Census 2000 Evaluation A.5.a
September 30, 2003

Analysis of the Linguistically Isolated Population in Census 2000

FINAL REPORT

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Frederic Allen Lestina, Jr.
Planning, Research and Evaluation Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

CONTENTS

List of Tables

Executive Summary

This evaluation has four objectives which are the following:

- to create a universe of households identified as linguistically isolated
- to determine how they were enumerated in Census 2000
- to examine the education attainment of the householder
- to examine geographic clustering at the tract and county levels

The enumeration of each household involves any of the following operations: paper mailback questionnaires, Internet responses, Be Counted Questionnaires, Telephone Questionnaire Assistance interviews, Coverage Edit Followup interviews, List/Enumerate, Update/Enumerate, Nonresponse Followup, and Coverage Improvement Followup. A household is classified as linguistically isolated if all household members age 14 years or older speak a language other than English and have limited English proficiency. Earlier studies have found that the inability to speak English well was a barrier to effective enumeration (Bruce and Robinson, 1999). In responding to legislation, specifically the Civil Rights Act of 1964, the Bilingual Education Act, and the Voting Rights Act, using census data, the Census Bureau developed an approach to identify a language spoken at home and measure self-reported English language ability (Siegel, Martin, and Bruno, 2000). "Linguistic Isolation" was defined by these measurements. This calculation revealed that the percent of linguistically isolated households had increased since 1990, from 3.2 percent to 4.1 percent (SE = 0.0049 percent). The evaluation also found that linguistically isolated households were less likely to self-respond than the non-linguistically isolated households. These self-responses included paper mailback questionnaires, Internet responses, Telephone Questionnaire Assistance interviews, and Coverage Edit Followup interviews. For linguistically isolated households, 61.0 percent (SE = 0.0636 percent) were self-responses. For non-linguistically isolated households, 72.6 percent (SE = 0.0112 percent) were self-responses.

There are 3,141 counties in the nation with at least one linguistically isolated household. Each of these has up to 35.0 percent of its households that are linguistically isolated. Of the 3,141 counties, 91.53 percent have less than five percent of their households that are linguistically isolated. There are eight counties in Texas with at least 25 percent of their households that are linguistically isolated.

Of the 64,960 tracts in the nation with at least one linguistically isolated household, 77.5 percent have less than five percent of their households that are linguistically isolated. There are 11 tracts in the nation where at least 75 percent of their households are linguistically isolated. They are as follow: one in Maricopa County, Arizona, one in Pinal County, Arizona, three in Los Angeles County, California, one in San Francisco County, California, one in Lafourche Parish, Louisiana, two in Bronx County, New York, one in Dutchess County, New York, and one in Charleston County, South Carolina.

Finally, householders in linguistically isolated households are less likely to have formal education beyond grade 12 than those in non-linguistically isolated households.

At the tract and county levels, the linguistic isolation variable may help with identifying areas for special enumeration procedures including language programs for the 2010 Census. Further analysis should be done by specific languages that are spoken at home in order to identify the level and if they are clustered.

1. BACKGROUND

This evaluation identifies non-English speaking households at the state and national level as well as the tract and county level. These non-English speaking households are known as linguistically isolated. The objectives of this evaluation are the following: (1) to create a universe of households identified as linguistically isolated; (2) to determine how they were enumerated in Census 2000; (3) to examine the education attainment of the householder; and (4) to examine the geographic clustering at the tract and county levels. This report does not include analysis by language group.

1.1 1990 Census

In 1990, a language other than English was spoken at home in 15.5 percent of all households in the country and 20 percent of those households or 3.2 percent of all households were linguistically isolated (Siegel et. al., 2000). This research raised some issues surrounding the validity of interpretations of linguistic isolation of households. In addition, the report suggested further investigation into the justification of applying the definition of linguistic isolation in social programs and in survey administrations. For instance, sometimes it does not make sense if the level of English proficiency of the household is only tied to members of the household 14 years or older no matter how proficient in English persons who are younger than 14 years. The evaluation cited that there was much evidence dealing with non-sampling error in the measurement of the elements of linguistic isolation. It cited that there could be shortcomings about the level of uncertainty in estimates for small areas arising from sampling variation in the Census and other surveys including American Community Survey. Such examples would be areas with high rates of Hispanic migration and political or social indifference to language needs. The concept of linguistic isolation is rather complex and needs to be examined closely in related evaluation papers.

Bruce and Robinson (1999) showed that linguistically isolated households in 1990 were very concentrated geographically. In addition, their results showed that for linguistically isolated households living in hard-to-enumerate areas, no language other than Spanish occurred with any frequency, and that national tabulations masked important differences in the language needs of individual states. This fact necessitates analysis of the results for each state as well as for the nation.

Appendix A contains some key words of this study and their definition.

1.2 Census 2000

For Census 2000, the linguistic isolation variable was determined at the household level. It is found in the Sample Edited Data File (SEDF). It is a recode variable based on other variables that checks on the following for all household members 14 years or older: the language spoken at home and the level of English proficiency.

It serves as a flag in the household level data indicating whether the household is linguistically isolated or not.

2. METHODS

The evaluation is based on data from the SEDF, updated April 2002. The SEDF contains 100 percent and sample data which have been edited and imputed for all persons in sample housing units. Contained on the SEDF is the linguistic isolation flag variable which indicates whether the household is linguistically isolated or not.

2.1 Source of Data

The evaluation uses the following household level data from the SEDF:

- household linguistic isolation flag
- return source
- housing unit weight

The return source data identify the operation in which the household was enumerated during Census 2000. The evaluation uses only one person level variable which defines the language spoken at home for each person. Since the linguistic isolation data were derived from the long form survey data, the universe in this study only contained households enumerated by the long form. Appendix B in this paper has information on the variables that are used in this study as well as all the values for each of the variables.

2.2 Estimation and Weighting

The data are weighted using the sample housing unit weight variable (HWT). The data are weighted since the analysis is based on the long form sample data. The values of the return source (RSOURCE) are collapsed into return types which define the operation in which the households were enumerated. The unweighted total of all households in return type j is n_j where

$$j = \begin{cases} 1 & \text{for paper mailback}^{1} \\ 2 & \text{for coverage edit followup} \\ 3 & \text{for nonresponse followup} \\ 4 & \text{for coverage improvement followup} \\ 5 & \text{for personal visit enumeration}^{2} \\ 6 & \text{for others}^{3} \end{cases}$$

The weighted total number of households is *A*.

$$A = \sum_{j=1}^{6} \sum_{i=1}^{n_j} hwt_i$$

where hwt_i is the housing unit weight for the household i in return type j and $i = 1,2,\ldots,n_j$.

The weighted total is equal to or greater than the unweighted total. The weighted total number of the linguistically isolated households is *B*.

$$B = \sum_{j=1}^{6} \sum_{i=1}^{n_j} hwt_i * LI_i$$

where hwt_i is the housing unit weight for the household i for return type j and

$$LI_i = \begin{cases} 0 & \text{if the household } i \text{ is not linguistically isolated} \\ 1 & \text{if the household } i \text{ is linguistically isolated} \end{cases}$$

The weighted total number of the non-linguistically isolated households is *C*.

$$C = \sum_{j=1}^{6} \sum_{i=1}^{n_j} hwt_i * nonLI_i$$

---

[1] including paper mailback questionnaires from mailout, Internet responses, Telephone Questionnaire Assistance (TQA) interviews, paper mailback questionnaires in Update/Leave areas, paper mailback questionnaires in Urban Update/Leave areas, foreign language questionnaires, and Be Counted Forms (BCF).

[2] including List/Enumerate, Update/Enumerate, and Remote Alaska.

[3] including T-Night, Group Quarters, Military Group Quarters, and Shipboard Group Quarters.

where hwt_i is the housing unit weight for the household *i* for return type j and

$$nonLI_i = \begin{cases} 1 & \text{if household } i \text{ is not linguistically isolated} \\ 0 & \text{if household } i \text{ is linguistically isolated} \end{cases}$$

The percent of linguistically isolated households is *D*.

$$D = \left[\sum_{j=1}^{6} \sum_{i=1}^{n_j} (hwt_i * LI_i) / \sum_{j=1}^{6} \sum_{i=1}^{n_j} hwt_i \right] * 100$$

The percent of non-linguistically isolated households is *E*.

$$E = \left[\sum_{j=1}^{6} \sum_{i=1}^{n_j} (hwt_i * nonLI_i) / \sum_{j=1}^{6} \sum_{i=1}^{n_j} hwt_i \right] * 100$$

The percent of linguistically isolated households in return type j is F_j.

$$F_j = \left[\sum_{i=1}^{n_j} (hwt_i * LI_i) / \sum_{i=1}^{n_j} hwt_i \right] * 100$$

The percent of non-linguistically isolated households in return type j is G_j.

$$G_j = \left[\sum_{i=1}^{n_j} (hwt_i * nonLI_i) / \sum_{i=1}^{n_j} hwt_i \right] * 100$$

2.3 Linguistically isolated households at the county and tract levels

For calculating the percent of the linguistically isolated households at the county level, households were grouped by the state j and county k. The unweighted total number of all households in the state j and county k is n_{jk} where

$$j = 1,2,..,51$$

and

$$k = 1,2, ..., m_j$$

where m_j is the number of counties in state j. For the state j and county k, the weighted total number of all households is *R*.

$$R = \sum_{i=1}^{n_{jk}} hwt_i$$

where hwt_i is the housing unit weight for the household i, for $i = 1,2,\ldots,n_{jk}$.

The weighted total number of the linguistically isolated households is S.

$$S = \sum_{i=1}^{n_{jk}} hwt_i * LI_i$$

where hwt_i is the housing unit weight for the household i, for $i = 1,2,\ldots,n_{jk}$ and

$$LI_i = \begin{cases} 0 & \text{if the household } i \text{ is not linguistically isolated} \\ 1 & \text{if the household } i \text{ is linguistically isolated} \end{cases}$$

The percent of linguistically isolated households for the county level is T_{jk}.

$$T_{jk} = \left[\sum_{i=1}^{n_{jk}} (hwt_i * LI_i) / \sum_{i=1}^{n_{jk}} hwt_i \right] * 100$$

The procedures for calculating the percent of the linguistically isolated households are similar for tract geography. However, both state and county are needed to uniquely identify the tract l. Hence, households were grouped by the state j, county k, and tract l. The unweighted total number of all households in the state j, county k, and tract l is n_{jkl} where

$$j = 1,2,..,51$$

and

$$k = 1,2, ..., m_j$$

and

$$l = 1,2,..., p_{kj}$$

where m_j is the number of counties in state j and p_{kj} is the number of tracts in county k of state j. For the state j county k and tract l the weighted total number of all households is *U*.

$$U = \sum_{i=1}^{n_{jkl}} hwt_i$$

where hwt_i is the housing unit weight for the household i, for $i = 1,2,...,n_{jkl}$.

The weighted total number of the linguistically isolated households is *V*.

$$V = \sum_{i=1}^{n_{jkl}} hwt_i * LI_i$$

where hwt_i is the housing unit weight for the household i, for $i = 1,2,...,n_{jkl}$, and

$$LI_i = \begin{cases} 0 & \text{if the household } i \text{ is not linguistically isolated} \\ 1 & \text{if the household } i \text{ is linguistically isolated} \end{cases}$$

The percent of the linguistically isolated households for the tract level is W_{jkl}.

$$W_{jkl} = \left[\sum_{i=1}^{n_{jkl}} (hwt_i * LI_i) / \sum_{i=1}^{n_{jkl}} hwt_i \right] * 100$$

2.4 Applying Quality Assurance Procedures

Quality Assurance procedures were applied to the design, implementation, analysis and preparation of this report. A description of the procedures used is provided in the "Census 2000 Evaluation Program Quality Assurance Process".

3. LIMITS

3.1 Nonsampling error in measurements in small areas

There has been some evidence of difficulty in estimating error for small areas regarding the measure of the elements of linguistic isolation (Siegel et. al., 2000). Therefore, the measurements are subject to substantial variation at fine geographic levels, like tracts.

3.2 Validity of the definition of "linguistic isolation"

There have been some concerns about the validity of the definition of "linguistic isolation" which may need to be studied further to clarify its application in social programs and in survey administrations. Social programs and survey administrations may have different standards in identifying linguistically isolated households. For instance, some of them identify linguistically isolated households where no household member 14 years or older speaks English "very well" even though someone younger than 14 years in the household may speak English "very well". The others allow the level of English proficiency to be applied to any household member younger than 14 years in identifying linguistic isolated households.

In some areas where bilingualism is official in governmental services, like Dade County, Florida where English and Spanish are both official languages, there might be different requirements in defining linguistically isolated households. In some cases, households where a language other than English is spoken and where no household member speaks English "very well" might not be classified as linguistically isolated because the household language is also an official language in the area. There are political and ideological arguments over whether communication in English is or should be adequate for government programs, but the importance of language as a barrier to survey administration needs direct assessment.

3.3 Influx of migrants and population mobility

There may be many linguistically isolated households that are unaccounted for in some geographical areas in the nation (Siegel et. al., 2000). These household respondents have not resided in the area at the time of enumeration and are likely to migrate to different areas in a short time. Some of them could be reluctant to be enumerated because of fears that they might be reported to law enforcement agencies that deal with undocumented immigrants.

3.4 Sampling techniques of linguistically isolated household data

The universe of the households in this study is based on sample data from the Census 2000 long form sample. The results of the analysis on linguistic isolation are subject to sampling variability. The linguistic isolation data are only available in the SEDF which has edited and imputed data.

4. RESULTS

4.1 What was the percent of all households that are linguistically isolated?

A linguistically isolated household is one in which all persons age 14 years or older who speak a language other than English do not speak English "very well". The percent of linguistically isolated households has increased due to new immigrants in recent years. In 1990, 3.2 percent of all households were classified as linguistically isolated. Spanish, Tagalog, Chinese, Korean and

Vietnamese-speaking households represent the fastest growing non-English speaking components of this linguistic isolation universe (Siegel et. al., 2000). These new immigrants with limited English proficiency have been thought to have a language barrier that affects any Census survey of households. For Census 2000, a language assistance program was developed to alleviate this problem. Table 1 shows the overall weighted percent of linguistically isolated households and the overall weighted percent of non-linguistically isolated households, both at the national level. The percent of all households that are linguistically isolated is 4.1 percent. The table also shows the percentage distribution of the linguistically isolated and non-linguistically isolated households across two response modes, self-response and personal visit response. The self-response mode refers to paper mailback questionnaires, Internet responses, Be Counted Forms, Telephone Questionnaire Assistance interviews, and Coverage Edit Followup interviews. The personal visit response mode refers to information filled out by the enumerator from any of the following operations: List/Enumerate, Update/Enumerate, Nonresponse Followup, and Coverage Improvement Followup. There is a small number of linguistically isolated and non-linguistically isolated households that could not be classified into self-response or personal visit response so they are not included in the table.

Table 1. Percent of Linguistic Isolation by Mode - National level Weighted

		Mode	
Linguistic Isolation	**Total**	**Self-Response**	**Personal Visit Response**
Total			
Column Percent	**100.0**	**100.0**	**100.0**
Row Percent	**100.0**	**72.1 (0.0111)**	**27.9 (0.0111)**
Linguistically Isolated			
Column Percent	4.1 (0.0049)	3.5 (0.0053)	5.8 (0.0110)
Row Percent	100.0	61.0 (0.0636)	39.0 (0.0636)
Not linguistically isolated			
Column Percent	95.9 (0.0049)	96.5 (0.0053)	94.2 (0.0110)
Row Percent	100.0	72.6 (0.0112)	27.4 (0.0112)

Data Source: SEDF ( ) denotes standard error

The table shows that, for Census 2000, 4.1 percent (SE = 0.0049 percent) of all households were classified as linguistically isolated. Linguistically isolated households were less likely to self-respond than non-linguistically isolated households, 61.0 percent (SE = 0.0636 percent) to 72.6 percent (SE = 0.0112 percent), respectively. Thus, linguistically isolated households were more likely to have personal visit response than non-linguistically isolated households, 39.0 percent (SE = 0.0636 percent) to 27.4 percent (SE = 0.0112 percent), respectively.

The table in Appendix C shows the rank of the states by the percent of linguistically isolated households. At the state level, the percent of linguistic isolation ranged from a high of 9.65 percent (SE = 0.0240 percent) for California to a low of 0.32 percent (SE = 0.0154 percent) for West Virginia. The five states with the highest percent of linguistically isolated households are the following: California, 9.65 percent (SE = 0.0240 percent); New York, 7.74 percent (SE = 0.0263 percent); Texas, 7.22 percent (SE = 0.0246 percent); Hawaii, 7.00 percent (SE = 0.1018 percent); and New Mexico, 6.49 percent (SE = 0.0768 percent). The five states with the lowest percent of linguistically isolated households are the following: Kentucky, 0.73 percent (SE = 0.0165 percent); Alabama, 0.70 percent (SE = 0.0160 percent); Montana, 0.68 percent (SE = 0.0286 percent); Mississippi, 0.54 percent (SE = 0.0175 percent); and West Virginia, 0.32 percent (SE = 0.0154 percent).

4.2 How were Linguistically Isolated households enumerated in Census 2000?

Table 2 has national level data, both estimates and percentages, for the linguistically isolated households and non-linguistically isolated households by the types of return. Return type includes the following: paper mailback[1], Coverage Edit Followup, Nonresponse Followup, Coverage Improvement Followup, and Personal Visit Enumeration[2].

The table shows that the return type with the largest percent of both linguistically isolated households and non-linguistically isolated households, is paper mailback, 57.7 percent (SE = 0.0644 percent) and 71.2 percent (SE = 0.0114 percent), respectively. Nonresponse Followup (NRFU) is second to the paper for both linguistically isolated and non-linguistically isolated, 35.5 percent (SE = 0.0624 percent) and 24.9 percent (SE = 0.0109 percent), respectively. As expected, these are the two biggest operations.

Of all households enumerated by Coverage Edit Followup, 9.0 percent (SE = 0.0598 percent) are linguistically isolated; Nonresponse Followup, 5.8 percent (SE = 0.0117 percent); Coverage Improvement Followup, 5.5 percent (SE = 0.0414 percent); Personal Visit Enumeration, 5.9 percent (SE = 0.0507 percent). The reason for Coverage Edit Followup having the largest percent of linguistically isolated households could be due to the large household followup component. For that component, these cases were originally mailback self-response enumerations that contained more than six persons. During Coverage Edit Followup we contacted these households to collect data for Persons 7 and higher.

---

[1] including paper mailback questionnaires from mailout, Internet responses, Telephone Questionnaire Assistance (TQA) interviews, paper mailback questionnaires in Update/Leave areas, paper mailback questionnaires in Urban Update/Leave areas, foreign language questionnaires, and Be Counted Forms (BCF).

[2] including List/Enumerate, Update/Enumerate, and Remote Alaska.

Table 2. Estimates and Percent of Return Type by Linguistic Isolation - National level Weighted

	Linguistically Isolated Households			Non Linguistically Isolated Households		
		percent			percent	
Return Type	**estimates[4]**	**row**	**column**	**estimates[4]**	**row**	**column**
Total	**4,362,318**	**4.1 (0.0049)**	**100.0**	**101,117,783**	**95.9 (0.0049)**	**100.0**
Paper Mailback[1]	2,523,837	3.4 (0.0053)	57.7 (0.0644)	72,004,826	96.6 (0.0053)	71.2 (0.0114)
Coverage Edit Followup	134,396	9.0 (0.0598)	3.1 (0.0226)	1,358,364	91.0 (0.0598)	1.3 (0.0028)
Nonresponse Followup	1,549,925	5.8 (0.0117)	35.5 (0.0624)	25,147,460	94.2 (0.0117)	24.9 (0.0109)
Coverage Improvement Followup	104,291	5.5 (0.0414)	2.4 (0.0200)	1,801,675	94.5 (0.0414)	1.8 (0.0033)
Personal Visit Enumeration[2]	48,792	5.9 (0.0507)	1.2 (0.0142)	785,499	94.2 (0.0507)	0.8 (0.0022)
Other[3]	1,077	5.1 (1.1984)	0.0	19,959	94.9 (1.1984)	0.0

Data Source: SEDF ( ) denotes standard error

4.3 Where were the households located geographically?

Table 3 shows the distribution of tract by the percent of households within the tract that are linguistically isolated. The first column contains percent ranges from 0.0 percent to less than 5.0 percent, 5.0 percent to less than 10.0 percent and so on. The estimates column shows the number of tracts having the percent of linguistically isolated households within the range. The percentages in the percent column represent the percent of tracts that are within the range of linguistically isolated. The table indicates for 77.53 percent of all tracts, less than five percent of their households are linguistically isolated. Assuming there is no substantial clustering of the linguistically isolated households, this is expected given that nationally 4.1 percent of all households are linguistically isolated. There are 11 tracts where at least 75 percent of their households are linguistically isolated. These represent 0.02 percent of all tracts.

[1] including paper mailback questionnaires from mailout, Internet responses, Telephone Questionnaire Assistance (TQA) interviews, paper mailback questionnaires in Update/Leave areas, paper mailback questionnaires in Urban Update/Leave areas, foreign language questionnaires, and Be Counted Forms (BCF).

[2] including List/Enumerate, Update/Enumerate, and Remote Alaska.

[3] including T-Night, Group Quarters, Military Group Quarters, and Shipboard Group Quarters.

[4] The numbers in this column are based on a sample. Therefore they are an estimate of the true value and contain sampling error.

They are in the following eight counties:

1. Maricopa County, Arizona
2. Pinal County, Arizona
3. Los Angeles County, California (3)
4. San Francisco County, California
5. Lafourche Parish, Louisiana
6. Bronx County, New York (2)
7. Dutchess County, New York
8. Charleston County, South Carolina

Table 3. Nationwide Distribution of Tracts by the Percent of All Households Within a Tract That Are Linguistically Isolated

Estimated Percent of All Households Within A Tract that are Linguistically Isolated	Tracts		
	Count	Percent	Cumulative Percent
Total	**64,960**	**100.00**	
0.0% to less than 5.0%	50,365	77.53	77.53
5.0% to less than 10.0%	6,596	10.15	87.69
10.0% to less than 15.0%	2,949	4.54	92.23
15.0% to less than 20.0%	1,747	2.69	94.92
20.0% to less than 25.0%	1,188	1.83	96.74
25.0% to less than 30.0%	803	1.24	97.98
30.0% to less than 35.0%	539	0.83	98.81
35.0% to less than 40.0%	336	0.52	99.33
40.0% to less than 45.0%	180	0.28	99.60
45.0% to less than 50.0%	111	0.17	99.78
50.0% to less than 55.0%	65	0.10	99.88
55.0% to less than 60.0%	39	0.06	99.94
60.0% to less than 65.0%	10	0.02	99.95
65.0% to less than 70.0%	14	0.02	99.97
70.0% to less than 75.0%	7	0.01	99.98
75.0% to less than 80.0%	2	0.00	99.99
80.0% to less than 85.0%	0	0.00	99.99
85.0% to less than 90.0%	0	0.00	99.99
90.0% to less than 95.0%	0	0.00	99.99
95.0% to 100.0%	9	0.01	100.00

Data Source: SEDF ( ) denotes standard error

Table 4 is similar to Table 3, but it represents the distribution at the county level. The table indicates for 91.53 percent of all counties, less than five percent of their households are linguistically isolated. Similar to the tract, this is expected given that nationally 4.1 percent of all households are linguistically isolated and there is no substantial clustering. There are eight counties where at least 25 percent of their households are linguistically isolated. These represent 0.25 percent of all counties. The eight counties are all in Texas. They are as follow: Hudspeth, Kenedy, La Salle, Maverick, Presidio, Starr, Webb, and Zavala.

Table 4. Nationwide Distribution of Counties by the Percent of All Households Within a County That Are Linguistically Isolated

Estimated Percent of All Households Within a county that are Linguistically Isolated	Counties		
	Count	Percent	Cumulative Percent
Total	**3,141**	**100.00**	
0.0% less than 5.0%	2,875	91.53	91.53
5.0% less than 10.0%	187	5.95	97.48
10.0% less than 15.0%	43	1.37	98.85
15.0% less than 20.0%	19	0.60	99.46
20.0% less than 25.0%	9	0.29	99.75
25.0% less than 30.0%	4	0.13	99.87
30.0% to 35.0%	4	0.13	100.00

Data Source: SEDF ( ) denotes standard error

4.4 What is the educational level of the householders?

Table 5 contains estimates and percentages of the householder's educational attainment level for linguistically isolated and non-linguistically households. The distribution shows that householders in the non-linguistically isolated households have higher education background than those in the linguistically isolated households. Up to the 12th grade/no diploma level, the percentages are higher for the linguistically isolated households than for the non-linguistically isolated households with one exception, 11th grade. Beyond that level, this reverses. Householders in linguistically isolated households are less likely to have high school diploma through doctorate degree than those in non-linguistically isolated households.

Table 5. Nationwide Distribution of Households by Educational Attainment Levels For the Householders in Linguistically Isolated Households and Non-Linguistically Isolated Households

Householder's Education Attainment Level	Linguistically Isolated Households		Non-Linguistically Isolated Households	
	estimates[1]	percent	estimates[1]	percent
Total	**4,362,318**	**100.00**	**101,117,783**	**100.00**
No schooling completed	357,710	8.20 (0.0358)	790,655	0.78 (0.0022)
Nursery school to 4th grade	195,210	4.47 (0.0269)	498,154	0.49 (0.0018)
5th grade or 6th grade	518,608	11.89 (0.0422)	1,258,034	1.24 (0.0028)
7th grade or 8th grade	348,169	7.98 (0.0353)	3,417,450	3.38 (0.0045)
9th grade	269,557	6.18 (0.0314)	2,324,314	2.30 (0.0038)
10th grade	145,790	3.34 (0.0234)	3,051,508	3.02 (0.0043)
11th grade	120,067	2.75 (0.0213)	2,936,248	2.90 (0.0042)
12th grade, no diploma	309,653	7.10 (0.0335)	3,414,054	3.38 (0.0045)
High school graduate	821,122	18.82 (0.0510)	27,910,534	27.60 (0.0112)
Some college, but less than one year	145,397	3.33 (0.0234)	7,281,509	7.20 (0.0065)
One or more years of college, no degree	331,227	7.59 (0.0345)	15,435,101	15.26 (0.0090)
Associate degree	143,851	3.30 (0.0233)	6,261,840	6.19 (0.0061)
Bachelor's degree	373,504	8.56 (0.0365)	16,498,235	16.32 (0.0093)
Master's degree	162,528	3.73 (0.0247)	6,452,019	6.38 (0.0061)
Professional degree	72,785	1.67 (0.0167)	2,362,761	2.34 (0.0038)
Doctorate degree	47,140	1.08 (0.0135)	1,225,367	1.21 (0.0027)

Data Source: SEDF ( ) denotes standard error

[1] The numbers in this column are based on a sample. Therefore they are an estimate of the true value and contain sampling error.

5. CONCLUSIONS AND RECOMMENDATIONS

Conclusions:

The percent of linguistically isolated households has increased since 1990, from 3.2 percent to 4.1 percent (SE = 0.0049 percent). Linguistically isolated households were less likely to self-respond compared to the non-linguistically isolated households. For linguistically isolated households, 61.0 percent (SE = 0.0636 percent) were self-responses and, for non-linguistically isolated households, 72.6 percent (SE = 0.0112 percent) were self-responses. The return type with the largest percent of both linguistically isolated households and non-linguistically isolated households, is paper mailback, 57.7 percent (SE = 0.0644 percent) and 71.2 percent (SE = 0.0114 percent), respectively. This would be expected given the percent of self-response for both the linguistically isolated and non-linguistically isolated households. Of all households enumerated by Coverage Edit Followup, 9.0 percent (SE = 0.0598 percent) are linguistically isolated; Nonresponse Followup, 5.8 percent (SE = 0.0117 percent); Coverage Improvement Followup, 5.5 percent (SE = 0.0414 percent); Personal Visit Enumeration, 5.9 percent (SE = 0.0507 percent). This indicates that among those operations involving an enumerator, Coverage Edit Followup has the highest percent of linguistically isolated households.

Of the 3,141 counties in the nation with at least one linguistically isolated household, 91.53 percent have less than five percent linguistically isolated households. There are eight counties with at least 25 percent of their households that are linguistically isolated. All of them are in Texas.

Of the 64,960 tracts in the nation with at least one linguistically isolated household, 77.5 percent have less than five percent linguistically isolated households. There are 11 tracts where at least 75 percent of their households are linguistically isolated. They are as follow: one in Maricopa County, Arizona, one in Pinal County, Arizona, three in Los Angeles County, California, one in San Francisco County, California, one in Lafourche Parish, Louisiana, two in Bronx County, New York, one in Dutchess County, New York, and one in Charleston County, South Carolina.

Finally, householders in linguistically isolated households are less likely to have formal education beyond 12th grade than those in non-linguistically isolated households.

Recommendation:

At the tract and county levels, the linguistic isolation variable may help with identifying areas for special enumeration procedures including language programs for the 2010 Census. Further analysis should be done by specific languages that are spoken at home in order to identify the level and if they are clustered.

REFERENCES

Bruce, Antonio and J. Gregory Robinson, Population Division, Census Bureau, (1999) *Targeting Census Tracts for the Census 2000 Language Program: Use of the Planning Database*

Lestina, Frederic and Deborah Griffin, (2001) *How are Non-English Speakers Enumerated in the American Community Survey?* Preliminary results presented to the American Community Survey Research Group (Census Bureau), January 17, 2001

Siegel, Paul, Elizabeth Martin, and Rosalind Bruno, US Census, (2000) *Language Use and Linguistic Isolation: Historical Data and Methodological Issues* Presented at the FCSM Statistical Policy Seminar, Bethesda, MD November 8-9, 2000

Whitworth, Erin, US Census, (2001) *How Selected Language Groups Coped with Census 2000* Proceedings of the Annual Meeting of the American Statistical Association, August 5-9, 2001

Appendix A: Key Words

These data are respondent-identified, that is, the respondent provides data for all household members.

Household language - In households where one or more persons age five years or older speak a language other than English, the household language assigned to all household members is the non-English language spoken by the first person with a non-English in the following order: head of household, spouse, parents, sibling, child, grandchild, other kin, companion, roommate, and other relatives. Thus, a person who speaks only English may have a language other than English assigned to him/her in tabulations of individuals by household language.

Linguistic Isolation - A linguistically isolated household is one in which all adults age 14 years and older speak a language other than English at home and have some limitation in communication in English. A household in which no person age 14 years or older who speaks a language other than English speaks English "very well" is labeled as linguistically isolated, including members under 14 years who may speak only English.

Return Type - A return type is the method of enumerating households in this study. Including in the study are the following return types:

- Paper mailback
- Coverage Edit Followup
- Nonresponse Followup
- Coverage Improvement Followup
- Personal Visit Enumeration

Householder - Head of the household, usually referred to as Person 1 for each household in the questionnaire. There are other household persons in the questionnaire and each of these describes his/her relationship to Person 1. Person 1 describes himself/herself as Householder.

Appendix B: Sample Edited Detail File (SEDF) Variable Definitions

RT - Record Type

2 = Housing Unit record
3 = Housing Unit person record
5 = Group quarters person record

MAFID - MAF and DMAF ID

characters 1-2 = state code when the MAF ID was assigned
characters 3-5 = county code when the MAF ID was assigned
characters 6-12 = control ID

LNGI - linguistic isolation

1 = not linguistically isolated
2 = linguistically isolated

QHIGH - Educational Attainment

01 = No schooling completed
02 = Nursery school to 4th grade
03 = 5th grade or 6th grade
04 = 7th grade or 8th grade
05 = 9th grade
06 = 10th grade
07 = 11th grade
08 = 12th grade, no high school diploma
09 = high school graduate
10 = some college but less than one year
11 = one or more years of college, no degree
12 = Associate degree
13 = Bachelor's degree
14 = Master's degree
15 = Professional degree
16 = Doctorate degree

QREL - relationship to the householder

01 = householder (self)
02 = spouse
03 = natural son or daughter
04 = adopted son or daughter
05 = step-son or step-daughter
06 = brother or sister
07 = parent
08 = grandchild
09 = parent-in-law
10 = son-in-law or daughter-in-law
11 = other relative
12 = brother-in-law or sister-in-law
13 = nephew or niece
14 = grandparent
15 = uncle or aunt
16 = cousin
17 = roomer/boarder
18 = roommate/housemate
19 = unmarried partner
20 = foster child
21 = other non-relative
22 = institutional group quarters person
23 = non-institutional group quarters person

RSOURCE - source of return

01 = paper mail back questionnaire from mail out
03 = paper mail back questionnaire from telephone questionnaire assistance (TQA) without ID
04 = paper mail back questionnaire from Update Leave
05 = paper mail back questionnaire from Update Leave ADD
06 = paper mail back questionnaire from Update Leave SUBSTITUTE
07 = paper mail back questionnaire from Urban Update Leave
08 = paper mail back questionnaire from Urban Update Leave ADD
09 = paper mail back questionnaire from Urban Update Leave SUBSTITUTE
10 = paper mail back questionnaire from Request for Foreign Language

11 = paper mail back questionnaire from Be Counted Form (BCF) marked as whole household
12 = paper mail back questionnaire from BCF marked as partial household

13 = paper enumerator questionnaire from List Enumerate
14 = paper enumerator questionnaire from Update Enumerate
15 = paper enumerator questionnaire from Update Enumerate ADD
16 = paper enumerator questionnaire from Update Enumerate Substitute

17 = paper enumerator questionnaire from Nonresponse Follow-up (NRFU)
18 = paper enumerator questionnaire from NRFU ADD
19 = paper enumerator questionnaire from NRFU SUBSTITUTE
20 = paper enumerator questionnaire from NRFU Whole Household
21 = paper enumerator questionnaire from NRFU Partial Household

22 = paper enumerator questionnaire from Coverage Improvement Follow Up (CIFU)
23 = paper enumerator questionnaire from CIFU ADD
24 = paper enumerator questionnaire from CIFU SUBSTITUTE

25 = paper enumerator questionnaire from T-Night
26 = paper questionnaire for Usual Home Elsewhere (UHE) from Service-based Enumeration (SBE) Individual Census Questionnaire (ICQ)
27 = paper questionnaire for UHE from Group Quarters (GQ) enumeration
Individual Census Report (CR)
28 = paper questionnaire for UHE from Military GQ enumeration Military Census Report (MCR)
29 = paper questionnaire for UHE from Shipboard GQ enumeration Shipboard Census Report (SCR)

30 = IDC (Internet) -Census short form survey only

31 = electronic Telephone Questionnaire Assistance (TQA) /Computer-assisted Telephone Interview (CATI) short form
32 = electronic TQA /CATI Be Counted Form (BCF) for whole household
33 = electronic TQA/CATI BCF for partial household

34 = electronic Coverage Edit Follow-up (CEFU) from long or short form survey
35 = electronic CEFU from BCF for whole household
36 = electronic CEFU from IDC

37 = paper enumerator continuation form - unlinked “orphan”

STATE - State Code

COUNTY - County Code

TRACT - Nonresponse Followup Tract

HWT - sampling housing unit weight

PWT - sampling person weight

Appendix C: State Linguistic Isolation Rank by Percent

Rank	State	Linguistically Isolated household percent	Rank	State	Linguistically Isolated household percent
-	**USA**	**4.14 (0.0049)**	26	Minnesota	1.85 (0.0206)
1	California	9.65 (0.0240)	27	Delaware	1.83 (0.0647)
2	New York	7.74 (0.0263)	28	Pennsylvania	1.80 (0.0144)
3	Texas	7.22 (0.0246)	29	Nebraska	1.79 (0.0340)
4	Hawaii	7.00 (0.1018)	30	Louisiana	1.73 (0.0256)
5	New Mexico	6.49 (0.0768)	31	Michigan	1.70 (0.0155)
6	New Jersey	6.19 (0.0371)	32	Oklahoma	1.56 (0.0234)
7	Florida	5.89 (0.0267)	33	Wisconsin	1.44 (0.0175)
8	Nevada	5.62 (0.0768)	34	New Hamsphire	1.42 (0.0412)
9	Arizona	5.61 (0.0458)	35	Maine	1.38 (0.0347)
10	Rhode Island	5.08 (0.0951)	36	Ohio	1.25 (0.0132)
11	Illinois	4.75 (0.0253)	37	Indiana	1.24 (0.0187)
12	Massachusetts	4.70 (0.0370)	38	North Dakota	1.22 (0.0429)
13	Connecticut	4.44 (0.0486)	39	Iowa	1.21 (0.0220)
14	District of Columbia	4.16 (0.1115)	40	South Dakota	1.16 (0.0401)
15	Colorado	3.39 (0.0362)	41	Arkansas	1.10 (0.0231)
16	Washington	3.35 (0.0314)	42	Missouri	1.02 (0.0158)
17	Oregon	2.88 (0.0371)	43	South Carolina	0.99 (0.0206)
18	Utah	2.66 (0.0474)	44	Tennessee	0.96 (0.0171)
19	Maryland	2.42 (0.0297)	45	Vermont	0.92 (0.0378)
20	Alaska	2.38 (0.0721)	46	Wyoming	0.86 (0.0485)
21	Georgia	2.32 (0.0234)	47	Kentucky	0.73 (0.0165)
22	Virginia	2.07 (0.0231)	48	Alabama	0.70 (0.0160)
23	Kansas	2.02 (0.0320)	49	Montana	0.68 (0.0286)
24	Idaho	1.91 (0.0471)	50	Mississippi	0.54 (0.0175)
25	North Carolina	1.89 (0.0194)	51	West Virginia	0.32 (0.0154)

Data Source: SEDF ( ) denotes standard error

The table above shows that 14 of the 50 states plus DC have percents of linguistically isolated households that are higher than the national-level percent. California has the highest percent of linguistically isolated households. West Virginia has the lowest percent.