Exploratory Analysis of English Learners' Identified Race/Ethnicity and Home Language in the School District of Philadelphia, 2019-20

Many staff in the Office of Evaluation, Research, and Accountability contributed to collecting, analyzing, and summarizing the information that appears in this brief. Major contributions were made by Theodore Wills, Ph.D., and Mark Lewis, Ph.D.

Key Findings

- As of the 2019-20 school year, Hispanic/Latinx English Learners (ELs) are the largest racial/ethnic groups of District ELs (55.2%) followed by Asian ELs (21.8%) and White ELs (12.7%).

- ELs classified as Hispanic/Latinx are linguistically homogeneous, with 92.5% classified as speaking Spanish, while other racial/ethnic classifications of ELs are linguistically more heterogeneous.

- When examining the racial/ethnic diversity of linguistic categories, two language groups of ELs, Arabic speakers and Portuguese speakers, stand out as particularly diverse across multiple racial/ethnic classifications.

Introduction

Reporting by racial/ethnic subgroup is common for accountability purposes in District-, Network- and school-wide analyses. This student group reporting is a crucial tool for investigating racial and ethnic disproportionalities, the progress of racial equity initiatives, and evidence of the impact of systemic racism on District students and families. Similarly, reporting and accountability are often disaggregated by English Learner (EL) status. Doing so supports efforts to serve students who require linguistic support, and also provides some indirect information about the diversity of students' cultures.

However, when student groups are identified in this way, it highlights differences *between* groups, but may also imply that the differences *within* those groups are unimportant, or even absent. This brief aims to surface some of the internal diversity within racial/ethnic groups, and within home language groups, by exploring the intersection of race/ethnicity and home language among our students.

There is, for example, considerable internal diversity within each of the District's standardized racial/ethnic categories. One dimension of that internal diversity is linguistic; that is, **students who share the same race/ethnicity might not share the same home language**. The commonly used racial/ethnic categories encapsulate diverse ethnolinguistic communities with potentially vastly different family backgrounds, immigration histories, and socioeconomic circumstances.

Perhaps even less attention is paid to the inverse statement that **students who share the same home language might not share the same race/ethnicity**. This is important, because assumptions (for example) that all speakers of Portuguese are likely to have similar cultural origins is challenged by the racial/ethnic diversity within this language group. It follows that efforts to identify, fund, and/or implement supports based on home language data alone may fail to meet the differentiated needs of our students.

This analysis investigates the linguistic diversity of racial/ethnic subgroups of ELs in SDP as a window on the internal diversity that characterizes these groups. These insights are critical to realizing the District's commitment to becoming a more culturally and linguistically responsive organization.

Exploratory Questions

Two exploratory research questions guided the analyses presented in this brief:

1. To what extent are racial/ethnic categories of ELs linguistically diverse?
2. To what extent are groups of ELs with the same home language racially/ethnically diverse?

Methods

This brief uses enrollment data from October of the 2019-20 school year to capture trends prior to any changes caused by the COVID-19 pandemic.[1] SDP home language data allows only one language to be recorded as a student's home language, so it does not capture families' multilingual practices. SDP race/ethnicity data is based on self-identification, and there is variation in how people self-identify with respect to the racial/ethnic options that are available in District data systems.

Findings

Home Languages of English Learners by Race/Ethnicity

Hispanic/Latinx ELs are the largest racial/ethnic group of ELs, followed by Asian and White ELs.

ELs classified as Hispanic/Latinx made up 55.2% of all SDP ELs in 2019-20 (Table 1). The next largest groups were Asian ELs (21.8% of District ELs) and White ELs (12.7% of District ELs). Black/African American ELs made up 7.3% of District ELs.

---

[1] Enrollment data was downloaded from the Enrollment - Oct 1 Snapshot Qlik App and was accessed 7/9/2021.

Table 1. Number of ELs classified in each major race/ethnicity in 2019-20 (n = 16,724)

Race/ethnicity	Number of ELs in 2019-20	Percent of all ELs in 2019-20
Asian	3,659	21.8%
Black/African American	1,214	7.3%
Hispanic/Latinx	9,245	55.2%
Multiracial/Other	490	2.9%
White	2,116	12.7%
All races/ethnicities	**16,724**	**100%**

Note: Multiracial/Other includes ELs classified as Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, and Unknown.

In the School District of Philadelphia, ELs classified as Hispanic/Latinx are linguistically homogeneous, but other racial/ethnic classifications of ELs are linguistically very heterogeneous.

ELs in SDP who are Hispanic/Latinx are a linguistically homogenous group, with 92.5% classified as speaking Spanish, another 6% classified as speaking Portuguese, and 1% as speaking English (Table 2). No other language group featured 100 or more ELs.

Table 2. Home languages spoken by Hispanic/Latinx ELs in SDP, 2019-20

Home Language	Total Number of EL Students with this Home Language	Number of Hispanic/ Latinx ELs with this Home Language	Number of Non-Hispanic/ Latinx ELs with this Home Language	% of Hispanic/ Latinx ELs with this Home Language
Spanish	8,752	8,555	197	92.7%
Portuguese	1,115	558	557	6.1%
English*	184	90	94	1.0%
All home languages with 100 or more ELs	**10,051**	**9,203**	**848**	**99.8%**[2]

* The category of EL students who speak English at home may be the result of how data about home language is collected. The home language variable only allows for a single entry per household. Thus, it is likely, but impossible to verify, that some of the 184 EL students listed here with a home language of English have multilingual home environments, and/or that English is a recent, or non-dominant component of their household's communication practices.

[2] There were an additional 17 Hispanic/Latinx-identifying ELs with Mayan languages for a home language, but Campbell-Montalvo (2021) suggests school districts regularly undercount Indigenous languages spoken by families classified as Hispanic/Latinx. As noted previously, SDP records only allow the recording of a single home language, so families who speak both Spanish and an Indigenous language would likely be recorded as Spanish-speaking. Campbell-Montalvo, R. (2021). Linguistic re-formation in Florida Heartland schools: School erasures of Indigenous Latino languages. _American Educational Research Journal, 58_(1), 32–67.

In contrast, ELs in SDP who are Asian are very linguistically diverse. About a third (35.6%) speak Mandarin, the most common home language of this group (Table 3). Notably, nine additional language groups had at least 100 members.[3]

Table 3. Home languages spoken by Asian ELs in SDP, 2019-20, Minimum 100 ELs

Home Language	Total Number of EL Students with this Home Language	Number of Asian ELs with this Home Language	Number of non-Asian ELs with this Home Language	% of Asian ELs with this Home Language
Chinese (Mandarin)	1,314	1,303	11	35.6%
Vietnamese	395	377	18	10.3%
Khmer	324	301	23	8.2%
Bengali	204	189	15	5.2%
Russian	651	160	491	4.4%
Chinese (Yue/Cantonese)	147	145	2	4.0%
Nepali	138	131	7	3.6%
Uzbek	238	130	108	3.6%
Pashto	163	128	35	3.5%
Malayalam	121	114	7	3.1%
Arabic	795	107	688	2.9%
All home languages with 100 or more ELs	**4,490**	**3,085**	**1,405**	**84.3%***

* There were an additional 575 Asian-identifying ELs, speaking a total of 44 additional home languages.

---

[3] Also note that for some language groups (notably Russian, Uzbek, and Arabic), there are many speakers who do not identify as Asian. This issue is explored more deeply in the second part of this brief.

ELs who are White non-Hispanic/Latinx have more linguistic diversity than Hispanic/Latinx ELs and less diversity than Asian ELs. Between 19.6% and 22.6% speak the most common languages of White ELs—Russian, Portuguese, and Arabic (Table 4; see Appendix A for additional tables for other race/ethnicity groups).

Table 4. Home languages spoken by White ELs in SDP, 2019-20

Home Language	Total Number of EL Students with this Home Language	Number of White ELs with this Home Language	Number of non-White ELs with this Home Language	% of White ELs with this Home Language
Russian	651	479	172	22.6%
Portuguese	1,115	461	654	21.8%
Arabic	795	415	380	19.6%
Albanian	150	144	6	6.8%
Ukrainian	137	134	3	6.3%
Uzbek	238	103	135	4.9%
Georgian	69	69	0	3.3%
Tajik	117	57	60	2.7%
Spanish	8,752	55	8,697	2.6%
Turkish	38	33	5	1.6%
All home languages with 100 or more ELs	**12,062**	**1,950**	**10,112**	**92.2%***

* There were an additional 166 White-identifying ELs, speaking a total of 39 additional home languages.

Race/Ethnicity of English Learners by Home Language

There are 10 home languages that accounted for at least one percent of all SDP ELs in 2019-20 (Table 5). The largest, Spanish, accounts for more than half of all ELs, while no other single language exceeds 10 percent.

Table 5. The 10 languages most commonly spoken by ELs, 2019-20 (n = 16,724)

Home Language	Number of ELs with this Home Language	Percent of ELs with this Home Language
Spanish	8,752	52.3%
Chinese (Mandarin)	1,314	7.9%
Portuguese	1,115	6.7%
Arabic	795	4.8%
Russian	651	3.9%
Vietnamese	395	2.4%
Khmer	324	1.9%
Uzbek	238	1.4%
French	228	1.4%
Bengali	204	1.2%

Portuguese- and Arabic-speaking ELs belong to multiple racial/ethnic classifications

In a parallel analysis to the *linguistic* diversity of a *racial/ethnic* category, it is also possible to examine the *racial/ethnic* diversity of a *linguistic* category. Two large language groups of ELs, Portuguese speakers and Arabic speakers, stand out as particularly diverse across multiple racial/ethnic classifications. It is important that analyses of EL students consider the intersection of racial/ethnic categories and home languages so that the heterogeneity of these populations is not overlooked. This is especially relevant in the District, as Portuguese and Arabic were the third- and fourth-most common languages spoken by ELs in 2019-20, respectively (Table 5, above).

Portuguese speakers are about evenly split between being classified as Hispanic/Latinx (50.0%) and being classified as white (non-Hispanic/Latinx) (41.3%, Table 6), with smaller representation among other races/ethnicities.

Table 6. Racial/ethnic classifications of Portuguese-speaking ELs, 2019-20 (n = 1,115)

Race/Ethnicity	Number of Portuguese speakers in this racial/ethnic group	% of Portuguese speakers in this racial/ethnic group
Asian	6	0.5%
Black/African American	29	2.6%
Hispanic/Latinx	558	50.0%
Multiracial/Other	61	5.0%
White	461	41.3%
All races/ethnicities	**1,115**	**100%**

Note: Multiracial/Other includes ELs classified as Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, and Unknown.

Arabic speakers are particularly diverse in their classification, with 52.2% classified as white, 25.6% classified as Multiracial/Other, and 13.4% classified as Asian (Table 7).

Table 7. Racial/ethnic classifications of Arabic-speaking ELs, 2019-20 (n = 795)

Race/Ethnicity	Number of Arabic speakers in this racial/ethnic group	% of Arabic speakers in this racial/ethnic group
Asian	107	13.4%
Black/African American	69	8.6%
Hispanic/Latinx	0	0%
Multiracial/Other	204	25.6%
White	415	52.2%
All races/ethnicities	**795**	**100%**

Note: Multiracial/Other includes ELs classified as Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, and Unknown.

Conclusions

The racial/ethnic categories typically used in reporting and accountability are intended to describe meaningful cultural groups within the United States. In the School District of Philadelphia, the categories alone do not capture the intersection of race/ethnicity and home language.

Subsequent analyses, in support of District equity goals, might further explore how home language communities differ in terms of academic and climate outcomes, On/Off/Near-Track status of schools attended by students, or other measures whose analysis would help reveal educational opportunities for the many linguistic communities from which English Learners come to attend the School District of Philadelphia.

Appendix A. Additional Home Language Groups by Race/Ethnicity Tables

Table A1. Languages spoken by Black/African-American ELs, 2019-20

Language	Number of Black/African American ELs who are speakers	Number of non-Black/African American ELs who are speakers	% of Black/African American ELs who are speakers
French	210	18	17.3%
Haitian Creole	180	3	14.8%
Creoles & Pidgins, Eng.-Based (Other)	118	13	9.7%
Creoles & Pidgins, French-Based (Other)	111	5	9.1%
Mandingo	75	3	6.2%
Fulah	70	4	5.8%
Arabic	69	726	5.7%
Swahili	60	11	4.9%
Spanish	54	8,698	4.4%
Portuguese	29	1,086	2.4%
45 additional languages	45	NA	19.7%
All languages	**1,214**	**NA**	**100%**

Table A2. Languages spoken by Multi Racial/Other ELs, 2019-20

Language	Number of Multiracial/Other ELs who are speakers	Number of non-Multiracial/Other ELs who are speakers	% of Multiracial/Other ELs who are speakers
Arabic	204	591	41.8%
Spanish	72	8,680	14.8%
Portuguese	61	1,054	12.5%
Pashto	19	144	3.9%
Vietnamese	11	384	2.3%
Russian	10	641	2.0%
Swahili	11	60	2.3%
Creoles & Pidgins, Eng.-Based (Other)	8	123	1.6%
Khmer	8	316	1.6%
English	7	177	1.4%
35 additional languages	45	NA	15.8
All languages	**490**	**NA**	**100%**

Note: Multiracial/Other includes ELs classified as Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, and Unknown.

Appendix B. Linguistic diversity measure for all races/ethnicities from 2015-16 to 2019-20

The linguistic diversity indicator (LDI) is a metric that provides one way to summarize a single impression of the linguistic diversity of a school's population of ELs.[4] In this appendix, we extend the LDI to summarize other populations of ELs, in this case the linguistic diversity of students of a specific race/ethnicity. The measure identifies the number of home languages required to account for 80% of ELs in a given school or population.[5]

How to read this table: For each year and each race/ethnicity group, the table provides the number of ELs. In addition, the last two columns provide that group's Linguistic Diversity score, and the specific languages that are represented in that score. For example, in 2014-15 there were 3,134 Asian ELs in the District. With those 3,134 Asian ELs, the largest language group was Mandarin, accounting for 24% of the total. In order to account for 80% of the total, it would be necessary to compile all of the EL's from the top eight language groups, plus most of the ninth largest group (resulting in the value of 8.86 for the Linguistic Diversity metric). Those nine languages are all listed in the last column.

School Year	Race/ethnicity	Number of ELs	Linguistic diversity measure	Languages spoken by 80% of this group (i.e., those included in linguistic diversity measure)
2014-15	Asian	3,134	8.86	Chinese (Mandarin) (24%), Vietnamese (14%), Khmer (12%), Undetermined (8%), Nepali (8%), Chinese (Yue/Cantonese) (6%), Pashto (4%), Malayalam (4%), Burmese (2%)
2014-15	Black/African American	1,162	9.49	Creoles and Pidgins, English-Based (Other) (20%), French (18%), Mandingo (8%), English (8%), Undetermined (6%), Spanish (6%), Fulah (6%), Arabic (4%), Kinyarwanda (4%), Haitian Creole (2%)
2014-15	Hispanic/Latinx	6,486	0.84	Spanish (94%)
2014-15	Multiracial/ Other	777	8.28	Arabic (42%), Spanish (22%), Creoles and Pidgins, English-Based (Other) (4%), Portuguese (2%), Undetermined (2%), Pashto (2%), Urdu (2%), Vietnamese (2%), Chinese (Mandarin) (2%)
2014-15	White	1,010	7.73	Arabic (24%), Russian (20%), Ukrainian (10%), Albanian (8%), Uzbek (6%), Portuguese (4%), Spanish (4%), Tajik (4%)

---

[4] As a result of a technology system vendor's limitations, the online enrollment form used in SDP uses a list of 331 languages derived from the ISO 639-2, a list of languages maintained by the Library of Congress for bibliographic purposes: https://www.loc.gov/standards/iso639-2/php/code_list.php Thus, the SDP list includes erroneous possibilities such as ancient or extinct languages and language families. However, these errors account for a very small number of students, so they do not affect the linguistic diversity measure described here, which only examines the subset of students speaking the most common languages. There are not standardized or widely used procedures for correcting a student's home language on record.

[5] For a full description of this metric, see: https://www.philasd.org/research/wp-content/uploads/sites/90/2022/04/Linguistic-Diversity-Indicator-April-2022.pdf

School Year	Race/ethnicity	Number of ELs	Linguistic diversity measure	Languages spoken by 80% of this group (i.e., those included in linguistic diversity measure)
2015-16	Asian	3,108	9.35	Chinese (Mandarin) (26%), Vietnamese (14%), Khmer (12%), Nepali (6%), Undetermined (6%), Chinese (Yue/Cantonese) (6%), Pashto (4%), Malayalam (4%), Bengali (4%), Burmese (2%)
2015-16	Black/African American	1,199	10.51	French (18%), Creoles and Pidgins, English-Based (Other) (18%), Mandingo (8%), Arabic (6%), Undetermined (6%), English (6%), Spanish (6%), Fulah (4%), Haitian Creole (4%), Creoles and Pidgins, French-Based (Other) (4%), Amharic (2%)
2015-16	Hispanic/Latinx	6,860	0.85	Spanish (94%)
2015-16	Multiracial/ Other	851	7.84	Arabic (38%), Spanish (22%), Portuguese (4%), Undetermined (4%), Creoles and Pidgins, English-Based (Other) (4%), Pashto (4%), Vietnamese (2%), Bengali (2%)
2015-16	White	1,130	7.92	Arabic (22%), Russian (20%), Ukrainian (10%), Albanian (8%), Portuguese (6%), Uzbek (6%), Spanish (4%), Georgian (4%)
2016-17	Asian	3,149	9.3	Chinese (Mandarin) (30%), Vietnamese (12%), Khmer (10%), Chinese (Yue/Cantonese) (6%), Nepali (6%), Bengali (4%), Pashto (4%), Malayalam (4%), Burmese (2%), Urdu (2%)
2016-17	Black/African American	1,213	9.8	French (16%), Creoles and Pidgins, English-Based (Other) (16%), Mandingo (8%), Haitian Creole (8%), Spanish (6%), Arabic (6%), English (6%), Creoles and Pidgins, French-Based (Other) (6%), Swahili (6%), Fulah (4%)
2016-17	Hispanic/Latinx	7,349	0.84	Spanish (94%)
2016-17	Multiracial/ Other	804	8.17	Arabic (36%), Spanish (26%), Portuguese (6%), Pashto (4%), Creoles and Pidgins, English-Based (Other) (2%), Chinese (Mandarin) (2%), Bengali (2%), English (2%), Vietnamese (2%)
2016-17	White	1,290	6.89	Arabic (26%), Russian (20%), Ukrainian (10%), Albanian (8%), Portuguese (6%), Spanish (6%), Uzbek (4%)
2017-18	Asian	3,314	9.06	Chinese (Mandarin) (30%), Vietnamese (12%), Khmer (10%), English (6%), Chinese (Yue/Cantonese) (6%), Bengali (4%), Nepali (4%), Pashto (4%), Malayalam (4%), Arabic (2%)
2017-18	Black/African American	1,314	9.5	French (14%), Creoles and Pidgins, English-Based (Other) (12%), English (12%), Haitian Creole (8%), Mandingo (6%), Creoles and Pidgins, French-Based (Other) (6%), Arabic (6%), Swahili (6%), Spanish (6%), Fulah (4%)
2017-18	Hispanic/Latinx	7,593	0.89	Spanish (90%)

School Year	Race/ethnicity	Number of ELs	Linguistic diversity measure	Languages spoken by 80% of this group (i.e., those included in linguistic diversity measure)
2017-18	Multiracial/Other	842	7.2	Arabic (38%), Spanish (22%), Portuguese (8%), English (4%), Pashto (4%), Creoles and Pidgins, English-Based (Other) (2%), Persian (2%), Swahili (2%)
2017-18	White	1,705	7.61	Arabic (20%), Russian (20%), English (10%), Ukrainian (8%), Spanish (6%), Portuguese (6%), Albanian (6%), Uzbek (4%)
2018-19	Asian	3,580	9.55	Chinese (Mandarin) (34%), Vietnamese (12%), Khmer (10%), Bengali (6%), Chinese (Yue/Cantonese) (4%), Nepali (4%), Pashto (4%), Russian (4%), Malayalam (4%), Arabic (2%)
2018-19	Black/African American	1,333	9.46	French (16%), Haitian Creole (12%), Creoles and Pidgins, English-Based (Other) (12%), Creoles and Pidgins, French-Based (Other) (8%), Mandingo (6%), Arabic (6%), Swahili (6%), English (4%), Spanish (4%), Fulah (4%)
2018-19	Hispanic/Latinx	8,418	0.87	Spanish (92%)
2018-19	Multiracial/Other	680	6.86	Arabic (36%), Spanish (22%), Portuguese (10%), English (4%), Pashto (2%), Creoles and Pidgins, English-Based (Other) (2%), Swahili (2%)
2018-19	White	1,851	6.68	Russian (26%), Arabic (24%), Portuguese (10%), Albanian (6%), Ukrainian (6%), Uzbek (4%), Spanish (4%)
2019-20	Asian	3,659	9.55	Chinese (Mandarin) (36%), Vietnamese (10%), Khmer (8%), Bengali (6%), Russian (4%), Chinese (Yue/Cantonese) (4%), Nepali (4%), Uzbek (4%), Pashto (4%), Malayalam (4%)
2019-20	Black/African American	1,214	9.83	French (18%), Haitian Creole (14%), Creoles and Pidgins, English-Based (Other) (10%), Creoles and Pidgins, French-Based (Other) (10%), Mandingo (6%), Fulah (6%), Arabic (6%), Swahili (4%), Spanish (4%), Portuguese (2%)
2019-20	Hispanic/Latinx	9,245	0.86	Spanish (92%)
2019-20	Multiracial/Other	490	7.44	Arabic (42%), Spanish (14%), Portuguese (12%), Pashto (4%), Swahili (2%), Vietnamese (2%), Russian (2%), English (2%)
2019-20	White	2,116	5.58	Russian (22%), Portuguese (22%), Arabic (20%), Albanian (6%), Ukrainian (6%), Uzbek (4%)

Note: Multiracial/Other includes ELs classified as Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, and Unknown.