Willkommens-Merkel, Chaos-Johnson, and Tore-Klose:
Modeling the Evaluative Meaning of German Personal Name Compounds

Annerose Eichel[1], Tana Deeg[1], Andre Blessing[1], Milena Belosevic[2],
Sabine Arndt-Lappe[3], Sabine Schulte im Walde[1]

[1] Institute for Natural Language Processing, University of Stuttgart
[2] Faculty of Linguistics and Literary Studies, Department Linguistics, University of Bielefeld
[3] English Linguistics and Trier Center for Language and Communication, Trier University

LREC-COLING 2024

Introduction

Personal Name Compounds (PNCs)

modifier → **Willkommens-** **Merkel** ← personal name
Welcome-Merkel

* Understudied from both a theoretical and computational perspective
* Important for understanding and generating texts from and for domains such as the news, social media, and political discourse

Research Questions

* Are PNCs **perceived as more positive or negative** compared to the respective full name?
* **How to model and measure** evaluative meaning computationally?
* What is the **impact of modifier meaning**?
* Do **factors such as age, party membership, and compound valence** influence PNC evaluation?

Data

Targets

321 German PNCs
Common/proper noun modifier + personal name
found in 1+ contexts

131 Names
Full name
corr. to PNCs

Represented Domains
Politics > Sports > Show Business > Others

Example Contexts

#Lanz This political constellation should never have come about in the first place, says Merz. Another declaration of war on Welcome-Merkel

Excellent! - "Pedal-to-the-Metal-Vettel" saves World Cup lead #Vettel

Context Corpora

Social Media
𝕏
100 tweets / target

PNC: 9,145 tweets
Name: 24,688 tweets

News Data
DWDS
27M sentences

PNC: 170 sent.
Name: 233,477 sent.

Modeling Approaches

Idea: Use **valence as a proxy** to capture the evaluative meaning of PNCs

Valence Norms

* Automatically generated norms for German by Köper and Schulte im Walde, 2016

Fine-tuned PLMs

* Base models (XLM-RoBERTA or BERT) + different fine-tuning goals (multilingual vs. monolingual, general vs. politicians' tweets, twitter vs. additional data)

Human Evaluation

* 5 German native speakers annotate 10% of data in context

Idea: Explore influence of **personal background information** on PNC meaning

Data Enrichment

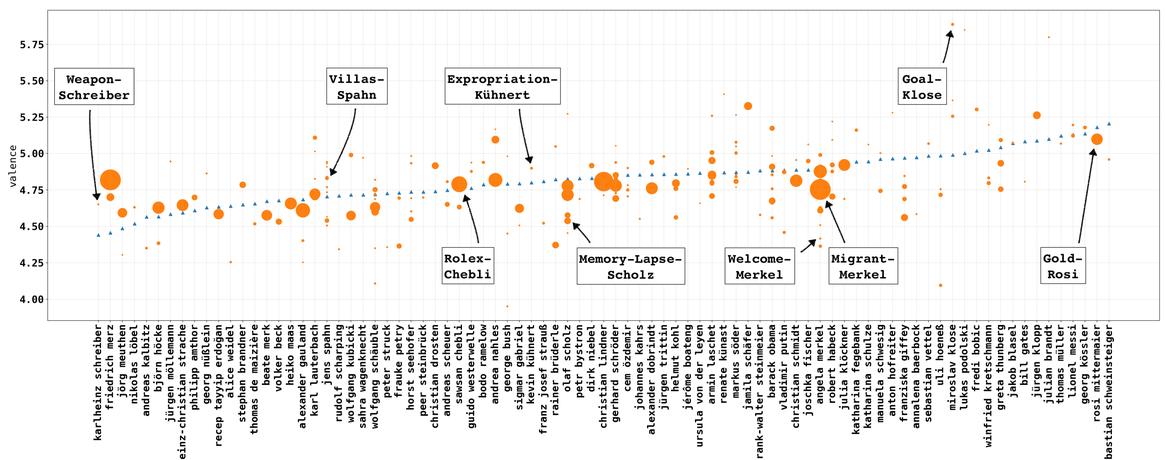
* Domain, age, nationality, birthplace, gender, political party, event frame, valence

Regression Modeling

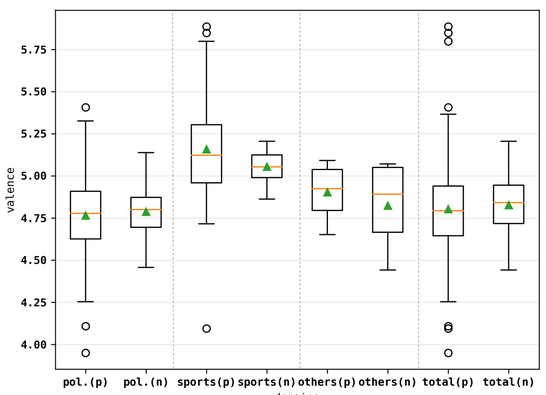
* univariate – multivariate – variable selection using Elastic Net

Capturing the Evaluative Meaning of PNCs

Valence frequency across domains captured through **valence norms**



Valence frequency across domains captured through **valence norms**

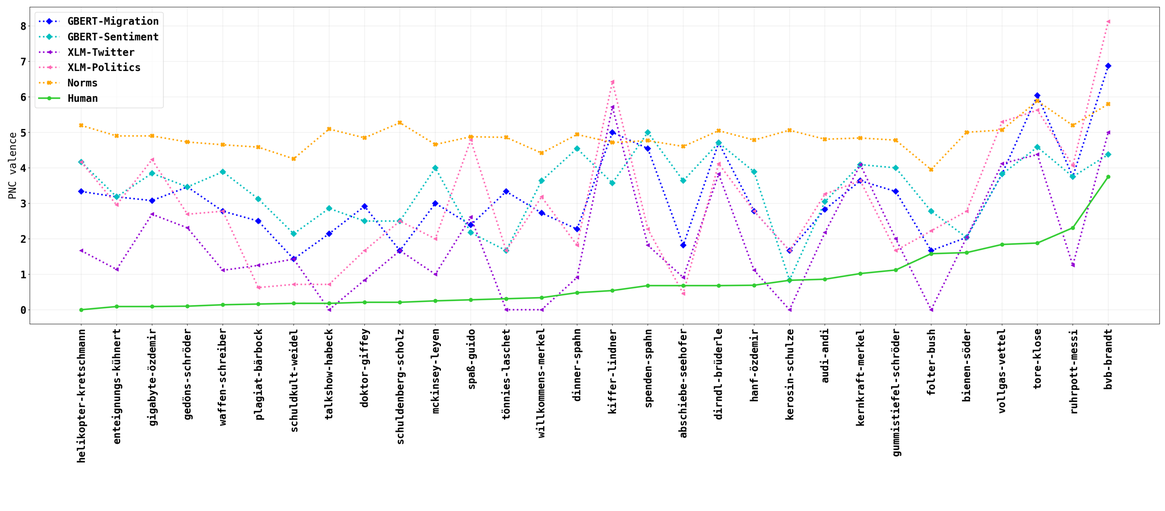


Cross-domain results using **PLMs**

	$\Delta > 0$	$\Delta < 0$
XLM-Twitter	90.32	9.68
XLM-Politics	82.49	17.51
GBERT-Sentiment	93.55	6.45
GBERT-Migration	89.86	10.14

Overview of **relative difference values** (Δ) **between PNC and name valence**
(Δ < 0: PNC bears a more negative meaning; Δ > 0: PNC more positively perceived than resp. name)

Comparison of PNC valence determined by **humans vs. PLMs vs. valence norms**



Univariate Linear Regression

* **PNC valence** highly significant
* **Modifier valence:** positive linear relationship
* **Age** shows significant inverse relationship
* **Party membership**: AfD members more likely to have negative relation

Multivariate Linear Regression

* **PNC valence** remains most relevant predictor
* **Male gender** – slightly more positive evaluation
* **US birthplace** hints towards more neg. relation
* **Political party membership:** inverse relations for CDU/CSU, FDP, The Greens, SPD

Conclusion

* First comprehensive computational exploration of modeling PNC meaning
* Developed two approaches based on valence norms and PLMs and conducted a human annotation study to compare results within and across domains
* Explored influence of personal background information on PNC evaluation