Parent-Child Politeness Across Varieties

Delaney Ervin, Giselle Castañeda, Debbie Odufuwa

Introduction (Delaney Ervin)

During the mid-seventeenth century, European powers were colonizing across the globe, and as a result, their respective languages were also imposed upon their colonies (Smyth, n.d.). With established hierarchies of colonization, the concept of "linguistic imperialism" was born, which is a theoretical construct that explains why some languages are more dominant than others (Zeng, Ponce, & Li, 2023). And while linguistic imperialism analyzes the dominancies of different *languages*, it also begs the question of *dialectical* biases. European French and Quebec French serve as one example of real, persisting dialectical prejudice. The stigma surrounding Quebec French is not as stark as it once was, but the "attitudes on the status dimension" of Quebecois remain disfavorable in comparison to Parisian French, often being regarded as less formal and therefore less prestigious (Kircher, 2012).

Formality and prestige falls within the umbrella of pragmatics in linguistics: a field of linguistics that focuses on conversational implicatures. It involves what a speaker implies and what a listener infers (Scott-Phillips, 2017). With this in mind, we wanted to know: how does politeness in parent-child interactions vary between standard and postcolonial varieties of the same language?

To answer the question, we trained a Naive Bayes classifier on a training set from a postcolonial variety (North American English and Mexican Spanish), and then tested the classifier on these language's respective standard variety (UK English and Castilian Spanish). Then, we set out to compare this classification to word vector translations and concordance table analysis, as Naive Bayes is missing an essential feature of pragmatics: context.

The Scholarly Background (Giselle Castañeda and Delaney Ervin)

There has been a study conducted to measure the markers of politeness in English natural language. The study, "The Politeness Package: Detecting Politeness in Natural Language" by Yeomans, Kantor and Tingley used different tools to extract measures of politeness in the English natural language (2018). The linguistic markers of politeness were measured and extracted during different social settings, like speed dating and negotiations. Politeness is involved in the ways humans communicate with each other and it can be indicated through words of interaction, like 'please, thank you, you are welcome,' etc. Also, politeness in human communication also comes from social context and cultural context. Essentially, the study measured how the politeness package varies between the context of the situation and the interaction between the individuals of the situation. The journal also states that more research needs to be conducted to measure the variety of politeness in each different social and cultural contexts, and research about measures of politeness has mainly been conducted for the English natural language, thus indicating this type of research needs to be conducted for other languages in general.

It is also important to establish the different forms of politeness . There is both negative politeness and positive politeness. Negative politeness is characterized by non-imposition on others, whereas positive politeness highlights friendliness (Stephan, Liberman, Trope, 2011). Distinguishing between these forms of politeness is important for this study because it imposed a serious challenge on our implementation. Negative politeness seeks to create distance, which often takes the form of not bothering other people–and therefore an overall lower speech production. A 2019 study used Twitter corpus data to distinguish whether the stereotype that the UK preferred negative politeness while the US preferred positive politeness was true (Culpeper, O'Driscoll and Hardaker). Using corpus linguistics tools, they found that the stereotypes were only partly supported, and that positive politeness is an important feature in both UK and USA interactions.

Steps

Phase 1: Selecting a Corpus (Delaney Ervin, Giselle Castañeda, Debbie Odufuwa)

We chose CHILDES as our corpus because it contains a large volume of parent-child interactions. This corpus served as a valuable source of data for our specific question because manners and etiquette, two features of politeness and pragmatics, are emphasized in childhood.

We chose transcripts of both mealtimes and "toy play" because both of these settings require pragmatic information and include pragmatic instruction by parents. The children's behavior is valuable to observe, but even more integral to answering our research question is the parent's choice of politeness instruction in the form of examples for their children.

We utilized transcripts from the North American English (Eng-NA), the United Kingdom English (Eng-UK), and the Spanish corpora. To accomplish our classification task in English, our training and development sets were derived from the Gleason, Garvey, Bliss, and Bohannon corpuses in the Eng-NA corpora and the test set was taken from the Edinburgh corpus in the Eng-UK corpora. For Spanish, our training sets came from the Montes corpus and the development set came from the Marrero corpus. For our project, we focused on the lines said by the target child, any additional child who spoke in the conversation and any of the parents involved in the conversation. Other speakers listed in the transcript were not considered as part of our data because we wanted to focus on what was said by either the child or the parent.

Phase 2: Annotating the Transcripts (Delaney Ervin and Giselle Castañeda)

The next phase was annotating the transcripts. We chose to analyze both English and Spanish, because our group has native speakers of both languages. For each transcript, we labeled a line of speech as either not polite (NOP) or polite (POL). It is important to note that not polite does not mean impolite in the context of this project; rather, it is an utterance that does not contain an overt politeness marker.

The annotations that we assigned to the North American English set were defined by the usage of modal verbs, hedging words, manners, and terms that marked uncertainty (Scott-Phillips, 2017). All of these speech acts “manage face” through either positive or negative politeness (Clelland & Haigh, 2023). The annotations that we assigned to the Mexican Spanish set were defined by verbal cues, mannerisms, verb conjugation, cultural context, and interrogative sentences.

We annotated around 1,800 lines of North American English and Mexican Spanish. Of this, 1,100 lines were used for training the classifiers, and 700 were used for a development set. In addition to the North American English and Mexican Spanish, we selected a transcript of UK English and Castilian Spanish. These served as test sets.

Phase 3: Building the Naive Bayes Classifier (Delaney Ervin)

We chose to classify our datasets with a Naive Bayes classifier for a few reasons. While Naive Bayes lacks the sophistication of a language model, it is a simple and efficient algorithm that does not require significant monetary or environmental resources. A Naive Bayes was also ideal for this project because we wanted to avoid the biases that language models impose. Language models are trained on data from the past, so using this as our sole classifier and annotation device would reinforce and perpetuate views from the past (Sagae, personal communication, 2024). The point of this research is to analyze parent/child interactions to challenge stereotypes of politeness and formality across dialects. Using an input for our classifier that only works with old, biased data would be counterproductive and only confirm existing stereotypes.

We also could have used a Perceptron, a linear classification that divides space between two classes. However, a Perceptron uses discriminative learning that figures out the difference between two classes, and then applies this to the features. Therefore, it is better fit for binary classification than multi-class. It looks at features and then sorts them into classes, while a Naive Bayes builds a model but does not directly compare classes. In this project, we did a binary classification, so a Perceptron could have been a viable option. However, we also recognize that politeness and pragmatics are not strictly binary. We wanted this project to be buildable, with one extension being classification of several varying levels of politeness. A Naive Bayes would be easier to alter for this potential future application, so we opted for it instead of a Perceptron.

In our classifier, we calculated priors using the training sets and then tested the accuracy of the classifier using the development sets. The accuracy of the English classifier was around 95%, while the accuracy of the Spanish one was about 85%.

Then, we classified different varieties of the same language. For American English, we found a rate of 3.6% politeness-marker usage. In UK English, the classifier labelled every utterance as “NOP” (not polite). In Mexican Spanish, we found a rate of 6.7% politeness. As with Castilian Spanish, the classifier labeled each utterance as “NOP” (not polite).

Phase 4: Large Language Model Classification (Debbie)

The large language model struggled to classify the given child and parent utterances from the transcripts. Even after providing a system prompt based on strategies used by Danescu-Niculescu-Mizil et al (2013) to identify the parts of each utterance that was annotated as being polite, the language model failed to produce a decent list. For instance, given the English training set from the Gleason corpora and the utterance 'Nanette , if you look underneath your hamburg(er) , you'll find some catsup .' which was annotated as POL, the model generated 'if, you' as the words that made that utterance polite. After classifying the data, we asked the model to produce a list of the twenty words or two to three word phrases that are likely to be used by a child or parent in conversation to denote politeness. Ultimately, the output did not exactly match what criteria we used to consider something as being polite and so the language model did not perform well.

Phase 5: Word Vectors (Debbie)

Our goal for using word vectors was to determine whether the terms for politeness we identified in North American English would correspond to the terms we picked in Mexican Spanish and vice versa. Additionally, we wanted to test and see if using word vectors could translate politeness across languages. We initially attempted to have a language model, Llama-3.1 8b, generate a list of words and/or two to three word phrases that could be markers for politeness when used in conversation between a child and an adult. We planned to take that list of words, obtain the frequencies for those words based on the frequency distributions, identify five words that had the highest frequencies out of the list of words, and use those specified words to find the nearest terms in the other language and see if the results aligned with our words in the other language. However, one limitation of our project was that the terms that the language model continued to produce a frequency of zero. This was because the model tended to produce an output with more than one word but our tokens in the frequency distributions did not include bigrams. We ultimately found that using the language model was not the best approach to answer our question.

Moreover, we arbitrarily picked five markers of politeness in Eng-NA and five markers in Mexican Spanish based on the guidelines we used to annotate our datasets in order to produce a list of the top twenty most similar words and their corresponding cosine similarity values. We expected that each of these markers would be near the top of the list in the other language since the words share similar meanings which means that they should also have similar distributions.

The markers in Eng-NA included 'please', 'if', 'thanks', 'could', and 'thank you' and the markers in Mexican Spanish included 'gracias', 'agradezco', 'por favor', 'muchísimas gracias', and 'gracias por todo.' In order to consider phrases as input, we considered taking the average of the word vectors in order to predict the most similar word (Le, Q., 2014).

From the given English words, our specified Spanish markers were predicted in the list of nearest words about 40% of the time. For example, when we provided the phrase "thank you" the only word that appeared in its list was "agradecerme" which had a cosine similarity of 0.327418. On the other hand, given the Spanish markers, the English words appeared at least once in the list of nearest words about 60% of the time. To demonstrate, given the word "'gracias', the words "thanks", "thanks," and "thankyou" appeared with cosine similarities of 0.409792, 0.289152, and 0.268237 respectively. From these results, we could conclude that using word vectors did a slightly okay job at translating politeness cross linguistically.

Phase 5: Concordance Tables (Delaney Ervin)

We chose a Naive Bayes because it allowed us to annotate our own data and expand on the continuum of politeness with multiple classes for future research. However, the Naive Bayes does not account for word order, word meaning given context, or negation–all of which cannot be ignored when analyzing pragmatics.

As a solution to this problem, we analyzed the concordance tables containing “if”. This word serves as both a hedging word that manages face by softening commands or bad news, as well as a genuine marker of uncertainty (Clellan & Haigh, 2023). Understanding the context in which the marker occurred is key to knowing what purpose it serves.

In the North American English development set, “if” occurred as a politeness marker ⅙ times. In this instance, the speaker is softening their disagreement with uncertainty.

“I do n't know if that one goes there , let 's see .”

In the UK English test set, “if” occurred as a politeness marker ⅓ times as a disagreement softener, similarly to the NA English set.

“I do n't think that was I 'm not sure if that was it but + ... [+ NAC]”

The patterning between UK and NA English provides qualitative insight into the research question; both varieties are using “if” as a politeness marker. However, the number of instances in each language is too small to form any quantitative conclusions about the rate of usage.

To find a corresponding work in the Spanish test set, we used a word vector. The most similar output returned was “si”. This word occurred zero times in the Mexican Spanish data set; however, in the Castilian Spanish test set, there were 29 instances of “if”. Out of these, zero were politeness markers.

Word vectors define the space where similar things are closer together, assign probabilities, and pick the closest word. “If” serves as an interesting example of the application of using word vectors as translators because the word plays an important pragmatic role in English that it does not frequently have in Spanish, as shown by these concordance tables.

Analysis (Delaney Ervin, Giselle Castañeda, Debbie Odufuwa)

The main goal of the project was to determine how politeness in parent and child interactions vary between standard and postcolonial varieties of the same language. Given some natural language in the postcolonial varieties of language like North American English and Mexican Spanish, we wanted to see how well we could classify politeness in the standard variants such as UK English and Castilian Spanish using the Naive Bayes classifier. Overall, we found that we could not predict the standard variant very well which indicates that markers for politeness must have differed greatly across the variants.

The North American English classifier did not work on the UK English set. There are a few possible reasons for this. First, it could have been that the “POL” and “NOP” were highly disproportionate in the priors., with “POL” being 0.0197 and “NOP” being 0.980. Therefore, the classifier sacrifices recall for precision. Another reason that this classifier might not have worked is that the value of negative and politeness differs slightly between North American English and UK English (Culpeper, O’Driscoll and Hardaker, 2019). UK culture prefers negative politeness, and we only counted occurrences of positive politeness, so our results do not encompass all politeness.

Additionally, the language model classifier also did not work well on the UK English and Castilian Spanish sets and may have done a poor job since. Given more time, data for both of the test sets could have been annotated in order to calculate the accuracy, precision, recall, and F-score in order to see how it compared to the Naive Bayes Classifier.

As stated above, a weakness of our project is the limited annotated training data. All of our data was annotated by hand, so we assigned one person to each language in an attempt to maximize productivity. However, even with these efforts, our data is limited. With more time and resources, we would have: (1) annotated more North American English and Mexican Spanish for the training, and (2) annotated UK English and Castilian Spanish to calculate the classifiers accuracy on the test sets.

Another drawback of our research altogether, but of the annotation step in particular, was that we were not able to perform multiple annotations of the same datasets, which is usually necessary. Cross referencing annotations ensures it is accurate and free from bias. However, we were not able to do so due to time constraints and linguistic ability. Only one of our team members was fluent in Spanish and familiar with the pragmatics of the language.

The main unforeseen problem was the lack of polite remarks in the training, development, and test sets. Because the rates of politeness were low in comparison to non-polite remarks, the classifiers' prior probabilities were extremely skewed towards "NOP". Given more time, we would expand the annotated training data as well as the qualification of "polite". Politeness is characterized differently by different cultures, so if we had time to further research and then factor in all forms of politeness into our annotations, the accuracy of the results would be improved.

For further consideration, the number of the utterances that were classified as being polite may have been low because the markers or words that make an utterance polite may vary across languages and even within language families. Another thing to consider is that markers for politeness could have been multiple words or entire expressions and not just singular words. Additionally, some theories suggest that politeness should instead be examined using a discourse-based approach rather than at the level of words in order to consider the context of a language (Brown, 1987). Context is a very important factor to take note of, especially with regards to considering politeness, because words do not act independently. With this in mind, politeness can be examined at different levels of granularity but we decided to focus on measuring the markers of politeness in natural language using words and short word expressions in order to take advantage of the Naive Bayes classifier and assumptions to make the most of our limited data. Considering other approaches may

Division of Work

1) **Picking the Corpus**

Every group member was involved in picking the corpus. During the meeting that we decided our research question together, we also chose a corpus and files.

2) **Annotating**

Giselle Castañeda annotated the Mexican Spanish transcripts. She also researched and developed politeness markers of this variety beforehand. Delaney Ervin annotated the North American English transcripts. She researched and developed politeness markers.

3) **Naive Bayes Classification**

Delaney Ervin and Giselle Castañeda developed the English and Spanish Naive Bayes classifiers.

4) **Importing Data**

Debbie Odufuwa developed the scripts to convert the annotated data transcripts to formatted CSV files.

5) LLM Classification

Debbie Odufuwa created the language model classifier, the system prompts, examples, and messages.

6) Word Vectors

Debbie Odufuwa developed and adapted the word vector code to get the most similar words from a given phrase based on the cosine similarities.

7) Concordance Tables

Delaney Ervin created and analyzed concordance tables for each variety, using the word vectors as input.

Ethical Considerations (Giselle Castañeda)

The only potential moral and ethical issue I can predict being a major issue is not obtaining consent from the target child's parent to record and transcribe their conservation and interaction. However, since that data obtained is from CHILDES, there must have been some type of consent already established. There are no obvious biases when obtaining the data or choosing our approach for the project. All our project ideas were well thought out, collective and collaborative. I feel that one so-called "bias" is only choosing languages that we have familiarity with, such as English and Spanish. English and Spanish are very commonly spoken languages, and a plethora of linguistic research has been done on these languages, so it can be perceived that more and more research is being conducted on the dominant languages of the world versus languages that are considered to be minority languages. Specifically to the Spanish language, another foreseen bias is choosing a certain variety of Spanish based upon cultural bias and prestige. For example in the project, the Spanish data used came from México and Spain. Castilian Spanish has the cultural bias of being the superior variety of Spanish that is spoken versus other varieties of Spanish spoken in the Spanish speaking world. Since only two varieties of Spanish are represented in the data set, then maybe, a Spanish speaker who does not speak the Spanish from México or Spain can view this project in a way to be discriminative of their Spanish variety. The work from the project could potentially benefit speakers of the Spanish and English language, but it can also benefit researchers studying the Spanish and English language, and their markers of politeness in the field of sociolinguistics. It could possibly develop and further research about how one's societal culture, culture of language and customs in the Spanish and English speaking world affect the way their proper language is spoken. There is not a specific population that can be harmed from this project. Possible harm that could occur if this project is further developed and researched is politeness markers only being studied for the English and Spanish language, and is not conducted on different languages. This type of research would only represent English and Spanish, other related Germanic and Romance languages, but those languages and language families account for a small fraction of natural languages that exist. It would fall into the trend of only conducting language research on dominant languages versus underrepresented, under researched languages.

Another form of ethical consideration is the overgeneralization of "UK" speakers and "North American" speakers. Both of the varieties contain several subdialects, so referring to them as one entity might lead to erasure of varieties that do not conform to mainstream versions.

Conclusions (Delaney Ervin, Giselle Castañeda, Debbie Odufuwa)

The project set out to answer a question using a specific approach. In doing so, we also learned about the problems with using language models to answer questions that require novel inputs. If we are seeking to explore stereotypes, relying on a biased corpus full of human prejudice would only reinforce what we already know. However, this project also highlighted the usefulness of finding a balance between methodologies. While a Naive Bayes classifier was better suited for the initial phases of our project, the support of a language model was useful for translating politeness, which then provided the syntax of politeness markers in concordance tables.

Given more time, the work could be continued through expanding what "polite" means. In this work, "polite" has a fairly narrow meaning. "Polite" was set and defined by word tokens such as "could, thanks, gracias, bueno." While certain words do define politeness in a conversation or text, it is important to consider that politeness is dependent on cultural and societal context, which does affect linguistic context. In this situation, "politeness" was defined by word tokens, but also it is missing the pragmatics and context of the text.

Overall, we can conclude that classifying some phenomenon (i.e. politeness) in one language or language variant does not mean that the phenomenon will reflect or be similar in another language. If we were to perform the experiment and do the project again, one major factor to consider would be testing more pairings of language varieties in order to see how our classification results compared between postcolonial and standard variants in those language families. For instance, we could have examined the politeness difference between High German and Swiss German or even between Hindustani and Urdu in order to consider additional language families instead of only examining Indo-European and Romance languages. One final point to mention is that politeness can be interpreted differently across various languages and so it is difficult to define a true standard for what makes something polite.

References

Brown, P., & Levinson, S. C. (1987). Politeness: Some universals in language usage. Cambridge University Press. https://psycnet.apa.org/record/1987-97641-000

Clelland, H. T., & Haigh, M. (2023). Politeness and the communication of uncertainty when breaking bad news. *Discourse Processes*, *60*(7), 479–501. https://doi.org/10.1080/0163853X.2023.2245310

Culpeper, J., O'Driscoll, J., & Hardaker, C. (2019). Notions of Politeness in Britain and North America. In E. Ogiermann & P. G.-C. Blitvich (Eds.), *From Speech Acts to Lay Understandings of Politeness: Multilingual and Multicultural Perspectives* (pp. 175–200). chapter, Cambridge: Cambridge University Press.

Detecting Politeness in Natural Language, journal.r-project.org/archive/2018/RJ-2018-079/RJ-2018-079.pdf. Accessed 5 Dec. 2024.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013). A computational approach to politeness with application to social factors. Annual Meeting of the Association for Computational Linguistics. https://api.semanticscholar.org/CorpusID:12383721

Gilder Lehrman Institute of American History. (n.d.). *Early European imperial colonization in the New World.* Retrieved December 6, 2024, from https://www.gilderlehrman.org

Le, Q., & Mikolov, T. (2014, June). *Distributed representations of sentences and documents*. In International conference on machine learning (pp. 1188-1196). PMLR. https://doi.org/10.48550/arXiv.1405.4053

KIRCHER, R. (2012). *How Pluricentric Is the French Language? An Investigation of Attitudes towards Quebec French Compared to European French*, *22*(3), 345–370. doi:10.1017/S0959269512000014

Scott-Phillips T. C. (2017). Pragmatics and the aims of language evolution. *Psychonomic bulletin & review*, *24*(1), 186–189. https://doi.org/10.3758/s13423-016-1061-2

Stephan, E., Liberman, N., & Trope, Y. (2010). Politeness and psychological distance: a construal level perspective. *Journal of personality and social psychology*, *98*(2), 268–280. https://doi.org/10.1037/a0016960

Zeng, J., Ponce, A. R., & Li, Y. (2023). English linguistic neo-imperialism in the era of globalization: A conceptual viewpoint. *Frontiers in psychology*, *14*, 1149471. https://doi.org/10.3389/fpsyg.2023.1149471