Testing the Fidelity of English-Spanish Translations in Election Materials

Michael Portal and Abril M. Brugo

Rice University
Abstract

The purpose of this research is to identify and test inconsistencies in language usage and meaning as they may appear in translated election material. Following amendments to the Voting Rights Act of 1965, the production of minority language election material became a federal requirement without clearcut federal oversight. Utilizing a 2 (Direction: Spanish-to-English or English-to-Spanish) x 2 (Formality: if the language was expected to be formal “voting language” or casual “everyday language”) Between study we sought to assess the quality of these translations. We analyze significant and recurring words used in ballots from Ballotpedia’s 2015 list of “Notable local measures” and their translations as provided by LA County’s Spanish Translation Glossary. While many of the words tested did not have a main effect of Direction or Formality, some words tested demonstrate deficiencies in translation or word usage that could jeopardize the fidelity of intended meaning. We examine these results alongside other important research, primarily concerning the use of plain-language on ballots. This paper encourages further study into the words chosen for ballots and the potential for non-American translations and computer learning programs to help resolve these issues moving forward.

Keywords: Election material, translation, fidelity, plain language, Voting Rights Act
Testing the Fidelity of English-Spanish Translations in Election Materials

“… translations are never perfect; something is always left out, hidden, covered up, well illustrating their fragmentary nature, their failure to travel intact, and, thus, their suitability for cultural manipulation.”

Edwin Gentzler (2008)

In 1975 the Voting Rights Act (VRA) was amended to expand and protect access to election materials for minority language citizens. Congress justified the amendment by observing that, “through the use of various practices and procedures, citizens of language minorities have been effectively excluded from participation in the electoral process” (Gupta, 2015). The addition of Section 203 sought to remedy this inequity by guaranteeing that all “materials or information relating to the electoral process” be made available in the “language of the applicable minority group” (Gupta, 2015). Originally intended to protect only four groups -- Spanish, Asian, Native American, and Alaskan Natives -- the amendment requires the provision of suitable election material for any population that is either over 10,000 citizens or at least five percent of the larger population. Unsurprisingly, the amendment dramatically increased the amount of translated electoral material produced each year; LA County, for example, provides election material online in 17 different languages (“Find My Election Information,” n.d.). This federal requirement persists today (“52 USC 10503: Bilingual election requirements,” n.d.).

With potentially 25 million or more Americans with “limited English proficiency” (Cha & Kennedy, 2014), the addition of Section 203 certainly increased access to the ballot box. However, the long and contested path to this access, and the way forward, remains highly contested. To Ancheta (2007) the legislation remains essentially passive/reactive in nature, insufficient to confront the realities of inequity in modern America. While the 1975 amendment may have produced “powerful mechanisms to ensure the right to vote and to increase the
participation of minority voters,” as she argues, it remains fundamentally an “antidiscrimination policy” and “not a substitute for an agenda that includes public policies under which both language assistance and English language learning are integrated into public services and the educational system” (Ancheta, 2007, pp. 32–33). This view appears consistent with other criticism levied by scholars like Yoshino (1999) or Abdelall (2004) and, more distantly, with larger discussions over governmental obligations alluded to in the likes of Meylaerts (2011) and Fishkin (2011).¹

Other criticisms of the legislation have more wholly rejected it in its (near) totality. One such example is the US House’s 1996 passage of the Emerson English Language Empowerment Act, which would make English the official language of the US and completely repeal the 1975 amendment to the VRA. Although the Act died in the Senate, the demonstrated support for the legislation represented to some, in text and deed, a “Nativist attack against immigrants who are not yet fluent in English” and a larger agenda to paint minority language speakers as “victimizers and aggressors” (Chiu, 2015). Support for these “English-only” programs, however, are a reoccurring phenomena; one such instance is Congresswoman Ginny Brown-Wait’s, a republican house member from Florida, 2006 comment on bilingual ballot provisions that have “long kept new citizens from increasing their knowledge of our language and from fully integrating into our society. Not only is it expensive to print ballots in a variety of different dialects and tongues, but it reinforces a fractious society” (Congress, 2010, p. 1174; Kibbee, 2016, p. 45).

¹ More pointed criticisms of “accommodation” systems can be located, philosophically, in the work of Foucault (2003) or Agamben (2005). Discussions of this question also arise in disabilities studies, see works like Thomas (2007).
It is for this reason that it bears repeating Howard Zinn’s belief that it is “[t]he responsibility of intellectuals is to discard the notion of ‘objectivity’ and the notion of ‘disinterested scholarship’” (“History as a Political Act,” 1998). Access to the ballot box occurs in a non-insignificant context; translations do not occur in a void (Ray, 2008). Bermann (2006), in studying legal translations of English-French in Louisiana, observed this reality directly. Translations and decisions over word choice are and were, as he noted, decided “by preferences that derive from and reflect the history against the background of which the translation occurred” (Bermann, 2006, p. 101). For historians on the subject (see Grofman and Davidson (1992) and Tucker (2016)) this context is a critical component of an “ongoing struggle” (Tucker, 2006, p. 259) for equity that dominates political life.

This study is embroiled in larger discussions on how this “struggle” manifests itself in the language of ballots, in the minutia of election activity that represents one (if not, the) most tangible dialogue between state and citizen. For many, the significance of this cannot be understated: equity in voting may be the true breaking point for the view that we are (or can be) fairly represented in the government (King, Stivers, & Box, 1998).

**Literature Review**

“… there is significant noncompliance across counties covered by Section 203 provisions, both in the provision of written materials for linguistic minorities as well as the availability of staff assistance in languages other than English.” Jones-Correa and Waismel-Manor in Henderson (2007)

Despite numerous instances where translations have estranged populations from voting -- for example, in 2002, Vietnamese voters in Orange County encountered an extremely poor translation of voter instructions, where the sample ballot’s use of “nontraditional idioms and [sic]
were so vague that it made election documents nearly impossible to understand” and, ultimately, made “the page of instructions on how to vote by absentee ballot undecipherable [sic]” (Reyes, 2002) -- little work has been done to study the quality of the access granted to minority language groups by the 1975 amendment to the VRA. While the law does have provisions governing the quality of the translations (“material provided in the language of a language minority group be clear, complete and accurate” (“28 CFR 55.19 - Written materials,” n.d.)), the system(s) for oversight remain ambiguous. Save for catastrophic translation errors, the federal government simply determines what and where multi-lingual election materials must be provided while “most of the details are left to state and local election officials” (Benson, 2007, p. 375).

Unfortunately, few have studied these details. While her work deals more specifically with readability than translation fidelity, Reilly (2015) represents the foundation for our research. Particular to our work here, she finds

“There are several complications to the translation process, and the difference between the English and translated versions create a chasm in voter accessibility… Whether this is accidental or deliberate, the difference in the reading ease of these ballot measures demonstrates that accessibility is still an issue for voters who speak minority languages” (Reilly, 2015, pp. 90–91).

For Reilly, overly formalized language, direct translations not concerned with accessibility, etc. all represent threats to equitable access for minority language voters and leads to a “clear difference in the versions of the ballot language presented to English-speaking and voters with limited English proficiency” (Reilly, 2015, p. 108). This is only partially a translation problem: highly complex language will depress participation and influence democratic activity/engagement independent of the election material having been translated or not (Reilly, 2015;
Reilly & Richey, 2009). This well-studied problem at the forefront of modern usability research provides a useful tangent to inform our research.

In this domain we draw particular attention to push for “plain language” on ballots. Simply, it is the push for language whose “wording, structure, and design are so clear that the intended readers can readily find what they need, understand it, and use it” (“Center for Plain Language - Make it clear,” n.d.). This approach is frequently called a reader-oriented method for writing and design. Laskowski and Redish (2006) outline best practices (for example, explaining consequences for actions before an action is taken) for such an approach and studied ballots from across the country, finding that “[m]ost, if not all, ballots, both on paper and on screen, violate at least some of the best practices.” This is critical since reading is already a complicated activity where even minor increases in “reading difficulties can affect voter education, and unavoidably affect the process of marking a ballot and casting the vote” (Summers, Chisnell, Davies, Alton, & Mckeever, 2014). Reinforced by Herrnson et al. (2006), errors or discrepancies in expression, intended action, or any of the other interactions contained in the voting systems effected by language can alter an election.²

The reality of this is increasingly well documented. Here, Redish et al. (2009) and Redish et al. (2010) present some of the most compelling research. These studies confirm concerns raised in Roth (1998) about usability, demonstrating that participants voted more accurately and had better attitudes about voting when using plain language systems. When testing voters with sample ballots of differing complexity in language (one ballot being “plain

² Others also working in this domain confirm this: “While all groups of voters are affected by poorly designed ballots and badly drafted instructions, these problems disproportionately affect low-income voters, new voters, and elderly voters. All too often, the loss of votes and rate of errors resulting from these mistakes are greater than the margin of victory between the two leading candidates” (Norden, Kimball, Quesenbery, & Chen, 2008).
language”), Redish et al. (2009) found the “difference in accuracy between the two ballots is marginally statistically significant (F1,43=3.413, p < .071) …[showing that participants] voted more accurately on the ballot with plain language instructions” (Janice G. Redish et al., 2009, pp. 30–31). The smaller 2010 study confirmed these findings, especially the results that have plain language improving voter attitudes: “voters recognized plain language, preferring short, simple words, short paragraphs, and clear explanations. When asked for an overall preference (Ballot A, Ballot B, no preference), 82% chose Ballot B, the ballot with plain language instructions” (Janice Ginny Redish et al., 2010, p. 98).

Improved attitudes can translate into votes and increased participation, addressing some of the reasons outlined by voters who are discouraged from voting or returning to vote (Magpantay, 2004, p. 48). More abstractly, this is why Meylaerts (2011) notes that “linguistic and translational territorality regimes puts non-translation as much as translation on the research agenda” (Meylaerts, 2011, p. 753). This is also why the study of political language goes beyond just the ballot box, like Hill and Moreno (2001) demonstrate when studying the way data collection can be influenced by decisions over language. In part, the ever evolving nature of language and the context for which it occurs explains why Newman (2007) and Barbas (2009) both believe that VRA language provisions ought to be constantly reformed in light of changes like technological advances or demographic shifts. It is only by questioning these systems and the direction of these changes that we hope to uncover and make intelligible (i.e. make useful for policy) these the injustice they may come with them (Aslani & Salmani, 2015; Schaeffner, 2010). In sum, tracking and analyzing these issues is a pressing issue at the core of understanding American governance and policy making (Lessig, 1993).
Our discussion of plain language provides utility for better understanding translations, but it is not without criticism (Penman, 1993). Some argue that law is simply not reducible to plain language or that complicated language is an inevitability, making the task to simplify language burdensome and wasteful (Assy, 2011). Others the enforcement of ideology by supporters of plain language that incorrectly “assert that only one style or variation of language use should be considered legitimate” (Turfler, 2015). However, these criticisms have not fully responsive to the role of plain language in translations where the benefits of the language are certainly more pronounced. This is fairly intuitive since “[p]lain language minimizes or explains jargon or terms of art that don’t translate well. There is less interpreting for the translator to do and therefore less opportunity for error or bias… [which] can save costs in translating to other languages and makes using the ballot easier and more efficient for all voters” (Chisnell, 2014, p. 9). Listed under areas for future studies in Redish et al. (2009), better understanding how these issues manifest themselves in the minutia of ballots (especially translated ballots) is the logical next step for understanding the contested role complex language plays in political life.

This is why research into the political language *writ large* is key to better understanding politics (Burnett & Kogan, 2015). In studying translations specifically, Marschall and Rutherford (2016) who, despite finding quantifiable benefits to the language assistance programs amended to the VRA, recognize that an increase in access to voting has not had changed the makeup of candidates or up-stream equity for those in minority language groups. These issues are larger and more complex than just writing style or word choice, they require a concerted engagement with the practical components and experiences of those at the margins. Their study reified the “need for vigilance if minority rights are to be protected and expanded” (Marschall & Rutherford,
2016, p. 605). It is the aim of our study to test these practices in context to, in part, remain vigilant of the errors and discrepancies that are systematically ignored in lieu of other questions in political science and usability studies (Alvarez, 2002).

**Methods**

**Subjects**

Participants for this study were recruited in three different ways. Some were recruited online through Experimetrix (a site where Rice students sign up for psychology experiments) and were given 1 credit for participation in the study. Other participants were approached through email advertisements to qualifying populations. Lastly, we also recruited participants by posting advertisements on Rice University social media pages. Our study had a random sample of 65 participants, only 40 of whom completed the survey in its entirety. Of our respondents that completed the survey 23 were Female, 16 Male, and 1 Prefer not to Disclose. Participants ranged between the ages of 18 and 67, with exactly 50% indicating Spanish as their first language and 50% indicating English as their first language. 21 participants translated words from Spanish-to-English, while 19 translated words from English-to-Spanish. 21 participants were then placed in the Formal condition, and 19 were placed in the Informal condition.

**Design**

To investigate the effects of ballot translations on voter comprehension and attitudes we conducted a 2 (Direction: Spanish-to-English, English-to-Spanish) x 2 (Formality: Formal, Informal) factorial design with between subject groups. Our experiment consisted of two between-subject independent variables. The first independent variable was Direction which had two levels: Spanish-to-English and English-to-Spanish. The second independent variable was
Formality with two levels: Formal and Informal. The dependent variable in this study was the Accuracy Score, which was determined by assigning a correct translation a score of “1” and an incorrect translation a score of “0”. These scores were then averaged together to calculate an Accuracy Score.

Participants were assigned to one of four conditions. Group 1 had to read and answer multiple choice questions concerning electoral terms in Spanish with the goal of identifying the word’s synonym in English as they would use in daily conversation. Group 2 had to read and answer multiple choice questions concerning electoral terms in Spanish with the goal of identifying the word’s synonym in English as they would expect to see on a formal ballot or voting document. Group 3 had to read and answer multiple choice questions concerning electoral terms in English with the goal of identifying the word’s synonym in Spanish as they would use in daily conversation. Group 4 had to read and answer multiple choice questions concerning electoral terms in English with the goal of identifying the word’s synonym in Spanish as they would expect to see on a formal ballot or voting document. Sorting the participants in this way allowed for us to analyze and control the data in two ways. First, it allowed us to determine potential discrepancies between the language people use in a formal and informal context and if that had an impact on the accuracy of translations. Second, allowing participants to only translate in one direction ensured that participants were not primed with the correct answer given that we only tested a subset of twelve words.

Data was then exported from Qualtrics into an Excel Spreadsheet. A composite Accuracy Score was calculated for each participant by averaging their responses. Data was then inputted into SPSS by setting up the variables of Direction (1: Spanish-to-English, 2: English-to-Spanish)
and Formality (3: Formal, 4: Informal). Participant’s responses to questions 1 to 12 were recorded individually in order to assess the effect of the independent variables on each individual word. A Univariate ANOVA was then run on the data to determine the effects of Direction and Formality on the composite Accuracy Score as well as the Accuracy Scores of Questions 1 to 12.

**Materials**

The experiment was administered through an online Rice Qualtrics survey. Participants were assigned to one of four conditions where they were given a short multiple choice questionnaire to translate words according to their condition.

The words we selected were a result of an analysis of the United States’ top 10 most popular ballot initiatives of 2015. We selected which words to use on the questionnaire by cross-referencing both the words importance and frequency of use. Given that these terms were in Spanish we identified the “correct” Spanish translation by using the LA Vote Spanish Translation Glossary. We populated the rest of the multiple choice answers (in both English and Spanish) by using Google Translate and selecting the top three results.
Procedures

Before participating in the study users were given a consent form to review that contained information on the benefits and risks of participation, purpose of the study, and contact information of the researchers. The consent form was followed by a short questionnaire. Participants were first asked a series of demographic questions that helped quantify how many years they had been speaking Spanish and English. We also tried to account for regional variations in Spanish by asking them to specify which region of the world spoke a Spanish that most closely resembled their own. Participants were then given a self-complied twelve item multiple choice questionnaire where they had to translate words into their closest equivalent. After participation in the study respondents were shown a debrief form.
Results

A two-way analysis of variance was conducted on the influence of two independent variables (Direction, Formality) on the accuracy of translations. Direction consisted of two levels (Spanish-to-English, English-to-Spanish) and Formality consisted of two levels (Formal, Informal).

Composite

Our experiment, overall, failed to support our hypothesis that direction and Formality would have an effect on the accuracy of translations. There was no main effect of Direction, $F(1,36)=2.56, p>0.12$, showing no statistical difference between Spanish-to-English translations ($M=0.58, SD=0.19$) and English-to-Spanish translations ($M=0.50, SD=0.16$). Formality, $F(1,36)=0.07, p>0.79$, also did not yield a main effect on the accuracy of translations showing no statistical significance between Formal ($M=0.54, SD=0.16$) and Informal ($M=0.54, SD=0.19$) conditions. There was also no interaction between Direction and Formality, $F(1, 36) =.75, p>0.39$. As indicated by Figure 1, accuracy in all conditions hovered around the same range and had an average accuracy of 53.7%.

![Figure 2. Effect of Direction and Formality on Translation Accuracy Scores](image-url)
**Abatement, Reducción**

When it came to the translation of the word Abatement/Reducción there was a significant effect of Direction at the p<.05 level, $F(1, 36) = 14.03, p < .001$. English-to-Spanish ($M=0.53$, $SD=0.51$) was translated with a much higher accuracy than Spanish-to-English ($M=0.05$, $SD=0.22$) which can be seen in Figure 2. There was no main effect of Formality, $F(1, 36) = 1.15, p > .29$, on the accuracy of translations showing no statistical significance between the Formal ($M=0.19$, $SD=0.40$) and Informal ($M=0.37$, $SD=0.50$) conditions. There was also no interaction that could subsume the main effect of Direction, $F(1, 36) = .03, p > .86$.

![Figure 3. The Effects of Direction and Formality on Translation Accuracy Scores on the word Abatement/Reducion](image)

**Accordance, Conformidad**

The translation of the word Accordance/Conformidad did not demonstrate any main effects of Direction, $F(1, 36) = .006, p > .93$, or Formality, $F(1,36) = p > .40$. There was no statistical difference between Spanish-to-English ($M=.24$, $SD=.44$) and English-to-Spanish ($M=.26$, $SD=.45$). There was also no statistical significance between the Formal ($M=.19$, $SD=.40$) and...
Informal ($M=.32, SD=.48$) conditions. There was also no interaction between Direction and Formality, $F (1, 36) = .1, p > .76$.

**Candidate, Candidato**

The translation accuracy for the word Candidate/Candidato did not show an effect of Direction, $F (1, 36) = 3.35, p > .75$ or Formality, $F (1,36) = .69, p > .41$. There was no statistical significance in the accuracy of Spanish-to-English: ($M=.86, SD=.36$) and English-to-Spanish ($M=1.00, SD=.00$) translations. There was also a lack of significance between the Formal ($M=.95, SD=.22$) and Informal ($M=.89, SD=.32$) conditions. This analysis also did not yield an interaction between Direction and Formality on translation accuracy, $F (1, 36) = .69, p > .41$.

**Contest, Contienda**

The translation of the word Contest/Contienda yielded a significant effect of Direction at the $p < .05$ level, $F (1, 36) = 7.8, p < .008$. Spanish-to-English translations ($M=.57, SD=.51$) had a higher accuracy rate than English-to-Spanish translations ($M=.16, SD=.37$). There was no main effect of Formality, $F (1, 36) = 4.0, p > 0.054$ showing a lack of statistical significance between the accuracy scores in the Informal ($M=.21, SD=.42$) and Formal ($M=.52, SD=.51$) conditions. There was no interaction that could subsume the main effect of Direction, $F (1, 36) = .16, p > .692$.

**Party, Partido**

Translation accuracy did not show a main effect for both Direction, $F (1,36) = .57, p > .46$, or Formality, $F (1,36) = 2.27, p > .14$. There was no statistical significance between Spanish-to-English ($M=.76, SD=.44$) and English-to-Spanish: ($M=.63, SD=.5$). Nor is there a significant statistical difference between the Formal ($M=.81, SD=.4$) and Informal ($M=.58, SD=.4$).
51) conditions. There was also no interaction between Direction and Formality, \( F(1,36) = .709, p > .71 \).

**Enact, Promulgar**

There was no effect of Direction, \( F(1,36) = 1.47, p > .23 \), or Formality, \( F(1,36) = 2.69, p > .11 \), on the accuracy of translations for the word Enact/Promulgar. There was no statistical significance between Spanish-to-English translations (\( M = .48, SD = .51 \)) and English-to-Spanish translations (\( M = .32, SD = .48 \)). There was also no statistical significance between the Formal (\( M = .29, SD = .44 \)) and Informal (\( M = .53, SD = .51 \)) conditions. There was also no interaction between both independent variables, \( F(1,36) = .25, p > .62 \).

**Enforcement, Aplicabilidad de la Ley**

There was a significant effect of Direction at the \( p < .05 \) level, \( F(1, 36) = 5.49, p < .25 \), for the translation of the word Enforcement/Aplicabilidad de la Ley. Spanish-to-English translations (\( M = .62, SD = .50 \)) had a higher accuracy rate than English-to-Spanish (\( M = .26, SD = .45 \)) translations. There was no effect of Formality, \( F(1, 36) = .269, p > .61 \), meaning there was no significant statistical difference between the Formal (\( M = .43, SD = .51 \)) and Informal (\( M = .47, SD = .51 \)) conditions. There was also no interaction that could subsume the main effect of Directionality, \( F(1, 36) = 0, p > .99 \).
Adopts, Adoptar

There was no effect of Direction, $F(1, 36) = .7, p > .41$, or Formality, $F(1,36) = 2.27, p > .08$, on the accuracy of translations for the words Adopt/Adoptar. There was no significant statistical difference between Spanish-to-English translations ($M = .90, SD = .30$) and English-to-Spanish translations ($M = .80, SD = .42$). There was also no significant difference between the Formal ($M = .95, SD = .22$) and Informal ($M = .74, SD = .45$) conditions. There was also no interaction between Direction and Formality, $F(1, 36) = .02, p > .88$.

Lawful, Legítimo

There was no effect of Direction, $F(1, 36) = .07, p > .79$, or Formality, $F(1,36) = .26, p > .62$, on the accuracy of translations for the word Lawful/Legítimo. There was no significant statistical difference between Spanish-to-English translations ($M = .24, SD = .44$) and English-to-Spanish translations ($M = .26, SD = .45$). There was also no significant difference between the Formal ($M = .29, SD = .46$) and Informal ($M = .21, SD = .42$) conditions. There was also no interaction between Direction and Formality, $F(1, 36) = 1.11, p > .30$.

Amendment, Enmienda
There was no effect of Direction, $F(1, 36) = 1.83, p > 0.18$, or Formality, $F(1, 36) = 2.13, p > 0.15$, on the accuracy of translations for the word Amendment/Enmienda. There was no significant statistical difference between Spanish-to-English translations ($M = .67, SD = .48$) and English-to-Spanish translations ($M = .47, SD = .51$). There was also no significant difference between the Formal ($M = .48, SD = .51$) and Informal ($M = .68, SD = .48$) conditions. There was also no interaction between Direction and Formality, $F(1, 36) = .05, p > 0.82$.

**Apply, Aplicar**

There was a significant effect of Direction at the $p < .05$ level, $F(1, 36) = 12.06, p < .001$. Spanish-to-English ($M = .81, SD = .40$) translations had higher accuracy rates than English-to-Spanish ($M = .42, SD = .51$) translations. There was also a main effect of Formality at the $p < .05$ level, $F(1, 36) = 5.95, p < .02$, showing that the Informal condition ($M = .74, SD = .45$) had higher accuracy ratings than the Formal condition ($M = .52, SD = .51$). There was also an interaction between Formality and Direction, $F(1, 36) = 21.668, p < .00$. As demonstrated in Figure 3, the Formal condition yielded a higher Accuracy Score in Spanish-to-English, while in the Informal condition has a higher Accuracy Score when translations happened from English-to-Spanish.

*Figure 5. The Effects of Direction and Formality on Translation Accuracy Scores on the word Apply/ Aplicar*
Agreement, Acuerdo

There was no effect of Direction, $F(1, 36) = .09, p > .76$, or Formality, $F(1, 36) = .32, p > .57$, on the accuracy of translations for the word Amendment/Enmienda. There was no significant statistical difference between Spanish-to-English translations ($M = .81, SD = .40$) and English-to-Spanish translations ($M = .84, SD = .37$). There was also no significant difference between the Formal ($M = .86, SD = .36$) and Informal ($M = .79, SD = .42$) conditions. There was also no interaction between Direction and Formality, $F(1, 36) = .02, p > .90$.

Discussion

When analyzing the Composite Accuracy Scores there is no effect of Direction and Formality. However, the Composite Accuracy score was 53.7%, which is both less than expected and less than an ideal translation accuracy rate.

Although a lot of words individually did not demonstrate a main effect of Direction or Formality on Accuracy Scores, when looking at Accuracy Scores for several isolated words their scores fell far above or below the average Accuracy Score. The analysis did not account for Accuracy Scores of one word compared to the average. However, we think it is meaningful to look at the instances where these accuracy scores are significantly above or below the average because it can provide insights into word choice, translation error, and the complexity of language. Interpreting the Accuracy Scores for individual words contextualizes claims about the difficulty of translation. To this end, there are multiple cases from our study of value.
First, there is the category of words that did not have a significant effect but had high accuracy scores. For example, Adopts/Adoptar. There was no significant effect of the independent variables on the Accuracy Score of Adopts/Adoptar, but the translations had an average Accuracy Score of 84.75% across conditions. A possible explanation for this is that the words Adopt-Adoptar have a more direct line of translation that makes it easy to identify in any direction. Given the other multiple choice options Adopts looks most like Adoptar making it the more intuitive answer over the answer choices “Assume” or “Espouse”. Here, translation along similarity of appearance might track meaning, but may also prioritize accessibility given people’s expectations for word choice. Does this trend hold for other words? Can translation services prioritize this quality of translations? Another example of this phenomenon is the Accuracy Scores for the word Agreement/Acuerdo, where although there were no effects of statistical significance, there was an overall Accuracy Score of 82.5%. The accuracy of response is surprising given the numerous distractors in the answer choices, which may suggest that certain words or definitions have a more direct or identifiable translation. This would engage in larger conversations about plain language where certain words just *are* more accessible given lack of unnecessary nuance, etc. The next test of this would be to study if this is simply the word at hand or if it is a result of the mental models employed by our participants, although this distinction may not hold under further inspection.

Another important case demonstrates opposite findings to above. These instances of low Accuracy Scores among other particular words may indicate that direct translations between English and Spanish are not as simple as others to conceptualize. One example is the Accuracy Scores for the words Lawful/Legitimo. Although it did not have an effect of Direction or
Formality, the overall Accuracy Score averaged at around 25% for all conditions. The low accuracy rate is a result for participants choosing the word “Legal” as the Spanish translation for the word “Lawful”. Although “Legal” is the more intuitive translation, English-to-Spanish dictionaries identify “Legitimo” as the correct translation of the word that should be used on voting materials. Here, nuance may be captured at the expense of expected (and, as we have alluded to, intuitive) reading language. This may confirm research into there being a very real distinction in normal and formal language, although what accounts for this is left untested here.

A second instance of this problem is the word Accordance/Conformidad, where there was a lack of statistical significance but the average Accuracy Score was 25.25%. This may be explained by the fact that most respondents indicated “Acuerdo” as the direct translation of the term “Accordance”, demonstrating yet again that the more intuitive translation is not consistent with the one outlined by the English-to-Spanish translation accuracy. This also is interesting given that relationship between Agreement and Acuerdo above. On one hand, difficulty translating this word may complicate our previous discussion about the potential for a more direct or clear translation displacing the possibility for confusion. On the other, Acuerdo seems to a much more approachable word, even if people indicate that it can relate to a multitude of words.

A final trend we would like to discuss is the fact that when there is statistical significance in our data it is usually an effect of Direction, specifically indicating a higher Accuracy Score in the Spanish-to-English direction. However, the trend deviates in exactly one instance, the translation of the word Abatement/Reduccion. The fact that this is the only case of English-to-Spanish having a higher accuracy is indicative of a transformation of meaning. There is no direct translation of the word Abatement into Spanish, and this may be exacerbated by the fact that
many native English speakers may fail to understand the meaning of the word Abatement from the onset. Determining that complicated words have a low accuracy of translation should not come as a surprise, although these findings are still useful in further confirming research in other disciplines on the value of clarity. The more interesting results, however, are that in every other case of statistical significance there is a main effect of Direction with Spanish-to-English having higher Accuracy Scores.

Contest/Contiendo is one of those instances, as well as Apply/Aplicar. One possible explanation of this occurrence is that Spanish is a more complex and nuanced language that exposes speakers to the roots of words and, potentially, to their fundamental meaning. Understanding the roots of words would provide immense utility for readers making translations on the fly, and would also explain why English-to-Spanish translations lacked the same accuracy given that readers might not be looking for the same type of connection to relate words. This can be seen as a consequence of Spanish having informal words that have a translation into English that is only used in formal contexts. For example, the word “Castigar” in Spanish is used in daily conversation to mean “Punishment”. Casitgar is very similar to the word Castigate in English, which is a word rarely found in casual contexts. Different language bases to understand and approach translations points to the multifaceted issue at stake here, as well as the importance for complex and human writing/translation services that can appreciate these distinction.

By better understanding the way different cultures commonly use words (see Google Translate reliance on textual base analysis to study the way language is actually used) we can potentially understand difference in cultures and the different experiences language groups may have when voting. Simply dismissing issues raised in the study as a subsidiary of plain language
research, however, would overlook the importance discrepancies like these may have in directing future studies in multiculturalism. It is not impossible to imagine a ballot comprised of words hand picked/crafted for the English voter that may estrange minority language voters looking for cognates or roots, for example. Further, its possible that a Spanish reader may have a larger advantage (even when disadvantaged) when voting as compared to Native American voters who may not have the same access or language structure to recognize similar/related vocabulary or roots. A renewed focus on language and the role it has democratic participation, thus, is an important step towards better understanding the role or translation direction and the roots of inequity.

Limitations

First, the small size and purview of our study dramatically curtails the ability for us to draw the conclusions or answers our background and literature review would hint towards. From the onset, studying particular words is a very limited test. Particular words, while important in understanding how an individual may understand (or misunderstand) a given task or ballot option, will never capture the total experience of one in a voting booth. While some have studied the particularities of language on the voting experience (Gafke & Leuthold, 1979), the perspective this provides on inequity is only cursory. Do these words really disenfranchise someone? Does not understanding this one word really turn people away from voting? How significant is the resulting attitude shift (if there is one) from running into a deficient translation? Learning about particular words and the mental models individuals have for those words only
touched the surface of larger and more complete studies conducted by people like Redish and
Reilly that can take the long view in evaluating the voter experience.

This limitation is particularly problematic given the complex nature of language and
words. Understanding what counts as a “correct translation” in the first place is already a dubious
task, even if our test is to simply determine the accessibility of LA County’s translation glossary.
Considering how these words may be living expressions and components of political and
governmental life the pure pursuit for “fidelity” (Lessig, 1993) may also be suspect. How do we
“test” language more broadly? Research into this question has plagued philosophers for some
time -- to see some contemporary discussion, see Searle in Black (2015) -- but remains an
ambiguous project by definition. Can we produce a more perfect translation? Does this trip the
problems raised by Turfler about the dangers of ideological purity in language studies? These are
all important questions that simply fall outside of the purview of our research project and design.

Second, the method used to conduct our experiment begs important questions about how
one can track inequity. For example, all of our study’s instructions were provided only in
English, which can prime respondents and alter their approach to a given task. This has been
documented in understanding political attitudes (Hopkins, 2011) and can estrange the very
populations our study was trying to engage, further politicizing the study (Bermann, 2006). (This
limitation can be resolved moving forward by having the survey alternate instruction language
based on the participant’s response to the questionnaire’s inquiry into the participants preferred
language to read in.) Further, our studies reliance on participant’s mental models for proper word
use may provide unusable data for which to make claims about inequity. Can we judge voter
disenfranchisement by simply accumulating the right and wrong responses of individuals with
necessarily different mental models for what constitutes “normal” versus “formal” language (Nida, 1969)? Does that distinction already carry critical assumptions about inequity that can direct responses and attitudes from the onset?

Our research project concedes ground to many questions of this sort. Larger and seemingly unanswerable questions about the nature of words, language, meaning, etc. simply cannot be answered by a study dealing with particular words. This is not grounds to reject our work. Rather, and in light of this reality, one can read this report and appreciate it as engaging in a larger context to evaluate the language employed by the state. To this end our results participate meaningfully with the way translations work (or don’t) and inform larger research goals interested in the way forward towards better election materials. We do not make larger claims about the way individuals engage civically, only that to better understand the political state would require more studies like this built upon the need for vigilance (Marschall & Rutherford, 2016) and continual drive for study (Alvarez, 2002).

**Future Research**

Moving forward our research would be well served by testing these terms in a context, like a voter guide or sample ballot. Understanding words in isolation, while interesting, does not fairly represent the means by which individuals actually consume information; looking at these words in context would allow for the study and test of tone, the influence of context clues, and the comprehension-speed dynamic (Hill & Moreno, 2001) to better understand the real experiences of voters. This would also enable researchers to better track voter attitudes on
language (and election concepts) which would be key to resolving some of our limitations concerning the study of civic engagement (Tucker, 2006).

Another important area unstudied here is the role machine translations may have in evaluating the political nature of translations or translators. While this project has largely revolved around testing if the errors caused by mistranslations is the fault of translations or ballot language, one could expand the test of translations by having learning programs like those employed by Google and Microsoft to see if there is a “perfect” translation or language that avoids user misunderstanding. While this is an issue touched on by some research (see: Li et al. (2014)), little has been done to seriously evaluate these machine translations in a political context or as a tool to understand inequity. These studies will, of course, have to confront important questions concerning the political nature of services like Google Translate when evaluating the means by which a seemingly apolitical network of data participates in the production of potentially normative evaluations of user participation, understanding, and access (Fuller & Goffey, 2012). Despite this, important questions remain. Do these translations services reduce error and bias? Does eliminating bias in translation alter the balance of access in a demonstrable way?

Lastly, studies moving forward should study the language employed on election and political materials in other countries to evaluate better word choice, confusion over language readability, etc. A comprehensive study of the translations and translation services employed by organizations like the United Nations and the International Monetary Fund would prove fruitful in evaluating access and language clarity. Although there are only a handful of countries that lack an official language, the United States being one of them, studying these translations (even in the
UN, who has 6 official and working languages) will have to engage many of the preliminary questions raised in the background and study about the direction of language and the political ramifications of this priority (Olson, 1991). Does writing originally in German or Spanish influence the way translations operate moving into English, French, or Russian? Do decisions about language priority and relevance influence translation clarity and fidelity?
References


