New Mexico and Colorado Spanish Survey [dataset]

Neddy Vigil

Garland Bills

Recommended Citation
Linguistic Atlas and Archive of the Spanish of New Mexico and Southern Colorado (UNM LoboVault hdl.handle.net/1928/20424)

Contents:  
Part 1: Document Description  
Part 2: Study Description  
Part 3: Data Files Description  
Part 4: Variable Description

Document Description

Citation

Title: Linguistic Atlas and Archive of the Spanish of New Mexico and Southern Colorado  
Identification Number: hdl.handle.net/1928/20424  
Authoring Entity: Bills, Garland (University of New Mexico)  
Vigil, Neddy (University of New Mexico)  
Producer: Bills, Garland  
Vigil, Neddy  
Copyright: Copyright University of New Mexico Foundation, 2011

Study Description

Citation

Title: Linguistic Atlas and Archive of the Spanish of New Mexico and Southern Colorado  
Identification Number: hdl.handle.net/1928/20424  
Authoring Entity: Bills, Garland (University of New Mexico)  
Vigil, Neddy (University of New Mexico)  
Producer: Bills, Garland  
Vigil, Neddy  
Software used in Production: Alpha 4  
Funding Agency/Sponsor: National Endowment for the Humanities  
Grant Number: RT-21502-91  
Funding Agency/Sponsor: National Endowment for the Humanities
Grant Number: RT-21502-93

Distributor: University of New Mexico, Center for Southwest Research

Version: This publicly available version of the New Mexico and Colorado Spanish Survey data is available in dBase (dbf) format and includes all supporting files necessary to recreate the functionality and work flow of the original Alpha 4 database application. Additionally, the primary data tables used in the database have been transformed to comma separated value (csv) format to support migration to alternative applications and architectures. Updated versions of the database and mapping utilities used by the original study team have further been provided in the form of MS Access, Quantum GIS and MapServer applications. Documentation for these applications is included. Finally, it should be noted that the source data set contained personally identifying information about the survey consultants. To protect the privacy of these individuals, this information has been removed and extracted fields have been documented.

Study Scope

Keywords: New Mexico languages, Colorado languages, Spanish language -- variation -- United States, Spanish language -- variation -- history, Spanish language -- spoken Spanish -- United States

Topic Classification: linguistics, Spanish language

Abstract: Data collection for the Linguistic Atlas and Archive of the Spanish of New Mexico and Southern Colorado began in 1991 and was funded by a National Endowment for the Humanities (NEH) grant for an estimated three and a half years. The data collection for the study represented a first of its kind project in multiple ways. First, while the unique characteristics of New Mexican Spanish had previously been the focus of linguistic study, much of the research was focused on particular regions or communities and findings were generally undocumented. There had previously been no comprehensive, systematic study of New Mexican Spanish across the entire state. Additionally, the methodology was unique in that the computer aided data entry and analysis strategies developed for the NMCOSS were relatively new to linguistic study.

Only native born Spanish speakers were interviewed for the project, which included all of New Mexico along with parts of Colorado which experienced an influx of Spanish speaking settlers during the nineteenth century. Preparatory research in the selection of interviewees was completed using US Census data, with follow up work being carried out within individual locales via contact with various cultural, civic, and community organizations. Ultimately, 357 consultants were interviewed in locations ranging across New Mexico and the sixteen southern Colorado counties.

Country: United States
Geographic Coverage: New Mexico, Colorado

Unit of Analysis: Individuals
Universe: Native born Spanish speaking residents of New Mexico and southern Colorado.

Methodology and Processing

Time Method: Data was collected between 1991 and 1993 and represents the New Mexican dialect of
Spanish as spoken during the middle and latter half of the twentieth century.

**Sampling Procedure:**

Excerpted from the investigators' National Endowment for the Humanities proposal:

The first methodological concern in sampling is delimiting of the geographical scope of the project and defining the geographical components to be sampled. As the project's central concern is with Traditional Southwest Spanish, the focus will be on the long-established Hispanic population of northern New Mexico and southern Colorado. However, we want to also examine the encroaching influences on this Traditional Spanish, especially the strong Mexican Spanish influence that has resulted from immigration and close contact with Mexico in the southern part of the state. Therefore, the survey will cover the entire state of New Mexico. This coverage will include the metropolitan areas of Albuquerque, Las Cruces, and Santa Fe, where the greatest influence of both English and Mexican Spanish is felt. In the state of Colorado, on the other hand, nearly half (28) of the 63 counties have fewer than 200 persons above age eighteen who reported Spanish as a home language in 1980. We propose to include in our survey only the sixteen counties of the southern third of this state. These sixteen counties have sizable Hispanic populations and include all but two of the Colorado counties with more than 10% Spanish home language claimants. While it is relevant to note that these selected counties account for only 28% of the Spanish home language claimants of Colorado, it must also be pointed out that 63% of the total live in the four metropolitan areas stretching north of the single metropolitan area included in our study (Pueblo); in fact, fully 25% of Colorado's Spanish home language claimants reside in Denver. It is beyond the scope of the present project to explore the Spanish kaleidoscope of these other urban areas; these and the Spanish speaking areas of Arizona, California, and Texas might be examined in an extension of the project with the participation of scholars from across the Southwest.

The geographical area just delimited contains some 310,000 persons above the age of eighteen who were reported to have Spanish as a home language in the 1980 Census. We propose to collect data from some 350 to 400 of these adult Spanish speakers, a sampling ratio of 1/775 to 1/885. In distributing the sample across the area, geographical representation, settlement history, and population density will be taken into consideration. In order to take advantage of the valuable data from the 1980 Census, the easiest grid system to utilize is one based on the county. The entire area is divided into three regions: a Central region roughly representing the Rio Grande drainage area and Western and Eastern regions on each side of this. These regions are divided into a total of twelve sectors based on geography and settlement history as interpreted from such standard sources as Erickson & Smith 1985 and Williams 1986. The earliest Spanish colonization, for example, was in Sector 5, followed shortly by settlement in Sector 6 and later expansion to Sectors 9, 10, and 4.

Factors other than history and geography need to be taken into consideration. Underwood 1974 and others have criticized traditional dialect geography for failing to include broad sociological sampling in the surveys. This project will include sex and age as criteria in the selection of subjects. In each locality we will select one male and one female from three generations of adult speakers (grouped at ages 18-39, 40-60, and 61 or more). We will not include socioeconomic considerations as a criterion in consultant selection. In this situation, variation in Spanish is likely to be more weakly associated with socioeconomic factors than the linguistic variation that has been documented in other areas. In the largely bilingual society to be examined here, it is English rather than a special variety of Spanish that is most associated with the upper portions of any socioeconomic scale. This is not to say, nevertheless, that there are no value judgments associated with the different varieties of Spanish available. There are indeed such perceptions of "goodness" (as Kravitz 1985 has demonstrated). We intend to tap this sociolinguistic variation, not through a priori selection of subjects based on socioeconomic status, but through ad hoc description of many aspects of the social status of each subject (education, occupation, leisure activities, housing, visible
manifestations of wealth and literacy, etc.) and through elicitation of different styles of speech during the interview.

The sampling procedure may be summarized as follows. The survey will include six persons from each of 60-65 localities, yielding a total of 350-400 consultants in the sample. Within each of the twelve Sectors at least four localities will be selected for sampling. An additional 15 or so localities will be chosen in those Sectors having the greatest number of Spanish speaking adults. The localities selected within each Sector will represent areas showing the heaviest proportion of Spanish speakers (that is, areas where individuals can be expected to have a fairly well developed verbal repertoire in Spanish). The pre-identification of these localities on the basis of national census information will be facilitated by Bills' extensive recent research on the 1980 Census (e.g. Bills 1989). Finally, within each locality, six consultants will be chosen to fill the specific sex and age group cells, one male and one female in each of the three adult age groups. Additional requirements in the selection of individual consultants will be those traditionally set in dialect geography studies: native speaker of Spanish (having acquired the language in early childhood), life-long resident of the immediate area (with only temporary residence elsewhere), no serious speech problems, talkative, and so forth (cf. Allen 1971).

Individual consultants will be identified, to the extent possible, through an expanding system of personal contacts. In some areas, perhaps especially a few isolated rural areas, it may be necessary for the interviewer to go in virtually blind, approaching a knowledgeable local such as parish priest or minister, postal official, or sheriff, in order to pinpoint appropriate consultants. It is our expectation, however, that most consultant candidates can be identified in advance (by name and address) through contacts, principally telephone contacts, prior to even entering the sampling site. Quality time for lining up consultants is built in to our detailed work plan; more than an hour and a half is allocated for the identification of each consultant, and in addition, eight hours is allocated for the conduct of each 4-5 hour interview. Thus, the work plan allows for about five hours per consultant just for identifying and arranging an interview with that consultant.

Mode of Data Collection:

Excerpted from the investigators' National Endowment for the Humanities proposal:

This project proposes to collect from each subject a wide range of data to document not only regional variation, but also social variation, stylistic variation, and proficiency-based variation (Bills & Vigil 1988 and Bills 1990 examine these dimensions of variability and how they may be tapped in this project). Notwithstanding the exceptional scope intended, we will be able to collect those data in an interview of just four to five hours duration, a somewhat shorter interview than that typical in dialect geography research (e.g., 6-20 hours for the Linguistic Atlas of New England, 8 hours for the Atlas Lingüístico de México). This compact interview is due to an elicitation instrument designed to meet the special needs of this project and to take maximum advantage of the time-saving capacity of the tape recorder as discussed by Pederson 1974, Underwood 1972, and others.

The interview can be divided into three sections distinguished by the kind of information to be gathered. In addition to a section documenting the personal background of the consultant, this instrument will include two major components: specific elicitation and free conversation. Though these two major components are discussed below as if they were isolated components, in the conduct of the interview, they are intermixed in an innovative fashion to make for a more interesting and less tiring interview. Since much of the specific elicitation is divided into semantic categories (plants, animals, household, transportation, and such), the interviewers will encourage conversational deviation into topics important to the consultant at numerous points in the sequence.

The specific elicitation questionnaire will contain a maximum of 800 items. Because these
items will be easy to access and will provide comparability across all informants, they will form
the linguistic basis for the atlas. These items will exhibit specific features primarily of lexicon
and phonology but also some of morphology and syntax as well. These features represent a
selection from three major sources: (1) those known by us--through personal experience and
citation in the literature--to show variation within the region; (2) those contained in the 1,000-
item questionnaire of the Atlas Lingüístico de México (Lope Blanch 1970a); and (3) those
documented for earlier periods of the Spanish of the region in Hills 1906, Espinosa 1909, and
other works cited in Teschner et al. 1975. Many items will, of course, be represented in all
three sources.

Virtually all lexical and phonological items will be elicited by means of pictures and real
objects. This is a far more efficient method than the traditional verbal elicitation which results
in the investigator producing ten words for each word elicited from the subject whose speech
is being investigated (Underwood 1972:216). Over the past three years we have done
extensive pilot testing of elicitation of items through pictures--both line drawings and
magazine photographs. We have carried out this testing with a wide variety of subjects from
all over the state (reported in Vigil et al. 1989), including elderly subjects ranging up to 82
years of age. We have encountered no problems or difficulties. On the contrary, the elicitation
was smooth, accurate, and highly efficient, yielding over 350 responses per interview of less
than one hour.

Some important items, however, especially for morphological and syntactic phenomena, are
not susceptible to representational elicitation. We will therefore make occasional use of the
traditional verbal completion techniques (e.g., "Ahora veo este caballo y ayer también lo . . ."
to elicit the preterit form vi/vide). Taking advantage of the bilingual character of this
population, we may also employ translation in a few cases where other methods of elicitation
are too unreliable and inefficient.

Dialect geography projects typically employ a much longer list of specifically elicited items
than the 800 planned here. For, example, the number of questionnaire items was 1,000 for
the Atlas Lingüístico de México, 1,350 for the Atlas Lingüístico y Etnográfico de Colombia
(Montes Giraldo 1964), 711 for the Linguistic Atlas of New England (Kurath et al. 1939), 1,500
for the Linguistic Atlas of the Gulf States (Pederson et al. 1974), and 1,322 for the Survey of
English Dialects in Britain. Our list is relatively short for two reasons: (1) A single item may
reveal two (or more) linguistic features. The word martillo 'hammer', for instance, provides
information on two important phonological features, syllable final /r/ and intervocalic /yl/. (2)
Our list will be limited to those items that we feel confident are known by most prospective
consultants. We will not waste time trying to elicit specialized terms from nonspecialists (e.g.,
asking lifelong city dwellers to identify farm implements, asking females to describe animal
slaughtering procedures traditionally carried out by males, or asking elderly rural citizens
about intricacies of low rider automobiles). Such information is, of course, very important to
both linguistic and cultural documentation. We must, therefore, tap it. In this survey, however,
that specialized knowledge will be elicited from the specialists, as explained in the description
below of the free conversation elicitation. In the specific elicitation process, interviewers will
also explore the consultants' awareness of forms other than their initial responses. Knowledge
of multiple forms can be expected to be quite common for lexical items. Wherever multiple
forms are elicited, the interviewer will attempt to get an assessment of the form the consultant
actually uses as well as an evaluative reaction to the alternate forms mentioned.

The second major component of the interview is the elicitation of "conversational" data. Here
the consultants will be encouraged to expound on topics in which they have the greatest
interest and expertise. The focus of elicitation will be aspects of personal history and
traditional activities related to that person's own experience. The kinds of topics to be pursued
are childhood games, leisure activities of youth, adult work activities, and dangerous and
humorous moments in the consultant's life. This elicitation will require the interviewers' finest
skills, as so comprehensively and lucidly discussed in Briggs 1986, a seminal examination of interview methodology deriving from fieldwork among Hispanics in the northern New Mexico village of Córdoba. The interviewers will have to be involved conversational participants who are profoundly interested in the consultants' contributions (even when heard for the fiftieth time) and who can astutely probe for details and try to secure comparable information across consultants having similar expertise.

At the conclusion of the interview the interviewer will switch to English in order to assess the subject's ability in English. This assessment, together with the self-report of English proficiency requested at the beginning of the interview, will make it possible to explore any correlations of Spanish use with bilingual proficiency.

Sources Statement

Weighting: Excerpted from the investigators' National Endowment for the Humanities proposal:

Coding will consist of preparing the relevant data for creation of a computer file. The data to be coded include all personal information about the consultant (except name), all specific elicitation responses, the tabulations for conversational use of selected phonological and grammatical features, and the indexing information for each tape. All data will be coded as numerical values and entered in the format of a standard program for statistical analysis and mapping. For the linguistic data, the coding process will first involve selection, that is, determination of how many variants of each variable and subvariable need to be recognized. For individual consultants, up to three alternate forms of each item (by degree of preference) may be entered. All data will be entered as separate variants initially.

File Description--F1
File: COUNTY.DBF
- Number of cases: 49
- No. of variables per record: 4
- Type of File: dBase III

File Description--F2
File: GRAMMAR.DBF
- Number of cases: 726
- No. of variables per record: 2
- Type of File: dBase III

File Description--F3
File: LEXICON.DBF
- Number of cases: 5404
- No. of variables per record: 2
- Type of File: dBase III
**File Description--F4**

**File: MASTER.DBF**
- Number of cases: 357
- No. of variables per record: 48
- Type of File: dBase III

**File Description--F5**

**File: RANGE3.DBF**
- Number of cases: 4
- No. of variables per record: 5
- Type of File: dBase III

**File Description--F6**

**File: RANGE5.DBF**
- Number of cases: 6
- No. of variables per record: 3
- Type of File: dBase III

**File Description--F7**

**File: RESPONSE.DBF**
- Number of cases: 262130
- No. of variables per record: 5
- Type of File: dBase III

**Variable Description**

*List of Variables:*
- **COUNTY** - COUNTY
- **CODE** - CODE
- **ST** - ST
- **SECTOR** - SECTOR
- **ENG** - ENG
- **SPAN** - SPAN
- **LOND** - LOND
- **NO_TAPES** - Number of Tapes
- **LONM** - LONM
- LONS - LONS
- LATD - LATD
- LATM - LATM
- LATS - LATS
- DATE - Interview Date
- CITY - CITY
- STATE - STATE
- INTRVIEWER - INTRVIEWER
- COUNTY - COUNTY
- INT_STATE - INT_STATE
- SECTOR - SECTOR
- URBAN - URBAN
- RES_YOUTH - Where Grew Up
- AGE_SPAN - Aged Learned Spanish
- AGE_ENG - Age Learned English
- ENG_AB_VER - English Ability
- SP_AB_VERB - Spanish Ability
- CHILD_LANG - Childhood Home Language
- OTHER_LANG - Note if not English or Spanish
- SP_SPOUSE - With Spouse
- SP_CHILDRN - With Children
- SP_WORK - At Work
- SP_FRIENDS - With Friends
- SP_YRS_STD - Years Formal Spanish Study
- YRS_SP_CNT - Years in Spanish-Speaking Countries
- READ_SP - Read Spanish
- WRITE_SP - Write Spanish
- PREF_SP_TV - Prefer Spanish TV
- SPTV_HR_WK - Hrs/Wk
- PREF_SP_RD - Prefer Spanish Radio
- SPRADIO_WK - Hours/Week of Spanish Radio
- ADJLOND - ADJLOND
- ADJLONM - ADJLONM
- ADJLONS - ADJLONS
- ADJLATD - ADJLATD
- ADJLATM - ADJLATM
- ADJLATS - ADJLATS
- LONGDEC - LONGDEC
- LATDEC - LATDEC
- LOCATION - LOCATION
- LOC_ST - LOC_ST
- PARAGOGIC - PARAGOGIC
- PALATAL - PALATAL
- ASPIRATION - ASPIRATION
- NUMBER - NUMBER
- ABILITY - ABILITY
- ACCENT - ACCENT
- LITERATURE - LITERATURE
- INCOME - INCOME
- NUMBER - NUMBER
- LANG_USE - LANG_USE
- ASSESS_AB - ASSESS_AB
- INT_NO - INT_NO
- ENG - ENG
<table>
<thead>
<tr>
<th>Variables</th>
<th>Width</th>
<th>Description</th>
<th>Variable Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>COUNTY</td>
<td>10</td>
<td>County name.</td>
<td>character</td>
</tr>
<tr>
<td>CODE</td>
<td>2</td>
<td>Two letter county name abbreviations.</td>
<td>character</td>
</tr>
<tr>
<td>ST</td>
<td>2</td>
<td>State in which the county is located.</td>
<td>character</td>
</tr>
<tr>
<td>SECTOR</td>
<td></td>
<td>Demographic sector in which the county is located.</td>
<td>numeric</td>
</tr>
<tr>
<td>ENG</td>
<td>20</td>
<td>English terms for items used in the grammar portion of the survey.</td>
<td>character</td>
</tr>
<tr>
<td>SPAN</td>
<td>22</td>
<td>Spanish equivalents for English grammar survey items as provided by consultants.</td>
<td>character</td>
</tr>
<tr>
<td>ENG</td>
<td>19</td>
<td>English terms for elicitation images used in the lexical portion of the survey.</td>
<td>character</td>
</tr>
<tr>
<td>SPAN</td>
<td>26</td>
<td>Spanish equivalents for English lexical survey items as provided by consultants.</td>
<td>character</td>
</tr>
<tr>
<td>Interview Number</td>
<td></td>
<td>Unique interview number.</td>
<td></td>
</tr>
</tbody>
</table>
Range of Valid Data Values: 1-366
Variable Format: numeric

**Number of Tapes**

Description: Number of tapes.
Range of Valid Data Values: 1-8
Variable Format: numeric

**LOND**

Instructions: Read the map!
Description: Degrees longitude of consultant location.
Range of Valid Data Values: 103-108
Variable Format: numeric

**LONM**

Instructions: Read the map!
Description: Minutes longitude of consultant location.
Range of Valid Data Values: 0-59
Variable Format: numeric

**LONS**

Instructions: Read the map!
Description: Seconds longitude of consultant location.
Range of Valid Data Values: 0-59
Variable Format: numeric

**LATD**

Instructions: Read the map!
Description: Degrees latitude of consultant location.
Range of Valid Data Values: 31-38
Variable Format: numeric

**LATM**

Instructions: Read the map!
Description: Minutes latitude of consultant location.
Range of Valid Data Values: 0-59
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Range of Valid Data Values</th>
<th>Variable Format</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LATS</strong></td>
<td>Instructions: Read the map!</td>
<td></td>
<td>numeric</td>
</tr>
<tr>
<td></td>
<td>Description: Seconds latitude of consultant location.</td>
<td>0-58</td>
<td></td>
</tr>
<tr>
<td><strong>Interview Date</strong></td>
<td>Description: Interview date.</td>
<td></td>
<td>numeric</td>
</tr>
<tr>
<td><strong>CITY</strong></td>
<td>Description: City of residence.</td>
<td></td>
<td>character</td>
</tr>
<tr>
<td><strong>STATE</strong></td>
<td>Description: State of residence.</td>
<td></td>
<td>character</td>
</tr>
<tr>
<td><strong>INTRVIEWER</strong></td>
<td>Description: Name of interviewer.</td>
<td></td>
<td>character</td>
</tr>
<tr>
<td><strong>COUNTY</strong></td>
<td>Description: County of residence.</td>
<td></td>
<td>character</td>
</tr>
<tr>
<td><strong>INT_STATE</strong></td>
<td>Description: State in which interview took place.</td>
<td></td>
<td>character</td>
</tr>
<tr>
<td><strong>SECTOR</strong></td>
<td>Description: Demographic sector, as defined by principle investigators, of consultant location.</td>
<td>0-12</td>
<td>numeric</td>
</tr>
<tr>
<td><strong>URBAN</strong></td>
<td>Instructions: Leave blank if question was not asked.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Description: Does the consultant live in an urban environment?

Variable Format: character

**Where Grew Up**

Width: 30

Description: Location where consultant lived as a youth.

Variable Format: character

**Aged Learned Spanish**

Width: 2

Instructions: Enter 1 for native speaker, 0 for never learned.

Description: Age at which the consultant learned Spanish.

Variable Format: character

**Age Learned English**

Width: 2

Instructions: Enter 1 for native speaker, 0 for never learned.

Description: Age at which the consultant learned English.

Variable Format: character

**English Ability**

Description: Consultant's self assessment of his/her English speaking ability as measured on a three point scale in which 1 = None/Very Little, 2 = Get Along, 3 = Good/Comfortable. A zero rating is available for N/A.

Range of Valid Data Values: 0-3

Variable Format: numeric

**Spanish Ability**

Description: Consultant's self assessment of his/her Spanish speaking ability as measured on a three point scale in which 1 = None/Very Little, 2 = Get Along, 3 = Good/Comfortable. A zero rating is available for N/A.

Range of Valid Data Values: 0-3

Variable Format: numeric

**Childhood Home Language**

Width: 2

Description: Language used in the consultant's home as a youth.

Variable Format: character

**Note if not English or Spanish**

Instructions: Enter other languages used in following four (4) ratings.

Description: For following four (4) questions pertaining to current language use, interviewer to note if consultant uses any languages other than English or Spanish.
Variable Format: numeric

**With Spouse**

Instructions: Compare English/other language use to Spanish.

Description: Current language(s) used with spouse as measured on a five point scale in which 1 = All English, 2 = Mostly English, 3 = Both Equally, 4 = Mostly Spanish, 5 = All Spanish. A zero rating is available for N/A.

Range of Valid Data Values: 0-5

Variable Format: numeric

**With Children**

Instructions: Compare English/other language use to Spanish.

Description: Current language(s) used with children as measured on a five point scale in which 1 = All English, 2 = Mostly English, 3 = Both Equally, 4 = Mostly Spanish, 5 = All Spanish. A zero rating is available for N/A.

Range of Valid Data Values: 0-5

Variable Format: numeric

**At Work**

Instructions: Compare English/other language use to Spanish.

Description: Current language(s) used at work as measured on a five point scale in which 1 = All English, 2 = Mostly English, 3 = Both Equally, 4 = Mostly Spanish, 5 = All Spanish. A zero rating is available for N/A.

Range of Valid Data Values: 0-5

Variable Format: numeric

**With Friends**

Instructions: Compare English/other language use to Spanish.

Description: Current language(s) used with friends as measured on a five point scale in which 1 = All English, 2 = Mostly English, 3 = Both Equally, 4 = Mostly Spanish, 5 = All Spanish. A zero rating is available for N/A.

Range of Valid Data Values: 0-5

Variable Format: numeric

**Years Formal Spanish Study**

Width: 2

Description: Number of years of formal Spanish language study.

Variable Format: character

**Years in Spanish-Speaking Countries**
**Read Spanish**

Description: Consultant's self assessed ability to read Spanish as measured on three point a scale in which 1 = None/Very Little, 2 = Get Along, 3 = Good/Comfortable. A zero rating is available for N/A.

Variable Format: numeric

**Write Spanish**

Description: Consultant's self assessed ability to write Spanish as measured on three point a scale in which 1 = None/Very Little, 2 = Get Along, 3 = Good/Comfortable. A zero rating is available for N/A.

Variable Format: character

**Prefer Spanish TV**

Description: To determine whether the consultant prefers Spanish or English television. Yes or no.

Variable Format: character

**Hrs/Wk**

Description: Hours per week spent watching Spanish television.

Variable Format: character

**Prefer Spanish Radio**

Description: To determine whether the consultant prefers Spanish or English radio. Yes or no.

Variable Format: character

**Hours/Week of Spanish Radio**

Description: Hours per week spent listening to Spanish radio.

Variable Format: character

**ADJLOND**

Description: Degrees longitude of consultant location, adjusted so that multiple consultants within a single residence would be represented separately within mapping applications.

Range of Valid Data Values: 103-108

Variable Format: numeric

**ADJLONM**

Description: Minutes longitude of consultant location, adjusted so that multiple consultants
within a single residence would be represented separately within mapping applications.
Range of Valid Data Values: 0-59
Variable Format: numeric

**ADJLONS**

Description: Seconds longitude of consultant location, adjusted so that multiple consultants within a single residence would be represented separately within mapping applications.
Range of Valid Data Values: 0-59
Variable Format: numeric

**ADJLATD**

Description: Degrees latitude of consultant location, adjusted so that multiple consultants within a single residence would be represented separately within mapping applications.
Range of Valid Data Values: 31-38
Variable Format: numeric

**ADJLATM**

Description: Minutes latitude of consultant location, adjusted so that multiple consultants within a single residence would be represented separately within mapping applications.
Range of Valid Data Values: 0-63
Variable Format: numeric

**ADJLATS**

Description: Seconds latitude of consultant location, adjusted so that multiple consultants within a single residence would be represented separately within mapping applications.
Range of Valid Data Values: 0-86
Variable Format: numeric

**LONGDEC**

Description: Decimal longitude coordinates of consultant's location. Originally a calculated field: -1*(ADJLOND+ADJLONM/60+ADJLONS/3600)
Range of Valid Data Values: -108.86--103.09
Variable Format: numeric

**LATDEC**

Description: Decimal latitude coordinates of consultant's location. Originally a calculated field: ADJLATD+ADJLATM/60+ADJLATS/3600
Range of Valid Data Values: 31.5-38.3
Variable Format: numeric
LOCATION
Variable Format: numeric

LOC_ST
Variable Format: numeric

PARAGOGIC
Width: 1
Description: Phonemes subject to paragoge.
Variable Format: character

PALATAL
Width: 1
Description: Palatal consonants.
Variable Format: character

ASPIRATION
Width: 1
Description: Phonemes subject to aspiration.
Variable Format: character

NUMBER
Description: Indexes for lookup tables used to categorize various self, interviewer and transcriber assessments included in the survey.
Range of Valid Data Values: 0-3
Variable Format: numeric

ABILITY
Width: 16
Description: Lookup values corresponding to consultants' self assessed reading and writing abilities.
Variable Format: character

ACCENT
Width: 6
Description: Lookup values corresponding to interviewer assessment of consultants' English and Spanish accents. These qualitative assessments have been removed from the publicly available data set at the request of the principle investigators.
Variable Format: character

LITERATURE
Width: 4
Description: Lookup values corresponding to interviewer assessment of material evidence of consultants' English and Spanish literacy. These qualitative assessments have been removed from the publicly available data set at the request of the principle investigators.
Variable Format: character
### INCOME

**Width: 13**  
**Description:** Lookup values corresponding to consultants' household income. These qualitative assessments have been removed from the publicly available data set at the request of the principle investigators.  
**Variable Format:** character

### NUMBER

**Description:** Indexes for lookup tables used to categorize various self, interviewer and transcriber assessments included in the survey.  
**Range of Valid Data Values:** 0-5  
**Variable Format:** numeric

### LANG_USE

**Width: 14**  
**Description:** Lookup values corresponding to consultants' self assessed use of Spanish at home, work, and socially.  
**Variable Format:** character

### ASSESS_AB

**Width: 6**  
**Description:** Lookup values corresponding to interviewer assessment of consultants' English and Spanish abilities. These qualitative assessments have been removed from the publicly available data set at the request of the principle investigators.  
**Variable Format:** character

### INT_NO

**Width: 4**  
**Description:** Interview number of the consultant providing the response.  
**Variable Format:** character

### ENG

**Width: 20**  
**Description:** English term or elicitation image used to prompt consultants.  
**Variable Format:** character

### SPANISH

**Width: 36**  
**Description:** Spanish equivalent provided by consultants for English term, elicitation image or prompt.  
**Variable Format:** character

### CODE

**Width: 19**  
**Description:** Codes designating preferred Spanish equivalents in the event of multiple usages. Data entry options for this field include 1st Choice, 1st Choice/P, 2nd Choice, 2nd Choice/P, Co-Equal Choice and Co-Equal Choice/P.  
**Variable Format:** character
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Range of Valid Data Values</th>
<th>Variable Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECNO</td>
<td>A unique identifier for each response.</td>
<td>0-262160</td>
<td>numeric</td>
</tr>
</tbody>
</table>