Dimensions of Mind using Syntactical Information

Anonymous ACL submission

Abstract

This work captures dimensions of agency and power using syntactical features extracted from text. We derive a syntactic language measure and show that this measure highly correlates with human judgement. We then use our measure to analyze how agency and power evolved differently for male and female first names over the last 100 years using a historical corpus.

1 Introduction

Studying gender stereotypes has been an area of interest for more than a century. Recently, sociolinguistics started considering language as an important reflection of stereotypes of any kind. Examples of gender stereotypes in language include women being described more often via their body than men (Hoyle et al., 2019), and journalists quoting male newsmakers more frequently than female newsmaker (Asr et al., 2021). Furthermore, gender stereotypes are subject to change over time. (Eagly et al., 2019). This paper introduces a syntactic language measure predictive of perceived agency and patiency of an entity. We then use this measure to investigate the change in language over the last 100 years using a Historical Corpus.

There is increasing psychological evidence that humans perceive separate dimensions of mind. Gray et al. (2007) show that lab participants perceive entities to have different amounts of agency and experience. In this work, we use a large-scale survey investigating the perceived agency and patiency of entities (Ash et al., 2021). We examine whether agency correlates with syntactic features of language. This is connected to Proto-Agents and Proto-Patients (Dowty, 1991), a simplification at thematic roles in government binding theory, at the interface of syntax and semantics. To counteract role fragmentation, Dowty (1991) introduces the concept of continuous proto-agents and patients and makes an argument that these can replace thematic roles. To assess how much of a proto-agent/patient a word is, he proposes a set of questions, e.g. does the entity in question possess sentience, is there volitional involvement? Dowty predicts that syntactic subjects are more proto-agents, whereas objects are more proto-patients. Kako (2006) and Reisinger et al. (2015) both test and confirm this prediction, across linguistic experts and crowd-workers.

These questions are partially similar to how psychologists measure perceived agency and patiency of entities, see (e.g. Gray et al., 2007; Ash et al., 2021; Strohminger and Jordan, 2021). If it is true that subjects are more proto-agentic, and proto-agency is correlated with perceived agency, for example syntactic subject frequency should be predictive of perceived agency. This paper investigates this empirically, and finds that indeed these are highly correlated. Having a predictive language measure allows us to tackle philosophical and legal implications of agency and patiency. It also enables us to document how these have changed over time.

Thematic roles depend on the frame – the same syntactic arguments differ in their thematic role dependant on the verb. We can control for this using agency and power frames derived in (Sap et al., 2017), which are e.g. applied in (Antoniak et al., 2019). We find that conditioning on these frames further improves our method, and especially allows us to derive multiple language measures. This work is similar to recent work using language measures to predict perceived agency and patiency (Ash et al., 2021). While we show in our experiments that syntactic features are more accurate, they are also more data hungry and rely on high enough frequency of already sparse features.

- We derive a language measure to measure agency and patiency of entities using syntactical features

- We show that the langauge measure highly

correlates with survey respondents rating agency and patiency of entities in the lab.

- We perform a longitudinal analysis, investigating the change of agency and patiency over a decade – where our results confirm prior work on other language based measures.

2 Method

To figure out the best syntactical language measure, firstly, we considered a list of all combinations with the forms of

$$(x_1 + x_2)/(x_3 + x_4), x \in \mathcal{A} \qquad (1)$$

as possible options for our language measure. (Note: x_i can also be $\emptyset$). Additionally, we added some rational combinations which are aligned with our intuitions to work well, such as nsubj/(nsubj+subj_pass+dobj) and (nsubj+agent) / (subj_pass+dobj+subj). By choosing an objective corpus and a library for parsing it, in our case COHA and SpaCy (Honnibal and Montani, 2017) , we collect counts for all dependencies of interest for entities included in our study. Subsequently, we calculate a language agency score for each combination by the parsing results.

The next step is to figure out and validate whether one of our language-based measures for agency matches up with human judgments. To find the answer, we reused a survey, based on (Ash et al., 2021) which is derived from a large number of study participants (3181 responses). This survey contains three pre-registered subsets (31 entities, 58 entities, and all 255 entities). The survey captures human intuitions about agency and patiency for a list of 255 entities. Afterward, we compute a Spearman's correlation term between the survey score and our language score for a set of survey entities. An overview of the results is as below:

Considering Table1, we propose () as our language measure. This, not only leads us to the best option between our combinations, but also validates that the objective combination is perfectly aligned with the people's perception of the Agency dimension with a correlation of ( ). Figure 1 reports our main validation results for the preferred measure of agency from the survey and from the language model. The figure shows a scatter plot for the human and language measures of agency for the top 31 entities, including a line of best fit indicating a clear positive relationship. According to the rank correlation, the language measures are highly predictive of the human measures, with a coefficient of () which can be considered statistically significant with p < 0.005.
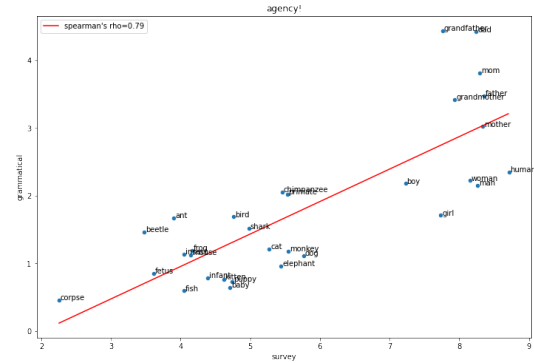

Figure 1: Correlation of Language Measure and Human Survey

3 Historical Analysis

As the first application of our measure, we apply it to a historical analysis for a completely different set of entities. To this extent, we continue to use the SpaCy library for dependency parsing with a NER accuracy of 85.85% and COHA(1900-2000) as our source corpus- containing 475 million words and balanced from decade to decade. For this part of the project, as in Figure 2, we use popular male and female names in the USA as our entities. (SSA) Thus, we collect the top 100 popular male/female names of each decade and use the average of calculated agency scores for that decade's male/female agency score (following the formula discussed in Section 2 but instead of using all COHA files, we would use an specific decade file). In advance of calculating the decades' agency score, we eliminate the redundancy of entities and fix the underlying structures.

For instance, we omitted :

- some names that might be used considerably over another concept (e.g. Angle, Rose, . . .)

- names that were gender-neutral(e.g. Alex, Taylor, . . .)

- names occurring less than 10 times as subject or object.

, and changed some dependency tags :

- regarding COHA genres, first names mostly would be in the companion of last name, therefore regarding the below parsing example, if

dependency of one entity is compound we would change its dependency to its head.

- considering some structures like the below example, if a child of dative is pobj we change the dependency label with pobj_dative

- if dependency of an entity and its head are pobj and Agent respectively, then we consider the dependency of that entity as Agent

With a not biased data in hand, we then plot female and male agency scores over time with 95% confidence intervals for each decade using the standard error of its averaged score. A comparison of the history of female and male agency scores is reported in Figure 2.
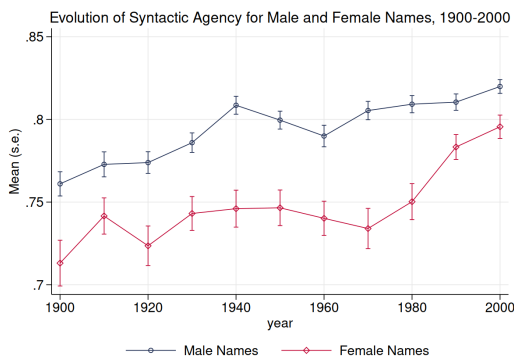

Figure 2: Longitudinal Study of Language Measure over Gender Names

4   Power Measures

As a second application of our measure, we investigate power and agency frames. We are inspired by (Sap et al., 2017) and (Antoniak et al., 2019). We count how often entities appear in power frames, agency frames and find that the resulting scores highly correlate with human judgement.

5   Conclusion

We presented an experimental framework for quantitatively studying the ways in which the language is gendered. This convergent evidence from naturally occurring language should increase our confidence that contemporary gender stereotypes thus convey a substantial female advantage in the agency dimension of mind. Moreover, in the basics, it demonstrates that the agency concept is a culturally embedded quantity that can vary over time and is not necessarily constant. Substantially,

studying this dimension in depth provides us with the opportunity to enriched our understanding of gender attributions. Our measure has opened a new window into people's latent attitudes without trying to gauge attitudes by survey from today's generation and we have already applied it over a period of history. As a vision for future study, it can also be applied across space dimension (e.g. different countries, regions,. . .) or cultural dimensions (e.g. politics, religion). Also in our study, there is a lake for the genre dimension (e.g., news, romance) of the text, even though genre is also likely to influence the language used to describe men and women, therefore it could be an area of interest for future works to subtle social attitudes and predict the direction of its change.

References

Maria Antoniak, David Mimno, and Karen Levy. 2019. Narrative paths and negotiation of power in birth stories. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW).

Elliott Ash, Dominik Stammbach, and Kevin Tobia. 2021. Dimensions of mind in semantic space. *SSRN Electron. J.*

Fatemeh Torabi Asr, Mohammad Mazraeh, Alexandre Lopes, Vasundhara Gautam, Junette Gonzales, Prashanth Rao, and Maite Taboada. 2021. The gender gap tracker: Using natural language processing to measure gender bias in media. *PLOS ONE*, 16(1):1–28.

David Dowty. 1991. Thematic proto-roles and argument selection. *Language*, 67:547–619.

Alice H Eagly, Christa Nater, David I Miller, Michèle Kaufmann, and Sabine Sczesny. 2019. Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *Am Psychol*, 75(3):301–315.

Heather Gray, Kurt Gray, and Daniel Wegner. 2007. Dimensions of mind perception. *Science (New York, N.Y.)*, 315:619.

Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.

Alexander Miserlis Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, Isabelle Augenstein, and Ryan Cotterell. 2019. Unsupervised discovery of gendered language through latent-variable modeling. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1706–1716, Florence, Italy. Association for Computational Linguistics.

3

Edward Kako. 2006. Thematic role properties of subjects and objects. *Cognition*, 101(1):1–42.

Drew Reisinger, Rachel Rudinger, Francis Ferraro, Craig Harman, Kyle Rawlins, and Benjamin Van Durme. 2015. Semantic proto-roles. *Transactions of the Association for Computational Linguistics*, 3:475–488.

Maarten Sap, Marcella Cindy Prasettio, Ari Holtzman, Hannah Rashkin, and Yejin Choi. 2017. Connotation frames of power and agency in modern films. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2329–2334, Copenhagen, Denmark. Association for Computational Linguistics.

Nina Strohminger and Matthew Jordan. 2021. Corporate insecthood.

A   Example Appendix

This is an appendix.