, C03

Nine-county Bay Area. Residents age 5+. ACS 2020-2024. Point-estimate ordering. Complete variable IDs, count MOEs and English-ability splits are in the data store.

All language categories

Rank	Category	Estimate	Share
23	Amharic, Somali, or other Afro-Asiatic languages	19,544	0.27%
24	Urdu	19,303	0.27%
25	Other Indo-European languages	17,280	0.24%
26	Italian	17,260	0.24%
27	Thai, Lao, or other Tai-Kadai languages	14,707	0.20%
28	Hebrew	13,276	0.18%
29	Bengali	13,057	0.18%
30	Ukrainian or other Slavic languages	11,430	0.16%
31	Khmer	9,970	0.14%
32	Yoruba, Twi, Igbo, or other languages of Western Africa	8,908	0.12%
33	Other and unspecified languages	8,567	0.12%
34	Yiddish, Pennsylvania Dutch or other West Germanic languages	6,708	0.09%
35	Greek	6,382	0.09%
36	Serbo-Croatian	6,333	0.09%
37	Polish	6,258	0.09%
38	Swahili or other languages of Central, Eastern, and Southern Africa	5,293	0.07%
39	Armenian	4,065	0.06%
40	Hmong	2,349	0.03%
41	Haitian	1,867	0.03%
42	Other Native languages of North America	1,637	0.02%
43	Navajo	149	0.00%

Sources: C01, C02, C03

Nine-county Bay Area. Residents age 5+. ACS 2020-2024. Point-estimate ordering. Complete variable IDs, count MOEs and English-ability splits are in the data store.

Inside the data store

Field or table	Meaning
geographies	Region, derived counties and constituent PUMAs
category_id	Original ACS estimate variable identifier
estimate / share	Count and percent of age-5+ geography total
moe90_approx	Approximate count margin of error, 90%
english_very_well	Count within non-English category
english_less_than_very_well	Combined English ability responses below very well
sources / findings	Source metadata; paragraph evidence links
school_records	EL-only aggregates, with geography and period
workplace_records	SF designated positions case, with source and period

Formats: JSON for the reader, SQLite for queries, and CSV exports for reuse. Source identifiers, retrieval dates and file hashes document the evidence trail. Dataset labels and null values are preserved.

Null English-ability values for “English only” mean the question is not applicable. They are not zero proficiency. Keep ACS, school and workplace tables separate when analyzing them.

Use the evidence. Keep the distinctions.

Suggested citation: Second Language Project. (2026). Languages of the Bay Area: Home, learning, work and community access. Research edition 1.1, 22 October 2026.

This edition was prepared through computer-assisted research, source extraction, analysis and publication production. It has not undergone independent academic peer review. Quantitative tables are reproducible from public sources; interpretations and recommendations remain open to community and subject-matter review.

The public reader includes the full text, searchable sources, geography filters, school and workplace evidence, and downloadable tables. Source references and methodological notes help readers examine how the findings were produced.

Print this reading-order PDF as a booklet using a printer's booklet setting, or use the separately imposed PDF. The imposed version is already arranged two pages per side: print actual size, landscape letter, duplex on the short edge. Do not apply booklet imposition a second time.

40 half-letter pages • 10 folded letter sheets

Print a two-sheet test before a full run. Fold, nest in sheet order and saddle staple. For professional printing, ask the printer to handle creep and finishing.

Continue the conversation.

Use this guide alongside local experience. Statistics can reveal patterns; people can explain what those patterns mean in their lives.

Explore the evidence

The online reader brings together the report, searchable sources, regional and local data, and separate school and workplace perspectives.

Share a question or request materials

Email the Second Language Project to ask about the guide or request a download key.

info@secondlanguageproject.org

secondlanguageproject.org

Many languages. Shared possibilities.

A guide to the languages people use at home, learn throughout life, and bring into work and community spaces.

Explore the Bay Area's languages, from regional patterns to the data and sources behind each finding.

For neighbors, educators, community organizers and anyone working toward more accessible conversations.



SECOND
LANGUAGE
PROJECT

LANGUAGE • CONNECTION • COMMUNITY

secondlanguageproject.org