

Not My President: How Names and Titles Frame Political Figures

Esther van den Berg*, Katharina Korfhage[†], Josef Ruppenhofer*[‡],
Michael Wiegand* and Katja Markert[†]

*Leibniz ScienceCampus, Heidelberg/Mannheim, Germany

[†]Institute of Computational Linguistics, Heidelberg University, Germany

[‡]Institute for German Language, Mannheim, Germany

{vdberg|korfhage|markert}@cl.uni-heidelberg.de

{ruppenhofer|wiegand}@ids-mannheim.de

Abstract

Naming and titling have been discussed in sociolinguistics as markers of status or solidarity. However, these functions have not been studied on a larger scale or for social media data. We collect a corpus of tweets mentioning presidents of six G20 countries by various naming forms. We show that naming variation relates to stance towards the president in a way that is suggestive of a framing effect mediated by respectfulness. This confirms sociolinguistic theory of naming and titling as markers of status.

1 Introduction

Framing is a field of research in communication theory and political science investigating how information is presented to audiences, especially in news media. According to a common definition, to frame is to “to select some aspects of a perceived reality and make them more salient in a communication text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation” (Entman, 1993, p. 52). Most work on framing has focused on issues and events, rather than entities (Card et al., 2015; Fulgoni et al., 2016; Field et al., 2018).

We therefore introduce *entity framing*, which we define as a presentation of an entity which intentionally or unintentionally promotes a particular viewpoint towards that entity. We focus on the framing of political figures on social media, in order to better understand computer-mediated civil political discourse.

Online political discussion has been said to have an increasing influence on the democratic process, including on the tone and civility of political debates (Persily, 2017; Ott, 2017). Tweets on political themes are indeed retweeted more often when their content is emotionally charged, and espe-

cially when they contain negative appraisals of political parties and figures (Dang-Xuan et al., 2013).

We explore one way in which respect or solidarity can be expressed towards political figures: the use of their names and titles. Sociolinguistic studies have suggested that names and titles convey status or solidarity (Allerton, 1996; Dickey, 1997). Of these functions we confirm the status-indicating function on a larger scale than in sociolinguistic studies and on social media data, by demonstrating that formality in naming is positively related to the stance of tweets towards the presidents.

We thus contribute:

- a corpus of stance-annotated tweets mentioning presidents of six G20 countries, which we make publicly available.¹
- quantitative evidence of the status-indicating function of names and titles on Twitter

2 Related work

According to sociolinguists, names and titles reflect two aspects of relationships: difference in status (based on e.g. age or professional role) and degree of solidarity (also referred to as intimacy or group membership) (Brown and Gilman, 1960; Allerton, 1996; Dickey, 1997).² Studies have also observed that naming patterns are context-specific and may be violated to achieve a specific communicative purpose (Ervin-Tripp, 1972; Dickey, 1997). These studies have been qualitative and/or based on real-time observations or

¹https://www.cl.uni-heidelberg.de/english/research/downloads/resource_pages/TwitterTitlingCorpus/twitles.shtml
To follow Twitter usage guidelines, we provide tweet ids rather than tweet texts.

²These observations were made for spoken language, where names are used either as a form of address to refer to the conversation partner or as a form of reference to refer to a third party. For reasons described in Section 5, we do not make this distinction.

interviews, whereas our work examines naming quantitatively and on social media data.

The concept of framing has been applied to a variety of issues and events (Card et al., 2015; Tsur et al., 2015; Fulgoni et al., 2016; Field et al., 2018), and in one case to the framing of entities (Card et al., 2016), but not previously on social media data. Use of social media to express political opinions has instead been studied to forecast elections (Burnap et al., 2016), political mobilisation (Weeks et al., 2017), and assess political polarization (Bail et al., 2018).

A prominent area of NLP that focuses on expressions of favour is stance detection, the detection of sentiment towards a specified target. Most systems focus on stance towards products, companies and abstract topics rather than persons (Somasundaran and Wiebe, 2010; Meng et al., 2012; Jiang et al., 2011; Mohammad et al., 2016).

The datasets for SemEval 2017 (Task A and B) (Rosenthal et al., 2017) and RepLab (Amigó et al., 2012, 2013, 2014) as well as the dataset created by (Taddy, 2013) do include a variety of person entities, but no stance detection work has investigated the influence of naming on stance.

3 Data

To study how names and titles affect stance towards political figures in social media, we created a corpus of 4002 English-language tweets that mention presidents by different naming forms and which are annotated for stance.

3.1 Collection and cleaning

We focused on leaders of G20 countries with a presidential system whose names followed the order *first-name last-name*. We collected tweets between 18 June 2017 and 30 August 2017 using three query types: last-name, #first-name and first-name + (last-name/country).³ After removing duplicates, we reduced the number of headlines in the data, as headlines are bound by journalistic style conventions with respect to naming (Siegal and Connolly, 1999). We defined as a *news tweet* any tweet from an account with the string *news* in the username or description. From country subsets with an above average number of news tweets we removed the excess number.

³For example, the queries for France were *macron*, *#emmanuel*, and *emmanuel AND (macron OR france)*.

Subcorpus	FE worker agr.	Expert agr.
France	0.77	0.78
Indonesia	0.80	0.91
Russia	0.77	0.72
South Africa	0.77	0.87
Turkey	0.44	0.65
United States	0.65	0.78

Table 1: Inter-annotator agreement for the on-target/off-target task (Krippendorff alpha): agreement among FE workers and agreement between two experts adjudicating tweets where FE worker judgment was not unanimous.

Subcorpus	Adj. tweets	Diff. w/ expert 1	Diff. w/ expert 2
France	281	0.07%	0.05%
Indonesia	290	0.04%	0.03%
Russia	121	0.06%	0.06%
South-Africa	227	0.04%	0.04%
Turkey	128	0.05%	0.04%
United States	192	0.07%	0.06%

Table 2: Adjudication for the on-target/off-target task of tweets where FE worker judgment was not unanimous: number of adjudicated tweets and percentage of tweets given a different label by either expert 1 or 2 than to the FE majority vote.

Manual inspection of 50 tweets per country subset revealed that one subset consisted of very homogenous tweets, and two others contained many tweets that did not refer to the intended target. These subsets were omitted from the data.

The tweets for the remaining six subsets - France, Indonesia, Russia, South Africa, Turkey and the United States - were then automatically labeled for their naming forms. Possible labels were: first name only (FN), last name only (LN), full name (FNLN), title and full name (TFNLN) and title and last name (TLN). Oversampling rarer naming forms, we sampled 1000 tweets per country.

To remove tweets that refer to a namesake rather than the intended target, we crowd-sourced three *on-target/off-target* judgments per tweet via Figure Eight (FE)⁴. If workers could not unanimously agree whether a tweet was on-target, we collected two additional judgments from the authors (Table 1). We compared the expert judgments to the majority vote from the FE annotations and found very few differences (Table 2). We thus consider the majority vote reliable. Off-target tweets were removed from the dataset, leaving 4002 tweets.

3.2 Stance annotation

Stance-annotations of the 4002 on-target tweets were collected via Amazon Mechanical Turk

⁴<https://www.figure-eight.com>

Subcorpus	Tweets	Workers	Agreement
France	638	39	0.55
Indonesia	477	27	0.58
Russia	754	66	0.49
South Africa	698	82	0.51
Turkey	692	53	0.62
United States	743	43	0.64
Overall	4002	204	0.58

Table 3: Statistics on stance annotation after removing least reliable annotators: number of tweets, number of workers and agreement among workers (Krippendorff’s alpha)

(AMT).⁵ Workers were required to pass an English proficiency and instruction comprehension test. They had to have a minimum number of completed HITs (500), a minimum HIT approval rate (97%) and a task-internal 97% accuracy rate based on gold questions making up roughly 4% of the data. Their compensation was \$0.02 per HIT for approximately 7 HITs per minute.

Each tweet was labeled by seven annotators. Inspired by the finding in Joseph et al. (2017) that political stance annotation on Twitter suffers when too little context is shown, we provided annotators with the tweet location, user photo, user name and user description. If the tweet was a response to another tweet, that tweet was shown also. The prompt was: *How would a supporter of President X feel about this tweet?* Possible answers were: *positive (+1)*, *neither positive nor negative (0)*, *negative (-1)* and *cannot read / understand*.

Our prompt is based on the reader-perspective elicitation prompt in Buechel and Hahn (2017). We expect it to better capture differences between tweets which are neutral in tone but reflect differently on the president, such as *‘Trump trailing in primaries’* vs *‘Jobs market improving under Trump’*. Crucially, the prompt also allows annotators to give different ratings to *‘President Trump visits France’* and *‘Trump visits France’*. As in Card et al. (2015), the perspective is anchored to that of a proponent of the target in order to combat the lower reliability of reader-perspective prompts (Buechel and Hahn, 2017).

After annotation we used Multi-Annotator Competence Estimation (MACE) (Hovy et al., 2013) to identify and remove the least reliable annotators. We collected an additional two judg-

⁵We chose AMT over FE for this task so we could include a questionnaire asking for country of residence, native language, age, gender, education level, and familiarity with twitter. A study of how annotator demographic impacts annotation is planned but goes beyond this paper.

ments per tweet for the country subsets with lowest agreement (Russia and South Africa). Table 3 shows the agreement scores. The data’s gold standard was obtained using MACE, which has been shown to retrieve reliable gold labels even under very unfavourable conditions.⁶

4 Framing through naming

We now examine the relation between the use of names and titles for presidents and stance towards them in the collected tweets. Sociolinguistic work suggests that naming expresses status or solidarity. Lower status and high solidarity are both signalled with less formal naming forms such as FN, while higher status and low solidarity are both signalled with more formal naming forms like TLN (Brown and Ford, 1961; Allerton, 1996; Dickey, 1997).

This dual social function gives rise to two possible main relations between naming and stance corresponding to the following hypotheses:

H0 Variation in naming and stance are not related.

H1 Naming primarily downplays or emphasises the president’s **status**. Therefore, formality of naming is positively related to stance.

H2 Naming primarily conveys the degree of **solidarity** with the president. Therefore, formality of naming is negatively related to stance.

Table 4 gives examples of tweets which can be interpreted to support either H1 or H2, or to support the existence of alternative, context-specific functions of naming, such as sarcasm.

We group tweets country-independently by naming form and perform a Kruskal-Wallis test of the difference in average stance. This reveals a statistically significant difference between the stance of tweets with different naming forms ($\chi^2(4)=424.67$, $p<0.001$). A post-hoc Dunn’s test with Bonferroni correction shows statistically significant differences between all naming forms ($p<0.001$) except for LN and FN, possibly due to the small size of the FN group. We reject the null-hypothesis that naming and stance are not related.

To examine which alternative hypothesis is more likely between H1 and H2, we rank the naming forms according to their formality:

⁶We conducted experiments with synthetic data to verify that MACE was likely to obtain a reliable gold standard from our data.

Function	Stance	Form	Tweet text
status	pos	TFNLN	<i>Dear President Joko Widodo, Happy Birthday. God bless you @jokowi</i>
status	neg	FN	<i>That's the truth!!! Double-standard #Donald at it again</i>
solidarity	pos	LN	<i>Duterte & Widodo are truly public servants. Saving their countries fr the menace of society.</i>
solidarity	neg	TLN	<i>President Trump probably won't like next week's newsstands</i>
sarcasm	neg	TLN	<i>Of course, I know, everything is sweetness & light in the wonderful democratic Paradise of President Erdogan!</i>

Table 4: Possible examples of a *status* or *solidarity* function of naming forms in tweets, as well as of an alternative function.

Subcorpus	FN	LN	FNLN	TLN	TFNLN
France	0.00 (10)	-0.29 (377)	-0.08 (117)	-0.04 (80)	0.04 (54)
Indonesia	-0.60 (15)	-0.03 (134)	0.14 (167)	0.08 (50)	0.37 (111)
Russia	-0.56 (54)	-0.71 (442)	-0.31 (122)	-0.26 (74)	0.24 (62)
South Africa	-0.50 (6)	-0.53 (405)	-0.40 (109)	-0.08 (106)	0.18 (72)
Turkey	-0.75 (4)	-0.67 (440)	-0.23 (124)	0.06 (82)	-0.17 (42)
United States	-0.80 (59)	-0.53 (363)	-0.50 (141)	0.15 (94)	0.03 (86)
Overall	-0.63 (148)	-0.52 (2161)	-0.21 (780)	-0.02 (486)	0.14 (427)

Table 5: Average stance and in brackets the absolute number of tweets containing naming forms from least to most formal.

FN < LN < FNLN < TLN < TFNLN.⁷ Table 5 shows that the average stance of tweets increases with each increase in formality. A Spearman's rank-order correlation test confirms a statistically significant positive correlation between naming formality and stance ($r_s(4002) = .32, p = .001$).

Furthermore, a chi-square test shows that the difference between the stance of tweets with and without a title in them (Table 6) is significant for each of the six subcorpora ($p < 0.05$).

These findings support the *status* hypothesis: due to naming mainly indicating status, status-indicating function of names, formality in naming is positively related to stance.

5 Discussion

Although we show a clear framing effect of naming and titling, our study has several limitations. First, we do not distinguish between address and reference. Our data contains both names used as forms of address (e.g. '*Making things "Great Again" huh #Donald?*') and as forms of reference (e.g. '*#Donald just cant handle competing for the title.*'). Studying these types separately would require additional manual annotation. In addition, this distinction is not as clear for Twitter data as for face-to-face conversations, as many tweets mix both functions.

Second, some of the naming forms occur only rarely, particularly FN. This hinders the finding of

significant differences between each of the naming forms for each individual country subset. Nevertheless, a significant difference in stance could be observed between tweets with and without titles in each subcorpus.

Third, we consider tweets from a limited time span. This means the content of the tweets and therefore the naming used in them may be influenced by the occurrence of specific events (e.g. Joko Widodo's birthday).

Fourth, we only consider English tweets. Tweets about presidents which are not well-known to native English speakers may be unrepresentative of local ways of referring to the president. They may also be more neutral in tone and may use (T)FNLN to be informative rather than respectful.

These limitations as well as certain social media/Twitter-specific properties (the character limit, the often unspecified audience) increased the chance that any primary function of naming would be lost among noise. It is therefore interesting to still see the clear trend across country subsets that informal naming of presidents co-occurs with perceived hostility, while formal naming co-occurs with perceived supportiveness of a tweet. This suggests that in tweets on politicians naming primarily emphasises status and conveys respect.

6 Conclusions and future work

We present an analysis of the way political figures are named in social media and how this naming relates to stance in a corpus of stance-annotated tweets mentioning presidents of six G20 countries.

⁷Based on the following criteria:

- 1) Naming with title is more formal than without title.
- 2) Longer names are more formal than shorter names.
- 3) Last names are more formal than first names.

Subcorpus	Without title	With title
France	-0.24 (504)	-0.01 (134)
Indonesia	0.03 (316)	0.28 (161)
Russia	-0.62 (618)	-0.03 (136)
South Africa	-0.50 (520)	0.03 (178)
Turkey	-0.58 (568)	-0.02 (124)
United States	-0.55 (563)	0.09 (180)
Overall	-0.45 (3089)	0.05 (913)

Table 6: Average stance and in brackets the absolute number of tweets without or with a title in their naming form.

Our analysis reveals a relation between the formality of names and the stance of tweets. More formal forms are significantly more frequent among positive tweets than less formal ones.

We thus confirm sociolinguistic claims that naming marks status and expresses respect that had not previously been investigated in a large, quantitative study, nor for social media texts. This study also represents the first approach to entity framing by providing evidence for a framing effect of naming.

Future work should investigate whether naming forms in address vs. reference impact stance differently, whether naming form usage differs depending on demographics and whether the naming trends found across the time span of our tweets can also be found across a longer time span. Also valuable would be a study of this effect in other languages and on different politician subgroups, such as female politicians. Studies such as [Uscinski and Goren \(2011\)](#) suggest that titles of female politicians are omitted more frequently and with different effect.

NLP work can use our corpus as further data for stance detection. Experiments in [Mohammad et al. \(2016\)](#) show that cross-target stance detection is very challenging. Our corpus can provide further training and testing data both for in-target and cross-target classification.

Acknowledgements

We thank the reviewers for their insightful comments. This research is funded by the Leibniz ScienceCampus Empirical Linguistics& Computational Language Modeling, supported by Leibniz Association grant no. SAS2015-IDS-LWC and by the Ministry of Science, Research, and Art of Baden-Wuerttemberg.

References

- David J Allerton. 1996. Proper names and definite descriptions with the same reference: A pragmatic choice for language users. *Journal of Pragmatics*, 25(5):621–633.
- Enrique Amigó, Jorge Carrillo-de Albornoz, Irina Chugur, Adolfo Corujo, Julio Gonzalo, Edgar Meij, Maarten de Rijke, and Damiano Spina. 2014. Overview of RepLab 2014: Author profiling and reputation dimensions for online reputation management. In *Proceedings of the Fifth International Conference of the CLEF Initiative*, pages 307–322.
- Enrique Amigó, Adolfo Corujo, Julio Gonzalo, Edgar Meij, Maarten de Rijke, et al. 2012. Overview of RepLab 2012: Evaluating online reputation management systems. In *CLEF 2012 Evaluation Labs and Workshop, Online Working Notes*.
- Enrique Amigó, Jorge Carrillo De Albornoz, Irina Chugur, Adolfo Corujo, Julio Gonzalo, Tamara Martín, Edgar Meij, Maarten De Rijke, and Damiano Spina. 2013. Overview of RepLab 2013: Evaluating online reputation monitoring systems. In *Proceedings of the Fourth International Conference of the CLEF Initiative*, pages 333–352.
- Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221.
- Roger Brown and Marguerite Ford. 1961. Address in American English. *Journal of abnormal and social psychology*, 62(2):375–385.
- Roger Brown and Albert Gilman. 1960. The pronouns of power and solidarity. In Thomas A Sebeok, editor, *Style in Language*, pages 253–276. MIT Press, Cambridge, MA.
- Sven Buechel and Udo Hahn. 2017. [Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis](#). In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, volume 2, pages 578–585.

- Pete Burnap, Rachel Gibson, Luke Sloan, Rosalynn Southern, and Matthew Williams. 2016. 140 characters to victory? Using twitter to predict the UK 2015 General Election. *Electoral Studies*, 41:230–233.
- Dallas Card, Amber E Boydston, Justin H Gross, Philip Resnik, and Noah A Smith. 2015. [The media frames corpus: Annotations of frames across issues](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, volume 2, pages 438–444.
- Dallas Card, Justin Gross, Amber Boydston, and Noah A Smith. 2016. [Analyzing framing through the casts of characters in the news](#). In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1410–1420.
- Linh Dang-Xuan, Stefan Stieglitz, Jennifer Wladarsch, and Christoph Neuberger. 2013. An investigation of influentials and the role of sentiment in political communication on Twitter during election periods. *Information, Communication & Society*, 16(5):795–825.
- Eleanor Dickey. 1997. Forms of address and terms of reference. *Journal of linguistics*, 33(2):255–274.
- Robert M Entman. 1993. Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58.
- Susan Ervin-Tripp. 1972. On sociolinguistic rules: Alternation and co-occurrence. *Directions in sociolinguistics*, pages 213–250.
- Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. [Framing and agenda-setting in Russian news: a computational analysis of intricate political strategies](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580.
- Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preotiuc-Pietro. 2016. An empirical exploration of moral foundations theory in partisan news sources. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, pages 3730–3736.
- Dirk Hovy, Taylor Berg-Kirkpatrick, Ashish Vaswani, and Eduard Hovy. 2013. [Learning whom to trust with MACE](#). In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1120–1130.
- Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. 2011. [Target-dependent twitter sentiment classification](#). In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 151–160.
- Kenneth Joseph, Lisa Friedland, William Hobbs, David Lazer, and Oren Tsur. 2017. [Constance: Modeling annotation contexts to improve stance classification](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1115–1124.
- Xinfan Meng, Furu Wei, Xiaohua Liu, Ming Zhou, Sujian Li, and Houfeng Wang. 2012. Entity-centric topic-oriented opinion summarization in twitter. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 379–387. ACM.
- Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. [Semeval-2016 task 6: Detecting stance in tweets](#). In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 31–41.
- Brian L Ott. 2017. The age of Twitter: Donald J. Trump and the politics of debasement. *Critical Studies in Media Communication*, 34(1):59–68.
- Nathaniel Persily. 2017. The 2016 US election: Can democracy survive the internet? *Journal of democracy*, 28(2):63–76.
- Sara Rosenthal, Noura Farra, and Preslav Nakov. 2017. [Semeval-2017 task 4: Sentiment analysis in Twitter](#). In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 502–518.
- Allan M Siegal and William G Connolly. 1999. *The New York Times manual of style and usage*. Three Rivers Press (CA).
- Swapna Somasundaran and Janyce Wiebe. 2010. [Recognizing stances in ideological on-line debates](#). In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124.
- Matt Taddy. 2013. Measuring political sentiment on Twitter: Factor optimal design for multinomial inverse regression. *Technometrics*, 55(4):415–425.
- Oren Tsur, Dan Calacci, and David Lazer. 2015. [A frame of mind: Using statistical models for detection of framing and agenda setting campaigns](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1629–1638. ACL.
- Joseph E Uscinski and Lilly J Goren. 2011. What’s in a name? Coverage of senator Hillary Clinton during the 2008 democratic primary. *Political Research Quarterly*, 64(4):884–896.
- Brian E Weeks, Alberto Ardèvol-Abreu, and Homero Gil de Zúñiga. 2017. Online influence? Social media use, opinion leadership, and political persuasion. *International Journal of Public Opinion Research*, 29(2):214–239.