
ATTITUDES TOWARDS SPANISH, ENGLISH AND CODE-SWITCHING IN TWO BORDERING CITIES IN SOUTH TEXAS

STATISTICAL ANALYSIS BY
JONATHAN PRESTO

MAY 4th, 2015

INTRODUCTION

1.1 Background and Problem Statement

Linguistic attitudes towards a spoken language could be attributed by certain members of certain national or cultural group. One would expect members belonging to similar cultural groups to hold similar perceptions or reactions to one spoken language verses another. Any variation in linguistic attitudes within a particular cultural group is an area of research interest. Some factors for variation could include age, gender or region. Understanding one's attitude or perception toward a particular language guise could have interesting applications for business or even government policy-makers, but the motivation for this study is largely driven by pure research in the social sciences.

South Texas comprises several regions, including bordering regions to Mexico, such as the city Laredo and the Rio Grande Valley ("RGV") region. Given South Texas close proximity to Mexico and the rich cultural history that go beyond the Mexican American war (1848), the two regions share many things in common, including a large Hispanic population and the prevalence of bilingual communities. What differentiates the two regions, however, is RGV is largely known for its agricultural practices, whereas Laredo is positioned as a prime location for international commerce and trade. These socioeconomic factors are expected to influence living conditions and social behavior, which in turn could affect linguistic attitudes toward the use of Spanish and code-switch.

The objective of this study is to measure any differences in implicit linguistic attitudes that may carry towards the use of English, Spanish and code-switch between the two target populations – people residing in Laredo and RGV. The research hypothesis to be tested is that residents from Laredo will hold more positive or favorable attitudes towards the use of Spanish and code-switch than residents from RGV across the characteristic levels of Solidarity, Status and Personal Appeal.

STUDY DESIGN

2.1 Research Subjects

The investigation was conducted as a cross-sectional study, a type of observational study conducted in one specific point in time. The client conducted survey questionnaires face-to-face as a means of data collection. A total of 200 university students were chosen as research subjects, 100 students from Laredo and the other 100 students from McAllen, a city within the Rio Grande Valley. The subjects came from a pool of students enrolled in an introductory course in Hispanic Studies. The professors who taught these courses volunteered their students to participate (students did not necessarily volunteer themselves).

2.2 Speakers

The investigation was carried out using an implicit approach known as "matched- guise" technique. Matched-guise is a sociolinguistic experimental technique first introduced by Wallace Lambert and his colleagues at McGill University in the 1960s to determine attitudes held by bilingual French Canadians towards English and French guises (Davies & Elder 2004:189, Agheyisi & Fishman, 1970, Lambert et. al, 1960:60).

Six speakers were recruited to produce a total of fourteen recordings. Four of the speakers were designated as “guise” speakers, which comprised of two males and two females. The remaining two speakers were designated as “filler” speakers, one male and one female. The four “guise” speakers are all bilingual. Each guise recorded three passages, same messaging, but in three different communications – English, Spanish and code-switch (switching between English and Spanish) – for a total of 12 recordings. In addition, the two fillers each recorded one message bringing the total to 14 recordings. The additional two recordings would add some deception to the listener to help with the impression that all guises are different. The research subjects were not informed that a single speaker could have recorded multiple messages. Having these various matched guises helped control for potential confounding effects of gender and/or voice quality. The factor of interest, after all, is the language used.

2.3 Procedure

Per classroom, all subjects listened simultaneously to a given sequence of 6 recordings (6 chosen from the 14 available). Each recording represented one combination of the three languages and two genders. A recording was played immediately followed by the opportunity to fill-out the measures questionnaire. The sequence was constrained such that no one speaker or language is to be played contiguously.

Each research subject completed a total of seven questionnaires – one questionnaire captured informational attributes about the subject such as age, sex, place of birth, information about parents and so forth. The six measure questionnaires captured the subject’s judgments or attitudes about the speaker by rating the speaker on a set of traits using a 1-to-5 scale (‘1’ being Not at All and ‘5’ being Very). Each trait was categorized into one of the three main characteristics: Solidarity, Status and Personal Appeal. Refer to APPENDIX A and APPENDIX B for details on the questionnaire.

STATISTICAL METHOD

3.1 Data Preparation

Data was collected from 187 of the 200 students selected from the study. Thirteen students did not submit their questionnaires. That leaves 96 from Laredo and 91 from RGV who submitted their questionnaires. Majority of the respondents were females in both regions; 80% in Laredo and 81% in RGV. The mean age for Laredo and RGV was 25 and 23 years old respectively. Some respondents were excluded for two main reasons: first being those individuals with a substantial amount of unanswered ratings (L047, L068, R019), and second, those individuals who did not live in their current location for 6 or more years were removed to help ensure the sample was well representative of their respective regions given the data collected at hand (upper 90th percentile of all respondents lived in their respective region for 6 or more years).

For each individual respondent, 9 additional mean-value columns (9 combinations of dialog and characteristic levels) were calculated in order to carry out inferences using t-Tests. The data set was first aggregated by respondent, then by the 3 dialog levels: English, Spanish, code-switch. For each of the 3 dialog levels, the mean of the trait scores for a given characteristic category was computed across the male and female speakers. To illustrate an

example, the column labeled *LAR_CS_Per* corresponds to the mean score of personal appeal traits *Intelligent, Hard-working, Good-looking, Funny and Open-minded* rated by a respondent from Laredo across the male and female speakers of code-switch.

To help control for potential measurement error at the individual level, the English dialog was used as a reference category (baseline category). A set of 6 adjusted ratings was derived from the 9 mean-value columns as described in the previous paragraph for each individual. The adjustment was done by subtracting a characteristic's English mean rating from the corresponding characteristic's Spanish mean rating. Similarly, a characteristic's English mean rating was subtracted from the corresponding characteristic's code-switch mean rating. Hence, an adjusted rating can have a positive or a negative value depending on its relative rating against the English rating (note: there are no adjusted ratings for English since they are redundantly equal to 0 for each individual). To illustrate another example, the column labeled *dLAR_CS_Per* corresponds to an adjusted rating computed by subtracting *LAR_EN_Per* from *LAR_CS_Per* for an individual respondent. APPENDIX C further describes these 6 adjusted ratings.

3.2 Matched Pairs

Each individual from Laredo was best matched with another individual from RGV by gender and age. A SAS macro that uses a case-control matching algorithm on propensity score performed the matching. Basic demographic information was collected and used to perform the matching. Age and gender were chosen to be the basis for matching. The matched-pairs algorithm resulted in 73 matched pairs having 96% matched by gender and 79% matched by age within ± 1 year. Using 6 years as the cut-off for location years resulted in a matched set having 90% of Laredo individuals living half or more of their lifetime at present location and 70% for RGV respectively. Having this cut-off helped ensure the subset of the individuals chosen is representative of their respective regions.

3.3 Data Analysis

Measuring favorability of Spanish and code-switch was performed using the derived adjusted ratings described in APPENDIX C. Given an individual and a characteristic category, a positive value for an adjusted rating would indicate a favorable response for Spanish or code-switch over English. In order to detect if Laredo holds more favorable ratings than RGV on Spanish and code-switch, the mean difference of the adjusted ratings (relative ratings) was computed between Laredo and RGV participants using the "matched-pairs" statistical method. To control for potential confounding effects of age and gender, two individuals, one from each region, were matched algorithmically by age and gender. An observation used for statistical inference was the calculated difference for one pair's adjusted ratings for each characteristic level. For instance, in order to measure Laredo's higher favorability than RGV's for Spanish verses English on solidarity characteristic: first compute the difference of two adjusted ratings, *dLAR_SP_Sol* minus *dRGV_SP_Sol*, for each *k*th pair from 1 to *n*; then average those *n* difference values and run a t-Test to see if the average difference equals 0 (null hypothesis) or not equal to zero (two-sided alternative). Higher positive values for a *t* statistic indicates Laredo having a higher favorability for Spanish (or code-switch) than that of RGV. Also worth mentioning is that exploratory analysis was done to show no large deviance from normality of the difference of adjusted ratings; normality assumption should be satisfied in order to carry out valid t-tests and confidence intervals. Refer to APPENDIX F for the normality plots.

RESULTS

TABLE 1 shows the mean difference between Laredo and RGV on their adjusted ratings by characteristic and by dialog comparison. A total of six matched-pair t-Tests were conducted using a two-tailed test at significance level 0.05 – two dialog comparisons (Spanish vs. English and code-switch vs. English) for each of the three characteristic levels (Solidarity, Status and Personal Appeal). There was no significant difference between Laredo and RGV in their respective ratings for Spanish vs. English guises on all three characteristics. Although Spanish guises were evaluated higher on Status relative to the other two characteristics, there was not enough statistical evidence to suggest a favorable response for Status alone. Similarly, there was no significant difference between Laredo and RGV in their respective ratings for code-switch vs. English guises on all three characteristics.

The 95% confidence intervals and corresponding standard errors for the mean difference values were also calculated (refer to Table D-1 of Appendix D). Because each interval contains 0, this further validates the conclusion for lack of sufficient evidence to support the research hypothesis. GRAPH A illustrates an example of a scatter plot of difference in adjusted ratings between paired subjects on personal appeal for code-switch. The observed points are uniformly scattered around the center of the graph. Because there is no distinct pattern, this validates lack of sufficient evidence to support the research hypothesis. APPENDIX E contains the remaining scatter plots.

TABLE 1

MEAN DIFFERENCE (LAREDO MINUS RGV) OF THE ADJUSTED RATINGS (RELATIVE RATINGS)

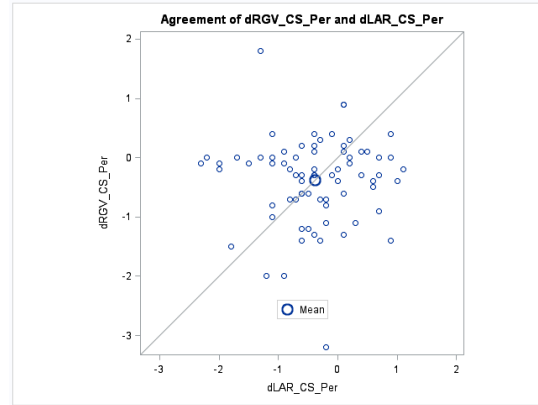
	Spanish vs. English			Codeswitch vs. English		
	Mean Difference	T statistic	P-value	Mean Difference	T statistic	P-value
Solidarity	0.0205	0.18	0.8582	-0.0317	-0.24	0.8125
Status	0.1659	1.96	0.0544	0.2018	1.55	0.1244
Personal Appeal	0.1198	1.17	0.2464	-0.0069	-0.05	0.9563

Note. Positive entries on mean difference values indicate that Laredo evaluated a particular characteristic category more favorably for Spanish (or code-switch) than that of Rio Grande Valley.

* indicates *t* is significant at the .05 level, on two-tailed tests.

GRAPH A

DIFFERENCE IN ADJUSTED RATINGS ON PERSONAL APPEAL TRAITS: ENGLISH VS. CODE-SWITCH



DISCUSSION

Based on the overall results from the matched-pairs t-Test, there is insufficient evidence to suggest that residents of Laredo and RGV differ in their linguistic attitudes towards Spanish and code-switch guise for a given characteristic category. Laredo does exhibit relatively higher scores for Spanish and code-switch guise in the characteristic for Status in comparison to Solidarity or Personal Appeal, but not enough to support the research hypothesis.

Measurement error at the subject level, if not accounted for, could have introduced a confounded effect, which could then lead to invalid t-statistics. An individual, regardless of residence, may for some reason, just like giving consistently high scores, or low scores. This effect was controlled for by setting the English dialog as a reference category (baseline category). So the adjusted scores reflect a subject's relative ratings with respect to the English rating.

The issue of *convenience sampling* known from conducting this cross-sectional survey is worth mentioning here. Because the pool of participants was chosen only from University students enrolled in classes for introductory Hispanic Studies, one must exercise caution when applying these results to the general population in Laredo and RGV. The participants may not be representative of the actual attitudes of the population (sample selection bias). One suggestion to help increase the sample's likelihood of representation is by taking a stratified random sample from each of the two populations. Stratification is important to ensure a weighted representation by certain demographics such as gender, age groups, socioeconomic status, and so forth, to closely resemble that of the general population.

Another suggestion to improve the study is to enhance the "practicality" of the rating questionnaire by including an Importance rating. For example, an individual may choose to rate the trait *Good-looking* with a '3', but that score may not necessarily indicate its relative importance to the rater in determining one's overall *Personal Appeal*. A participant who rates *Good-looking* with a '3' followed by an Importance of '5' as oppose to one who also rates *Good-looking* with a '3' but followed by an Importance of '1' may indicate a difference of attitudes towards that trait. So coupling each Rating score with an Importance score ('1' being *not very important* and '5' being *very important*) may introduce some additional insights to the researcher.

CONCLUSION

The aim of this study was to detect for differences in linguistic attitudes between two target populations in South Texas – Laredo and the Rio Grande Valley. The researcher hypothesized that residents of Laredo hold more positive attitudes towards the use of Spanish and Code-Switch language than those residents of the Rio Grande Valley.

University students enrolled in Introductory Hispanic Studies were chosen as research subjects to participate. Each student submitted a set of survey questionnaires. The analysis was designed around a matched-guise technique, controlling for the factors of language, location, respondent gender, respondent age, speaker and speaker gender. Each subject was asked to rate several traits categorized by three characteristics (three response variables) used to implicitly measure linguistic attitudes – solidarity, status and personal appeal.

Statistical inference was carried out using matched pairs t-Tests. A student from Laredo was matched with another from RGV by age and gender *and* only including those pairs living in their respective locations for 6 or more years (to help improve regional representation given the data at hand) . The t-Tests resulted in no significant evidence to support the research hypothesis for any of the three characteristic categories.

There is opportunity to further advance this research study. A variety of other information was collected about the research subjects but not used the analysis, such as birthplace of student, birthplace of parents, or language preferences under various settings. Some of this information can be structured as covariates to help improve matching of pairs, or even be used to drive further insight into the population when designing a method having a more representative sample.

REFERENCES

Agheyisi, R.; Fishman, J. (1970): “Language Attitudes: A Brief Survey of Methodological Approaches”, *Anthropological Linguistics*, 12. pgs. 137-157.

Davies, A. & Elder, C. 2004, *The Handbook of Applied Linguistics*, Blackwell Publishing, UK.

Lambert, W.E., Hodgson, R.C., Gardner, R.C., Fillenbaum, S. (1960): “Evaluation Reactions to Spoken Language”, *Journal of Abnormal and Social Psychology*, 60, 44-51.

Parsons, L.S., Using SAS ® Software to Perform Case-Control Match on Propensity Score in an Observational Study. Paper 225-25

APPENDIX A – INFORMATIONAL QUESTIONNAIRE

Informational Questionnaire

Answer the following questions accordingly:

1. Sex: (Circle one) F or M
2. Age: _____
3. Where were you born?
4. Where were your parents born? _____
5. Where does your father work? _____
6. Where does your mother work? _____
7. How long have you been living in Laredo or Rio Grande Valley? (Indicate in either months or years) _____
8. At what age did you learn how to speak: English _____ Spanish _____
9. What language(s) do you use for the following relationships? (Fill in the right column with the corresponding language for each relationship.)

Parents	
Siblings	
Grandparents	
Friends	
Neighbors	

10. Do you ever mix Spanish and English when speaking? _____

11. If yes, do you mix Spanish and English in conversations with the following relationships? (Circle either Yes or No)

Parents	Yes or No
Siblings	Yes or No

Grandparents	Yes or No
Friends	Yes or No

12. Do you write in Spanish, English, or both for the following: (Fill in the right column with the corresponding answer.)

School/ Work	
Journal Entries	
Letters/E-mails	
Facebook and other social media	

13. Is there anything you can say in Spanish that you can't say in English, and vice-versa?

14. In what language do you watch TV for the following: (Fill in the right column with the corresponding answer.)

News	
Sitcoms	
Talk shows	
Movies	

15. In what language do you read for the following: (Fill in the right column with the corresponding answer.)

Magazines	
School textbooks	
Internet	
Recreational reading	

APPENDIX B – MEASURES QUESTIONNAIRE

On a scale of 1 to 5, rate the speaker that you have just heard on the following characteristics.
1 being “Not at all” and 5 being “Very” :

Not at all [1 - 2 - 3 - 4 - 5] **Very**

(Circle only one number)

Solidarity	
Honest	1 - 2 - 3 - 4 - 5
Kind	1 - 2 - 3 - 4 - 5
Likeable	1 - 2 - 3 - 4 - 5
Friendly	1 - 2 - 3 - 4 - 5
Sincere	1 - 2 - 3 - 4 - 5
Religious	1 - 2 - 3 - 4 - 5
Status	
Rich	1 - 2 - 3 - 4 - 5
Old-Fashioned	1 - 2 - 3 - 4 - 5
Educated	1 - 2 - 3 - 4 - 5
High Class	1 - 2 - 3 - 4 - 5
Urban	1 - 2 - 3 - 4 - 5
Personal Appeal	
Intelligent	1 - 2 - 3 - 4 - 5
Hard-working	1 - 2 - 3 - 4 - 5
Good-Looking	1 - 2 - 3 - 4 - 5
Funny	1 - 2 - 3 - 4 - 5
Open-minded	1 - 2 - 3 - 4 - 5

TABLE B-1
SIXTEEN MEASURED TRAITS

SOLIDARITY	STATUS	PERSONAL APPEAL
Honest	Rich	Intelligent
Kind	Old-Fashioned	Hard-working
Likeable	Educated	Good-looking
Friendly	High Class	Funny
Sincere	Urban	Open-minded
Religious		

APPENDIX C – DERIVED NUMERICAL FIELDS

TABLE C-1

<i>Match_ID</i> (Unique identifier of Laredo individual matched with RGV individual)	
Column Name	Description
<i>Laredo_Respondent_ID</i>	Unique identifier of Laredo individual
<i>LAR_CS_Per</i>	Average score rated by Laredo individual on Personal Appeal traits for speakers of Codeswitch
<i>LAR_CS_Sol</i>	Average score rated by Laredo individual on Solidarity traits for speakers of Codeswitch
<i>LAR_CS_Sta</i>	Average score rated by Laredo individual on Status traits for speakers of Codeswitch
<i>LAR_EN_Per</i>	Average score rated by Laredo individual on Personal Appeal traits for speakers of English
<i>LAR_EN_Sol</i>	Average score rated by Laredo individual on Solidarity traits for speakers of English
<i>LAR_EN_Sta</i>	Average score rated by Laredo individual on Status traits for speakers of English
<i>LAR_SP_Per</i>	Average score rated by Laredo individual on Personal Appeal traits for speakers of Spanish
<i>LAR_SP_Sol</i>	Average score rated by Laredo individual on Solidarity traits for speakers of Spanish
<i>LAR_SP_Sta</i>	Average score rated by Laredo individual on Status traits for speakers of Spanish
<i>dLAR_CS_Per</i>	Difference of scores on CS minus English for Personal Appeal
<i>dLAR_CS_Sol</i>	Difference of scores on CS minus English for Solidarity
<i>dLAR_CS_Sta</i>	Difference of scores on CS minus English for Status
<i>dLAR_SP_Per</i>	Difference of scores on Spanish minus English for Personal Appeal
<i>dLAR_SP_Sol</i>	Difference of scores on Spanish minus English for Solidarity
<i>dLAR_SP_Sta</i>	Difference of scores on Spanish minus English for Status
<i>RGV_Respondent_ID</i>	Unique identifier of RGV individual
<i>RGV_CS_Per</i>	Average score rated by RGV individual on Personal Appeal traits for speakers of Codeswitch
<i>RGV_CS_Sol</i>	Average score rated by RGV individual on Solidarity traits for speakers of Codeswitch
<i>RGV_CS_Sta</i>	Average score rated by RGV individual on Status traits for speakers of Codeswitch
<i>RGV_EN_Per</i>	Average score rated by RGV individual on Personal Appeal traits for speakers of English
<i>RGV_EN_Sol</i>	Average score rated by RGV individual on Solidarity traits for speakers of English
<i>RGV_EN_Sta</i>	Average score rated by RGV individual on Status traits for speakers of English
<i>RGV_SP_Per</i>	Average score rated by RGV individual on Personal Appeal traits for speakers of Spanish
<i>RGV_SP_Sol</i>	Average score rated by RGV individual on Solidarity traits for speakers of Spanish
<i>RGV_SP_Sta</i>	Average score rated by RGV individual on Status traits for speakers of Spanish
<i>dRGV_CS_Per</i>	Difference of scores on CS minus English for Personal Appeal
<i>dRGV_CS_Sol</i>	Difference of scores on CS minus English for Solidarity
<i>dRGV_CS_Sta</i>	Difference of scores on CS minus English for Status
<i>dRGV_SP_Per</i>	Difference of scores on Spanish minus English for Personal Appeal
<i>dRGV_SP_Sol</i>	Difference of scores on Spanish minus English for Solidarity
<i>dRGV_SP_Sta</i>	Difference of scores on Spanish minus English for Status

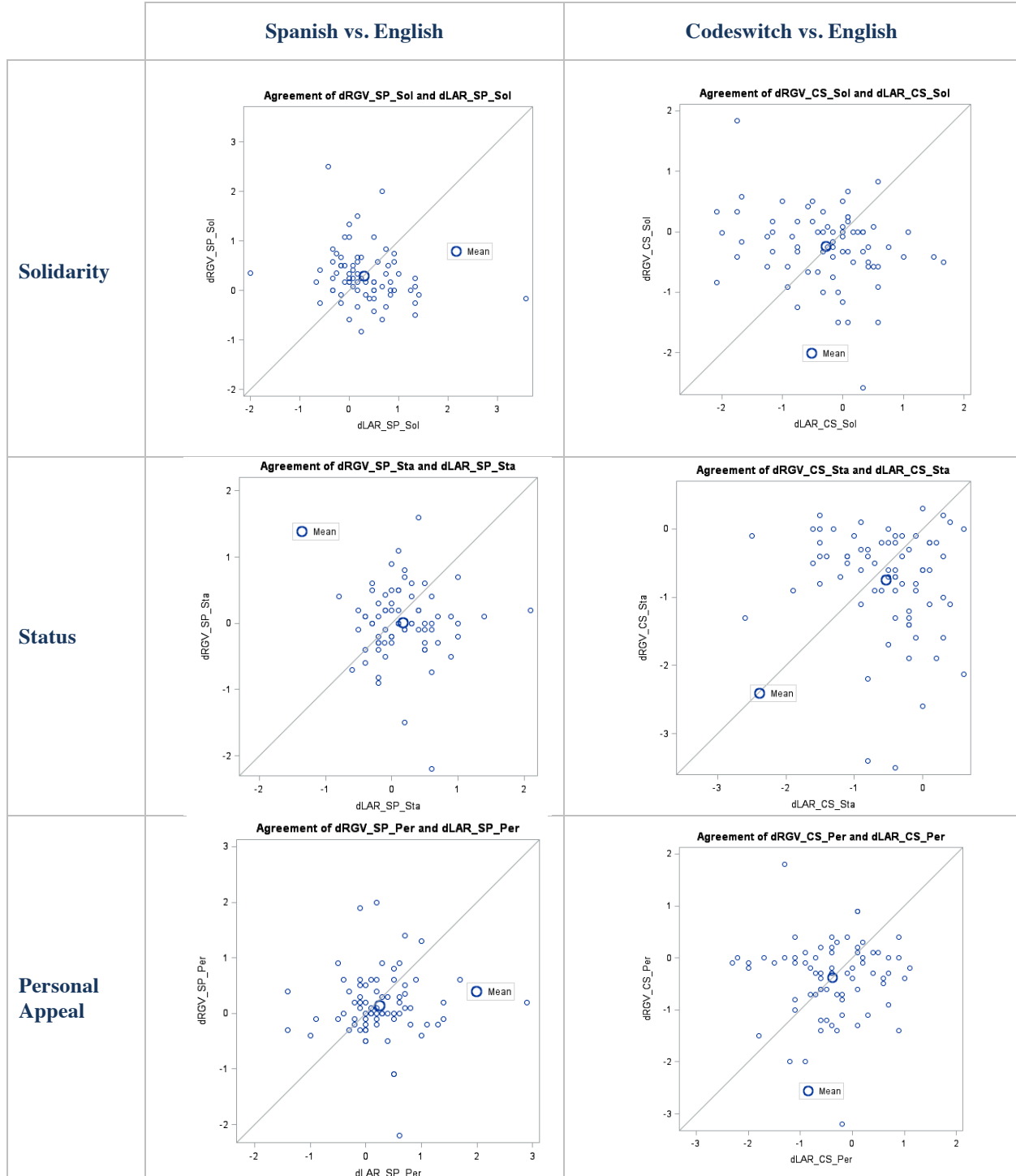
APPENDIX D – CONFIDENCE INTERVALS AND STANDARD ERRORS

TABLE D-1

95% CONFIDENCE INTERVALS AND STANDARD ERRORS OF THE MEAN DIFFERENCE (LAREDO MINUS RGV) ON THE ADJUSTED RATINGS (RELATIVE RATINGS), 72 DEGREES OF FREEDOM.

	Spanish vs. English			Codeswitch vs. English		
	Lower	Upper	Standard Error	Lower	Upper	Standard Error
Solidarity	-0.2077	0.2488	0.1145	-0.2975	0.2340	0.1333
Status	-0.00320	0.3350	0.0848	-0.0569	0.4606	0.1298
Personal Appeal	-0.0845	0.3241	0.1025	-0.2554	0.2417	0.1247

APPENDIX E – GRAPHICS FOR DIFFERENCE OF ADJUSTED RATINGS



APPENDIX F – NORMALITY IN DIFFERENCE OF ADJUSTED RATINGS

