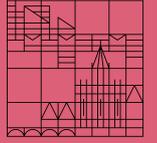


Inequality

magazine

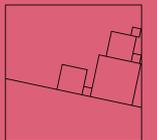
N°5

Universität
Konstanz



Information, Language, Power

Cluster of Excellence
The Politics of Inequality



In_equality magazine

**Information,
Language, Power**

*Research Magazine of the
Cluster of Excellence
“The Politics of Inequality”
at the University of Konstanz*

6 Editorial
36 News
46 Imprint
52 Perspectives

CONTEXTUALIZED

8 Language and Inequality.
 On the Relation between
 Linguistics & Social Sciences

(M. Butt, R. Eckardt)

INSIGHTS FROM OUR RESEARCH

12 Language Matters.
 Inequality amongst the Sámi
 Minority in Norway & Sweden

(T. Kupisch, A. Lloyd-Smith, F. Bergmann, R. Yasar)

INSIGHTS FROM OUR RESEARCH

18 Targeting the Voter. How
 Italian Political Parties Phrase
 Identity Appeals (S. Zanotto)

INSIGHTS FROM OUR RESEARCH

24 Language and Power.
 Computational Analysis
 Can Help Us Understand
 Inequalities in Communication

(S. Eckhard, I. Espinoza, S. Frenzel, A. Hautli-Janisz, W. Siskou)

TALES FROM THE LAB

30 All You Need Is Love?
 Politicians' Charisma
 and the Effect of Emotional
 Messages (J. Vári)

INSIGHTS FROM OUR RESEARCH

42 Subtle Messages.
 Artificial Intelligence
 Enhances Research on
 Framing, but Is still Far
 from Perfect (Q. Yu, A. Fliethmann)

WIDE ANGLE

48 Nations, Languages,
 and Speakers. Different
 Ways of Understanding
 Native Tongue and Its Role
 for Communities (J. Errington)

Table of Contents



Gabriele Spilker is Professor of International Politics—Global Inequality and Co-Speaker of the Cluster.

Marius R. Busemeyer is Professor of Political Science at the University of Konstanz and Speaker of the Cluster of Excellence “The Politics of Inequality.”

Claudia Diehl is Professor of Microsociology at the University of Konstanz and Co-Speaker of the Cluster.

Editorial

DEAR READERS

Language is more than a medium of everyday communication—language is an expression of identity and belonging, a tool in political struggles, and a stylistic device that can create charisma. By the way we speak and by what we (don’t) say, we address certain groups and exclude others. Thus, language is a tool that can generate inequality and render it (in)visible. That is why we dedicate the fifth issue of the *In_equality Magazine* to this complex constellation: information, language, and power.

Linguists in our Cluster explore multiple aspects of *language and inequality*. One example is language extinction: Almost every 14 days, one of the world’s estimated 7,000 languages loses its last native speaker. Language death not only means a loss of cultural diversity but also a major loss of identity for the communities affected. Often, they are forced to act in a language other than their mother tongue, thereby suffering social and economic disadvantages. This concerns indigenous populations as well as politically persecuted people or migrants.

Another field where language is key is artificial intelligence (AI) and new related forms of communication. Specifically, ChatGPT has been dominating the headlines lately, and AI-assisted (analytical) methods are now a standard component of our research as well. In linguistics, AI is used primarily in computational linguistics to study inequality. Cluster researchers use it to ana-

lyze large amounts of text, study the impact of subtle differences in rhetorical devices, or examine how power is expressed in language. Nevertheless, for all the advantages of computer-assisted linguistic analysis, it is also necessary to keep in mind its error-proneness and its inability to understand the subtleties of human speech. That is why we believe it is essential to have *linguistics* on board as a discipline to help explore the political dimensions of inequality in our Cluster.

Speaking of onboarding, the Cluster recently welcomed many new colleagues. The managing office saw particularly strong staff turnover in the past months. We are very pleased that Thomas Wöhler took over the management of our Cluster in March of this year.

Not least because of these staff changes, this issue of our magazine is published somewhat later than usual, but just in time for the summer break. We hope you enjoy reading it—on your beach vacation, on a balmy summer evening, or to entertain yourself during lonely lunch breaks in the office. /

Yours
MARIUS R. BUSEMEYER,
CLAUDIA DIEHL &
GABRIELE SPILKER

The current issue of the In_equality magazine deals with “Information, Language, Power.” *But how are these major themes interconnected? And why does social science research need a linguistic perspective?* A contextualization by linguists Miriam Butt and Regine Eckardt.

CONTEXTUALIZED

8

Language and Inequality. *On the Relation Between Linguistics & Social Sciences*

(M. Butt, R. Eckardt)

Political science is a discipline that is intimately concerned with communication: communication between those governing and the governed, communication between politicians or elected officials and voters, as well as the determination of which language the government communicates in the first place, to name just a few (but major) areas. And yet political science does not make knowledge of linguistics mandatory nor is the use of linguistic insights and tools widespread in political analysis. The cluster addresses this odd disconnect by bringing linguists into the mix of researchers who are otherwise primarily from the social sciences.

As can be seen from the excerpt of linguistically focused work in this magazine, linguistic methodology goes far beyond rough-and-ready methods such as sentiment analysis, word counts, and topic modelling. These are often thought of as linguistic because they involve words, but in fact they involve next to no deep linguistic insights and eschew an understanding of linguistic structure. For example, sentiment analysis finds its roots in commercially oriented applications designed to gauge customer reviews (positive vs. negative). The approach equates to a “bag-of-words” methodology: all words in a text are treated as items in a bag, without any structure between them and the text is ranked according to the sentiment associated with individual words. The method is very error-prone because it does not take any sort of linguistic structure (e.g., what word is modifying which and how) into account and also does not deal with logical relations, i.e. realizing that “not wonderful” is actually a negative sentiment of a kind.

While some work presented in this issue might use rough-and-ready tools like sentiment analysis as a first cut at a problem, linguistic analysis typically goes much deeper in examining what meaning is explicitly conveyed as part of a communication and what meaning is conveyed implicitly through what is said and, often more importantly, through what is not said. What matters in public debate is not only what one talks about and what one does not talk

about (topic framing) but also how one phrases things. For example, it has long been shown that if some medical measure is described as “80 % will survive,” it is likely to receive more support than if it is described as “20% will die” even though these statements are logically equivalent: it is a matter of framing.

Linguistic cues generate opinionated messages at a more subtle level than that captured by sentiment analysis. If one describes a measure as “hard, but necessary” one appeals to the addressee’s ability to endure. However, the same measure presented as “necessary, but hard” signals that hardships have been noted and will be taken into account in formulating policies. In this vein, Germany’s vice chancellor Robert Habeck’s recent attempt to sell his green energy plan might have done better to take such flavorings into account. Also important are the speaker’s underlying points of comparison in the discussion of an issue. Take the issue of refugees: if a nation is likened to a vessel, it sounds rational to declare that “the boat is full” and to refuse refugees entry. Nations, however, could equally fittingly be compared to companies, and in that setting slogans like “the bigger, the better” open up very different political avenues for debate. The project on Framing Inequality studies the migration debate in Germany and Europe, aiming to integrate different strata of opinionated language in a single model of framing to address just such issues. →

9



↑ The Microphone as a Symbol: Linguistic insights should be more integrated into social science research, argue the authors of this article.

Successful politicians (and diplomats) are generally skilled communicators and framers. This can also involve strategies such as not answering direct questions when the answer might be detrimental to their cause (a famous example is Madeleine Albright's disastrous response to the direct question if it was overall worth it to have half a million Iraqi die because of US sanctions), finding the right vocabulary and level of complexity in addressing voters, and finding the right dialect or accent to address them in. Famously, Margaret Thatcher shed her native accent in favor of English Received Pronunciation (the upper class standard) to sound more intelligent, convincing and trustworthy. In contrast, Nicola Sturgeon's pronounced Scottish accent undoubtedly plays to her voter base. There is also a general perception in the UK that people with Scottish accents are more trustworthy and charming, but whether this is based in actual fact and what makes up the overall charisma of a politician in the first place is not really known. The charisma project investigates charismatic communication by using a wide array of computational and experimental methods. The issue of accents and the use of dialects or minority languages also arises centrally in the work on the role of language in nation building (language policy) and the treatment of minority languages and its effect on speakers, for example with respect to the Sami speakers in Scandinavia as investigated in the cluster.



Miriam Butt is a Professor of General and Computational Linguistics at the University of Konstanz and a PI at the Cluster.



Regine Eckardt is a Professor at the Department of Linguistics at the University of Konstanz and a PI at the Cluster.

Beyond the mode of delivery, the content of politicians' speeches is, of course, of interest. One project focuses specifically on differences in communication strategies across Italian political parties. Another project investigates how closely the language of governments matches that of the governed, a question taken up in a very innovative research paradigm on interactions between civil servants and the public in contexts of public services. Using sophisticated computational methodology, the project dives deeply into issues of language structure such as sentence and vocabulary complexity, cooperative vs. uncooperative communication and powerful vs. powerless discourse strategies.

And what about large language models (LLMs) such as ChatGPT, which is currently making headlines? Rather than reiterating what these systems cannot do (e.g., reason, distinguish truth from falsehoods), we want to warn politicians to be aware of possibilities of abuse. LLMs are basically "stochastic parrots," a term coined by our colleague Emily Bender (University of Washington) and co-authors. This means that they are very sophisticated predictive machines trained on language data. LLMs can thus imitate and multiply whatever mode of speech/text they are trained on. They can be used to simulate online supporters of whatever political position and parrot whichever rhetorical means to push any program. Needless to say that such "utterances" don't reflect new majorities, as little as they announce

new truths. All the more important is what is not usually discussed in the media, like LLM's use of immense amounts of CO2 producing resources, matters of ethics and the increase of the digital divide with respect to majority vs. minority languages (such as Sami). These are questions that are extremely relevant from an inequality perspective—perhaps they are to be put on our agenda. /

Members of the Sámi minorities in Norway and Sweden who identify with their native culture experience discrimination, especially when they use the Sámi language in public. While this is true in both countries, a comparison shows that specific minority policies affect the level of inequality. *Spending more on the enhancement of Sámi language and culture—as Norway does—leads to a higher level of (self) esteem and equality.*

12

Linguistic inequality refers to the idea that languages and their speakers are not evaluated equally from a social and political perspective, often with the practical consequence that speakers of “less important” languages do not receive appropriate support and treatment. Worldwide, minorities face barriers to using their languages, which delimits their educational and economic success. Children of a minority may never learn to read in their native language, or, in extreme cases, not even understand what their teacher is saying. The reasons for linguistic discrimination are often political.

INSIGHTS FROM OUR RESEARCH

Language Matters.

Inequality amongst the Sámi Minority in Norway & Sweden

(T. Kupisch, A. Lloyd-Smith, F. Bergmann, R. Yasar)

Today, linguistic rights are seen as human rights. The United Nations Declaration on the Rights of Indigenous Peoples states that “Indigenous peoples have the right to establish and control their educational systems and institutions providing education in their own languages, in a manner appropriate to their cultural methods of teaching and learning” (Article 14, 1). Parents and teachers are increasingly conscious of the intimate connection between an individual’s mother tongue and identity formation. Nevertheless, linguistic diversity is threatened: only half of the world’s estimated 7,000 languages might survive the next century. A language is endangered when it is not passed on to children as a first language. As older speaker generations pass away, fewer speakers remain.

Linguistic inequality is not only a matter of the Global South (e.g., former colonies where the colonizers imposed their languages on the local people); it also happens in the Global North, where it is less expected. A case in point are the indigenous Sámi. Sámi languages are spoken in the arctic regions of Finland, Norway and Sweden, as well as the Kola Peninsula in Russia. Until the 1950s, the Sámi faced harsh discriminatory policies and assimilation pressure. As a consequence, all nine remaining Sámi languages are endangered, with less than 30,000 speakers. None of them are monolingual and few are fluent. Low proficiency levels pose a problem for language transmission to the next generation. Over the past →

13



↑ The colors of the Sámi flag symbolize different elements in life and are also found in the traditional Sámi costume

Figure 1: Maps of municipalities surveyed in the Nordic Peoples Survey, and proportion of respondents with Sámi ethnicity

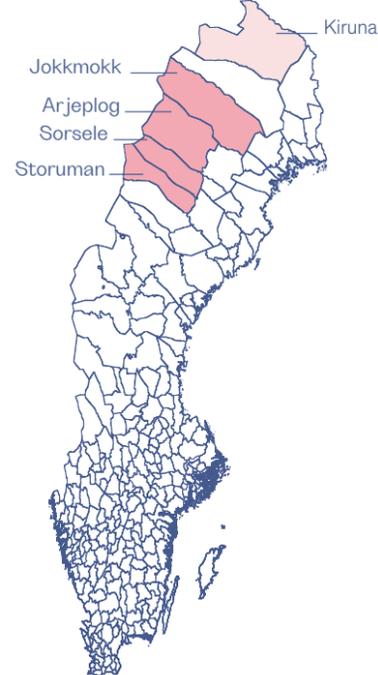
Norway

- 0 to 25%
- 25 to 50%
- 50 to 75%
- 75 to 100%



Sweden

- 0 to 25%
- 25 to 50%
- 50 to 75%
- 75 to 100%



decades, Sweden and Norway have employed strategies to revitalize the Sámi languages, including the ratification of the UN Declaration and the European Charter for Regional or Minority Languages and state-run schooling through the medium of Sámi in some municipalities. Despite being commendable, these measures have not yet succeeded in reversing the effects of discriminatory policies in the past. Speaker numbers continue to drop. Feelings of inequality remain.

In the “Nordic Peoples Survey” we reached out to adult residents in 20 northern municipalities in Norway and Sweden (Figure 1): 2,396 people in Norway, 44.7% of whom are ethnic Sámi—self-identifying as Sámi or reporting Sámi ancestry—and 3,020 in Sweden (28% ethnic Sámi). Language use across generations has dropped dramatically. While around 75% of Sámi respondents’ grandparents were Sámi users, just 20% of today’s adult generation uses a Sámi language (occasionally); only 11.3% (Norway) and 4.18% (Sweden) consider their proficiency “native-like.”

In addition to this worrisome linguistic situation, we found in our survey distinct perceptions of social inequality among the Swedish Sámi. When asked where they see themselves on a ladder representing society, self-identifying Swedish Sámi place themselves at a significantly lower position than respondents from the majority population. The Sámi and non-Sámi respondents, however, do not report different income levels, suggesting that these perceptions are not of economical nature. When asked where they, as an ethnic group, would be positioned on a ladder that reflects wider society’s esteem and respect for certain groups, Swedish Sámi position themselves lower than where the majority positions them.

In Norway, no such differences were found. The differences between Sweden and Norway regarding language proficiencies and perceptions of inequality might reflect the respective policies towards the Sámi. While both countries have taken measures to protect the languages, the decreasing numbers of speaker highlight the need for further efforts. Norway seems to act more resolutely: Norwegian and Sámi were given equal status (in some administrative areas), while in Sweden, the Sámi languages are categorized as minority languages (alongside Finnish, Meänkieli, Romani, Yiddish). School education through the medium of Sámi is more readily available in Norway.

There are also differences in expenditure for Sámi culture by the Norwegian and the Swedish Sámi Parliaments (popularly elected institutions with representative tasks, structured as a government agency). The Norwegian Sámi Parliament spends more on language and culture, and the Sámi in Norway are more integrated in the national decision-making processes. Thus, Norway’s policies appear to confer a higher social status to the Sámi, creating higher esteem for Sámi issues, which might lead to more equal status perceptions. Nevertheless, feelings of inequality remain an issue in both countries.

Since household income showed no economic difference between Sámi and non-Sámi respondents, we considered social and cultural explanations for perceived inequalities. Indeed, we found that discrimination experience was significantly predicted by proficiency in a Sámi language, its use within and beyond the family context (Figure 2).

Sámi who do not frequently use their language in public experience less discrimination. This can mean several things. First, using a minority language in public or having an accent in the majority language (due to frequent use of the minority language) may reveal one’s ethnicity. Second, those with stronger ties to Sámi culture could be more perceptive of ethnic discrimination. Third, the lack of opportunities to use and develop one’s language constitutes a form of discrimination in itself. Although Sámi has recently become more visible thanks to →

Figure 2: Experience of discrimination with respect to Sámi language use



political initiatives, many challenges remain: ensuring high-quality learning conditions at schools, the demographic situation (the Sámi are a minority in most areas), and administrative obstacles preventing local implementations of state policies. Consequently, many Sámi speakers feel that language policies deprive them of their linguistic heritage while insufficient resources and the absence of their language in the public sphere imply that their language—and by implication, their ethnicity—is not equal to that of the majority.

The case of the Sámi demonstrates that language policies are key to perceptions of inequality and that not all forms of social inequality can be tackled with socioeconomic measures alone. Establishing solely material equality—the approach typically taken by governments—does not adequately target historically rooted inequalities. Frequent use of a Sámi language is associated with experiencing more discrimination, and those who have given up their language experience less discrimination. Similarly, people with Sámi ancestry but who do not self-identify as Sámi perceive less inequality. But the price they have paid is their indigenous identity. Of course, asking minorities to give up their languages and identities

in order to avoid inequality can hardly be a reasonable strategy to achieve social equality. Our research indicates that the connection between discrimination and language is linked to a lack of esteem for Sámi culture and a lack of opportunities to develop their languages. Thus, ingraining the languages more assertively and positively not only within the minority community but in the society at large could contribute to dismantling discrimination.

Isolated language revitalization initiatives have shown remarkable results. However, they must be backed up by the state, even if this requires an investment, with comprehensive tuition in the language, high-quality teacher training, teaching material creation and schools' and municipalities' compliance with policy measures. The visibility of minority languages in the media and mainstream culture can be increased, all school students should receive education about minorities in their country. Such efforts can potentially tip the scale from language

decline to growth. Due to intergenerational gaps, the middle generation is an important target group for revitalization. They should be equipped with meta-linguistic and writing skills, and teacher education in minority languages must be incentivized to high training standards. The linguistic inequality we have observed is exemplary of many other cases in the world, where minorities have been deprived of their languages. /



Tanja Kupisch is a Professor of Romance Linguistics at the University of Konstanz, Assistant Professor at UiT and a PI of the Cluster.



Anika Lloyd-Smith is a postdoctoral researcher at the Cluster and works at the Department of Linguistics at the University of Konstanz.



Fabian Bergmann is a doctoral researcher in the Cluster project "Ethnic Policies"—Remedy for Between-Group Inequalities?



Rusen Yasar is a postdoctoral researcher in the Cluster project "Ethnic Policies"—Remedy for Between-Group Inequalities?

INSIGHTS FROM OUR RESEARCH

Targeting the Voter. *How Italian Political Parties Phrase Identity Appeals*

(S. Zanotto)

18

One way of convincing voters in public speeches is the appeal to identity. Analysis of public discourses of Italian politicians from 1950 to 2020 shows the subtle ways of using words that appeal to the voters' identities. The comparison of public statements across the political spectrum reveals the differences between the left and the right in wording concerning class, profession, and social status.



→ Ballot boxes in an Italian polling station

19

In politics, words matter. This is especially true for the words spread to a mass audience, like televised political interviews and parliamentary speeches. TV offers politicians a powerful platform for addressing the audiences' preferences and for communicating their ideology, their party, their political agenda.

How exactly do politicians attract attention and convince the audience? One key strategy under consideration is related to a phenomenon called "Identity Appeal". When using an identity appeal, a politician speaks of a particular part of the population for showing that she represents them, and therefore reassuring them that she is the best choice. For example, how many times have you heard a politician speak of the need to raise the pensions for retirees? Without speaking of the actual realization of this statement, the

politician is assuring that voting for her would improve life conditions for the majority of the elderly. If you are a retiree who struggles with rising living costs, you feel that this sentence is referring to you, or at least to a part of your identity.

To understand the way different political parties each use specific words to appeal to the audience's identities, I use automatic tools for text analysis, enabling me to analyze a great quantity of political discourses. This allows me to better understand how Italian parties speak to their (potential) voters, but it also offers data that other researchers can use for conducting comparative studies for their own countries.

In this article, I will consider only the analysis of the identities that Italian politicians call out when referring to different occupations or, to be more general, to the social situation of a person. That is why in my analysis we can see more specific references in politicians' discourses such as teachers and also more general calls to a larger part of the voters such as workers. →

20 The first step is to define and understand which are the words used for referring to people's identities. To do so, I create a list of words containing all the terms that refer to the occupational status of a person and automatically count them in the texts. Having the possibility of analyzing 1,500 spoken discourses of Italian politicians from 1950 to 2020, adding up to a total of around three million words, the first important step is to understand the absolute frequency of identities' occurrences in the dataset, therefore counting how many times specific parties appeal to an identity and which are the main targets in their speeches. Regardless of the party and the party's political orientation, the most frequent targets in the speeches are all very similar. This holds especially true for the word "workers" (in Italian lavoratori); in brief, if you are a worker, and most of the voters are, you are a politician's target.

However, just looking at the number of times your identity is addressed in speeches is not enough: how am I described by these politicians? Am I young, tired, angry, or needy? For identifying such a description, I use two different methods. The first one is

called co-occurrences, and looks at the neighboring words to "workers": adjectives, verbs and nouns that are used around the target word for describing it more specifically. This method again is a counting-based method: for example, for the most important current Italian parties, the closest word to their own idea of "workers" is always "enterprises."

The second method is called word-embedding, and looks at which words have similar surroundings to our target "workers" in the text. When we apply this method and check the probability of two words to be similar in comparison to their closest surrounding words, we consider a wider range of textual context. That is, we find all the possible words that could substitute the word "workers." Here we learn not only which are the most similar words to our target "workers," that is which kind of worker is appealed by singular parties; we also get an idea of how close the descriptions of "workers" are among parties in terms of relative distance between themselves. For example, a factory worker could see that in some parties the term "worker" is closer to "entrepreneurs" than to "factory worker." From this, she/he could conclude that this party probably represents

Figure 1: Word Embeddings of "workers" by Italian Political Parties

Word embeddings calculate high dimensional vectors over all properties of a word as found in a given text or corpus. These high dimensional vectors can be compared

as to how close they lie to one another in a vector space. It has been shown that spatial closeness corresponds to lexical semantic similarity, in other words how close or related the meanings of two words are.



the interests of entrepreneurs rather than those of factory workers. This is an example of how the use of a more general language on identity appeals permits to maintain a positive face in front of the majority of voters, while subtly conveying the interests of few of them.

Indeed, this second method shows differences among the principal parties in referring to workers: The more leftist idea of workers of the Partito Democratico (Democratic Party, PD) is far away from the more rightist one of the Lega Nord (Northern League), and of Forza Italia (Come On, Italy, FI); the perspective of the popu-

list party Movimento 5 Stelle (Five Stars Movement, M5S) on workers is a mix of the above-mentioned positions, but still more similar to the right than to the left, or at least it was until 2020.

To give you an idea of the most similar words to specific parties' idea of workers, we see in the figure that the Democratic Party talks about employees (dipendenti), self-employed (privati) and wages (salari); the more rightist Lega speaks of disabled people (disabili), shops (negozi) and dreams (sogni); Berlusconi's party Forza Italia speaks of unemployed people (disoccupati), retirement contributions (con- →

tributi) and the rich (ricchi); the Movimento 5 Stelle identifies workers with savers (risparmiatori), debts (debiti) and savings (risparmi). This analysis shows how general occupational terms like “workers” might actually refer to a smaller group like “employees” or “the rich”. Future research can potentially identify subtle discriminative language of politicians in referring to immigration and migrants and explain how refugees are described and with whom or what they are associated.

Linguistic analysis of public discourse is the first step towards fully understanding the intricate ways of political appeals, transfer of ideologies, persuasion—and what parties politically make of the perceptions and expectations of voters. That the party of the late Berlusconi appeals way more to rich people than the others is an important part of the puzzle that can be made transparent through linguistic research. A next step would be the use of insights like these by political scientists. In the end, a thorough understanding will contribute towards making subtle manipulations of citizens and voters more transparent and can serve as reliable information to form an opinion on parties, programs, and propaganda. /



↑ Voters look at election posters for a local election in Italy



Sergio Zanotto is a PhD candidate in Linguistics at the University of Konstanz and an independent doctoral researcher at the Cluster.

When a civil servant engages with a citizen, language can be an instrument of power, and the relation can be marked by inequality.

A detailed analysis of verbal interaction between civil servants and citizens shows, however, a high degree of responsive speaking behavior and a lack of evidence for biases. The civil servants adapted

the level of speech complexity to their counterparts in a flexible manner.

24

INSIGHTS FROM OUR RESEARCH

Language and Power.

Computational Analysis Can Help Us Understand Inequalities in Communication

(S. Eckhard, I. Espinoza, S. Frenzel, A. Hautli-Janisz, W. Siskou)

Language is central to reflecting and expressing social hierarchies. For instance, people in a higher power position are allowed to interrupt other speakers. They also have the freedom of choosing who speaks next and they themselves can dominate a conversation by extending somebody's speaking time with of interruptions or by evaluating others' speech acts with statements such as "that's right" or "very good." There are many ways in which interactional power manifests itself in communication and they are found everywhere, from informal communication between parents and children to formal interactions in the workplace.

With the advancement of more sophisticated techniques in Computational Linguistics, there is now an opportunity to complement qualitative and quantitative work in the social sciences with scalable automatic processes for measuring language. This also holds for identifying interactional power in larger datasets of real communication and sheds light on how power manifests itself in language and what effects it has on the interlocutors. So far there are several strands of research in computational linguistics for identifying and quantifying interactional power and its manifestations in language. Sentiment analysis, the computational study of how people convey emotions through language and how emotions shape our behavior, can be a useful tool for measuring linguistic expressions of power, such as negative evaluations of target groups, moral attitudes and the detection of microaggressions in interactions. We also find large-scale computational studies of dehumanizing language,

a phenomenon which is known to lead to extreme intergroup biases, hate speech and even physical violence. Applications of this have also been used to measure and analyze the use and reproduction of ethnic and gender stereotypes and biases. For instance, computational approaches are starting to track the construction of power hierarchies in discourse through the analysis of how low-status speakers are addressed in comparison to high-status speakers. This has yielded empirical evidence for racial profiling by police officers in the US, which shows that black community members are systematically treated with less respect.

Measuring the communication between citizens and state representatives is also at the core of the Cluster project „A Computational Linguistic Analysis of Public Service Encounters.“ Public Service Encounters (PSEs) are situations where a citizen turns to the state to claim a civil right, or to seek help or support, for instance when applying for social welfare benefits. Public Service Encounters depict a situation of structural dependence, since public officials are in a greater position of power compared to the citizens they interact with. Even though civil servants are constrained by different policies and regulations, they possess a certain amount of leeway in their decision-making, which is necessary for them to tailor their decisions to the needs of the citizens they interact with. This leeway puts them in a position of power, since they have a certain amount of freedom in choosing how to allocate resources among citizens trying to access public services, such as public housing or other welfare services. →

25

It is established knowledge in social psychology that public officials, like all humans, are inevitably subject to unconscious biases which affect how they interact with citizen clients in these encounters. These biases should condition systematic discrimination and unequal public service delivery, plausibly along demographic lines such as gender, ethnicity, or age. Direct observations of citizen treatment in public service delivery, however, hardly exist. Instead, researchers usually draw such conclusions indirectly based on macro level outcome data, or by means of survey research and experiments. In our project, we use recordings of conversations between civil servants and citizens that are transcribed and anonymized, and we measure with automatic means how public

officials engage with the citizens they interact with. Of central importance in these settings is the comprehensibility of the conversation: we assume that hierarchy is in parts manifested by the complexity of the language, backed up by ongoing efforts around the globe to put in place regulation to reduce this aspect of inequality in communication. Since 2002, for instance, the German Act on the Equality of People with Disabilities (Gesetz zur Gleichstellung von Menschen mit Behinderungen), as well as the Barrier-Free Information Technology Ordinance (Barriere-

freie-Informationstechnik-Verordnung) regulate the accessibility of written texts exchanged between people with disabilities and the federal government.

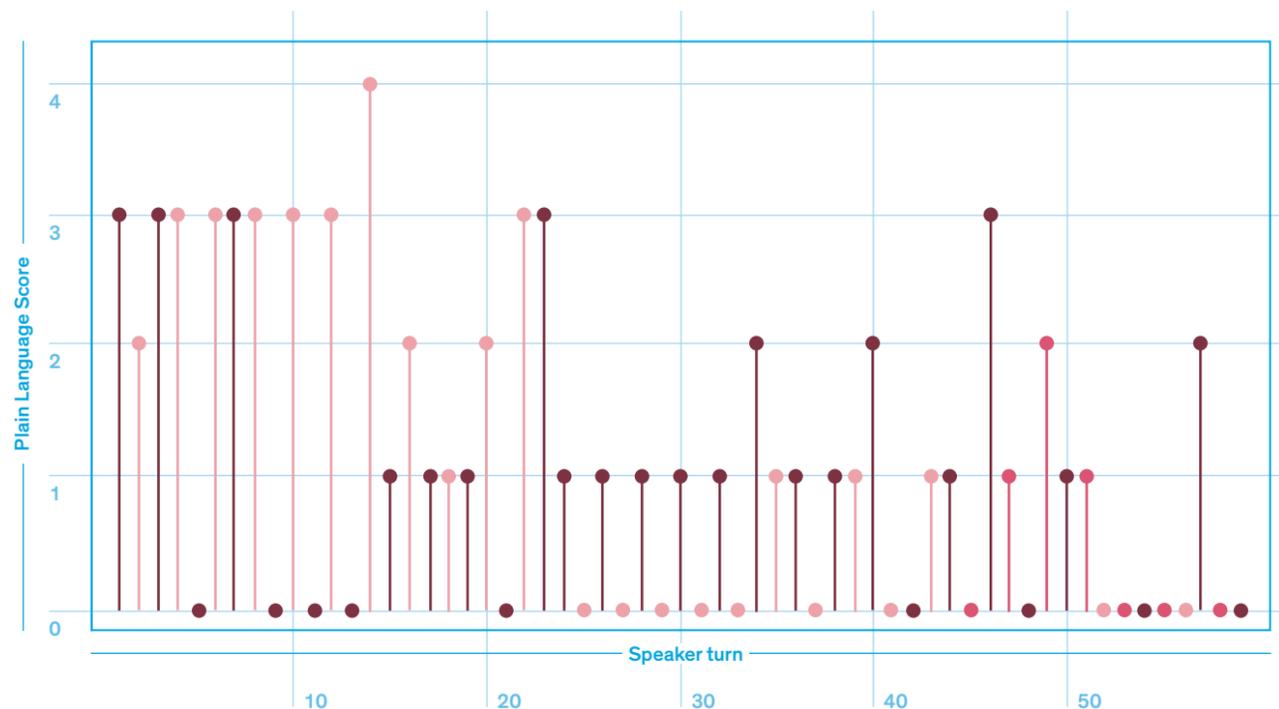
Starting from the criteria available for written communication (Netzwerk Leichte Sprache, founded in 2006), we systematically investigate administrative spoken language during public service delivery and propose a so-called plain language score that allows us to measure the comprehensibility of speaker turns. This language score is based on a multitude of dialogical features, from word-level features like jargon and specialized vocabulary, to sentential structures like the passive and participle construction, which are known factors for increasing language complexity. The details of the approach were published in the proceedings of the annual conference of the German Society of Computational Linguistics and Language Technology 2022.

As part of the EXC project we conducted an analysis of 52 PSEs recorded in 2021 and 2022. The investigation shows that on average, the language of public service employees and their clients is relatively similar. An investi-

gation of the dynamics of individual PSEs, however, yields a more nuanced picture, namely that there are changes in complexity during an encounter: In cases where the clients use less complex language, we do see public service employees reacting by reducing their language complexity. At the same time, more complex language by the clients yields sequences of more complex utterances by public officials. The lollipop chart in the Figure below illustrates the conversation dynamics of one such public service encounter. In a nutshell, we find that the civil servants leading these conversations we analyzed, had a responsive speaking behavior in the sense that they adapted to the speech complexity of the citizens they were interacting with. We do not find evidence for biases leading to strong hierarchies affecting these conversations in an unequal manner. →



↑ Bureaucracy in Germany



By incorporating the results of a survey on client satisfaction that was elicited in addition to these transcripts, we have the unique opportunity to understand the way in which civil servants engage with citizens. This next step in our research will also help us understand where and how unequal treatment originates in communication in public service encounters—and to what extent and how power is manifested in this type of communication. /

Figure 1: Lollipop Chart
 Visualization of conversation dynamics based on the plain language scores of employees and clients within a public service encounter

- Employee
- Client 1
- Client 2



Steffen Frenzel studies the master's programme "Speech and Language Processing" at the University of Konstanz and works as student assistant in the Cluster project "Inequality in Street-Level Bureaucracy."



Ingrid Espinoza is a doctoral researcher in the Cluster project "Inequality in Street-Level Bureaucracy: Linguistic Analysis of Public Service Encounter."



Wassiliki Siskou is a doctoral researcher in the Cluster project "Mobilizing Inequalities: From Grievances to Conflict."



Annette Hautli-Janisz is a Junior Professor of Computational Rhetoric and Natural Language Processing at the University of Passau and associated researcher at the Cluster.



Steffen Eckhard holds the Chair of Public Administration and Public Policy at Zeppelin University Friedrichshafen and is a PI of the Cluster.

TALES FROM THE LAB

All You Need Is Love? *Politicians' Charisma* 30 *and the Effect* *of Emotional Messages* (J. Vári) 31

Politicians who use inclusive language and emotionally pleasant words tend to be regarded as more charismatic than others whose language contains more exclusive and less emotional words. This emotion effect depends greatly on whether we consider written or spoken speeches. This is the result of a linguistic analysis of speeches by British politicians combined with a perception experiment.

↓ In future, we are working with Avatars, computer animated graphical representations

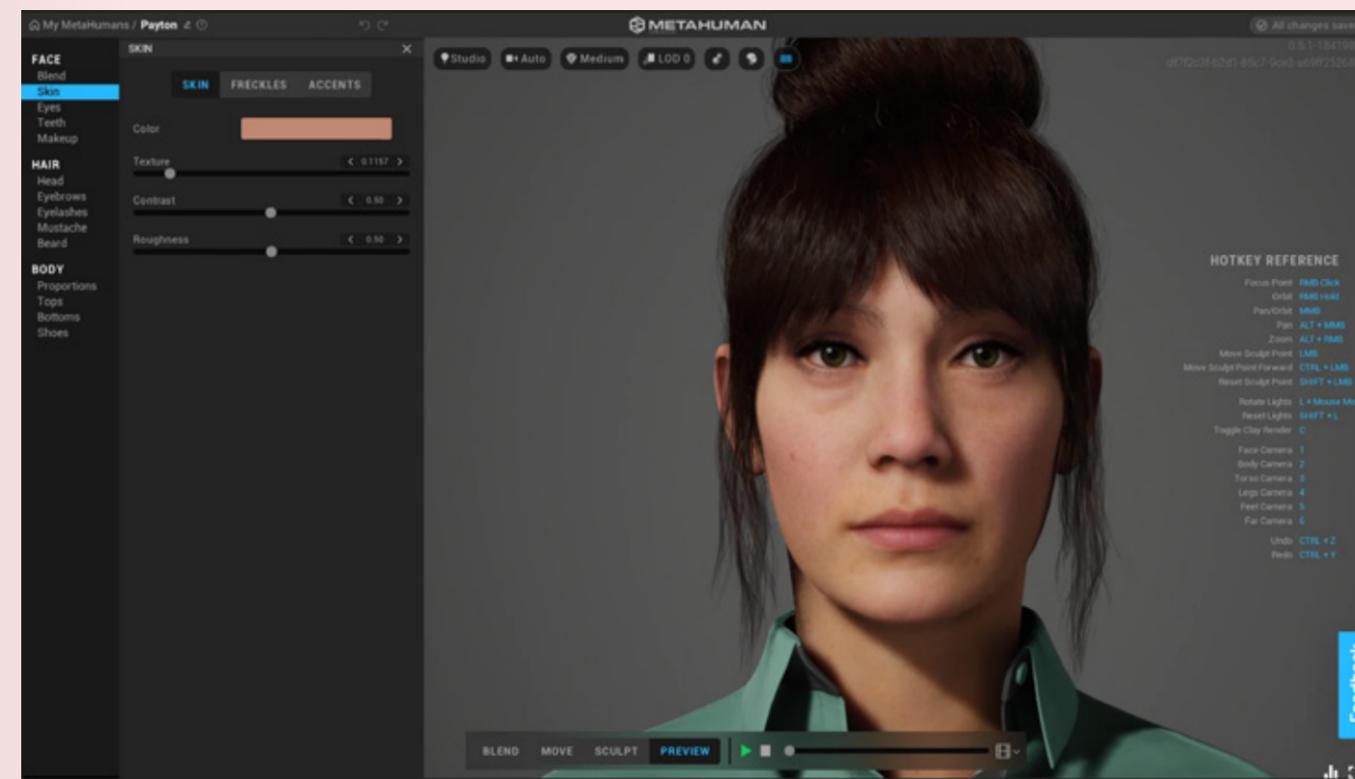
People tend to agree: Obama has it, Martin Luther King had it as well as Helmut Schmidt—it being a special power called “charisma.” The concept of charisma was introduced to the social sciences by Max Weber in the 20th century. For Weber, charisma is a personal quality of leaders who can inspire and influence followers in times of crisis. His, and other definitions of charisma, often include an important role of emotions. A study conducted by Joyce E. Bono and Remus Ilies found in 2006 empirical evidence that the display of positive emotions improved speakers’ charisma ratings.

Our research team wanted to build on this study’s findings, but also explore other aspects of emotional as well as charismatic messages. For example, besides emotional positivity—which Bono and Ilies looked into—does also emotional negativity impact on charisma? What particular role does language play for charisma? Most importantly, we were particularly interested whether the positive effect of emotions on charisma holds up also in written texts and not only in speech.

In order to investigate these questions, we took speeches from four British politicians on the topic of Brexit: David Cameron

(Conservative Party), Nigel Farage (UK Independence Party and later Brexit Party /Reform UK), Theresa May (Conservative Party) and Nicola Sturgeon (Scottish National Party). We focused only on the audio channel of the speeches, and not the visual, and extracted smaller samples of the whole speech which consisted of around 13 words on average. We then transcribed these excerpts in Standard British orthography. In the end, we had two stimuli in two different modalities: audio excerpts as well as written excerpts. Using these materials, we wanted to investigate whether politicians’ emotional language impacts on charisma. In order to do so we needed to find measures to establish how emotional our Brexit extracts were.

But what are even emotions? And how can one establish objectively that one message is more emotional than another one? →



Social psychological theories define emotions to have three dimensions: emotional valence (how pleasant), emotional dominance (how potent), and emotional arousal (how engaging). We wanted to look at the emotional dimensions of words in messages and see whether there was a link to charisma. We specifically focused on content words. These are words which have a meaning by themselves (e.g., apple) as opposed to function words which have a structural grammatical meaning (“the”). For example, the content word “love” is emotionally very valent (very pleasant), but the word’s emotional arousal and emotional dominance is medium.

We establish these emotion ratings by looking at earlier research, for example by Margaret Bradley and Peter Lang. They asked participants to rate around 1,000 English words on emotional valence, arousal and dominance. Based on these ratings, we were able to measure the emotionality of each sample of the Brexit speeches. We simply added the emotion ratings on all three dimensions, i.e. valence, arousal, and dominance for each content word in each extract. This way, we received for each extract three emotion ratings: one for its emotional valence, one for its arousal, and one for its dominance. We also standardized these emotion ratings for the number of words in an extract, so we would not end up having higher emotion ratings just because some extracts were longer.

What about function words? Inclusive pronouns (such as “we”, “us”) were good candidates to evoke emotions. Previous research had shown that inclusive pronouns can create a shared, collective identity. This collective identity creation could potentially entail positive feelings. But were inclusive pronouns also linked to charisma via emotions? In order to explore this question, we needed to measure somehow the pronouns in the extracts from the Brexit speeches. We calculated for each extract a ratio between occurrences of inclusive vs. occurrences of exclusive pronouns. We also standardized this pronoun ratio for the length of the extract, i.e. the number of words.

All that’s missing now is the concept we were mainly interested in: charisma. We recruited 149 participants who either read or listened to a smaller selection of the excerpts of the Brexit speeches and rated the messages by using an online questionnaire and a scale from “not at all” to “very” charismatic. Participants were asked how charismatic they perceived politicians to be based on the audio and written extract, without explicitly giving away politicians’ identity.

32

Then, the magical moment came: We related all these measures, i.e. emotional valence, arousal and dominance of content words, inclusive vs. exclusive pronoun ratings, to charisma ratings and explored this relationship. We were indeed intrigued by the results of our statistical analysis: Our results provided further evidence for Bono and Ilies findings as well as many theoretical works: emotionally positive messages lead to more charisma. However, this was not the case for the emotional dimensions of dominance and arousal and inclusive pronouns, but only for the emotional valence (pleasantness) dimension. We also found that emotionally negative words lead to lower charisma ratings and conversely, the more emotionally positive words (e.g., “love”) the politicians used, the more charismatic they were perceived to be. Is this the end of the story? All you need is love—to be charismatic?

Remember, we were particularly interested in the modality of the messages: Do emotional messages in writing affect politicians’ perceived charisma differently compared to au-

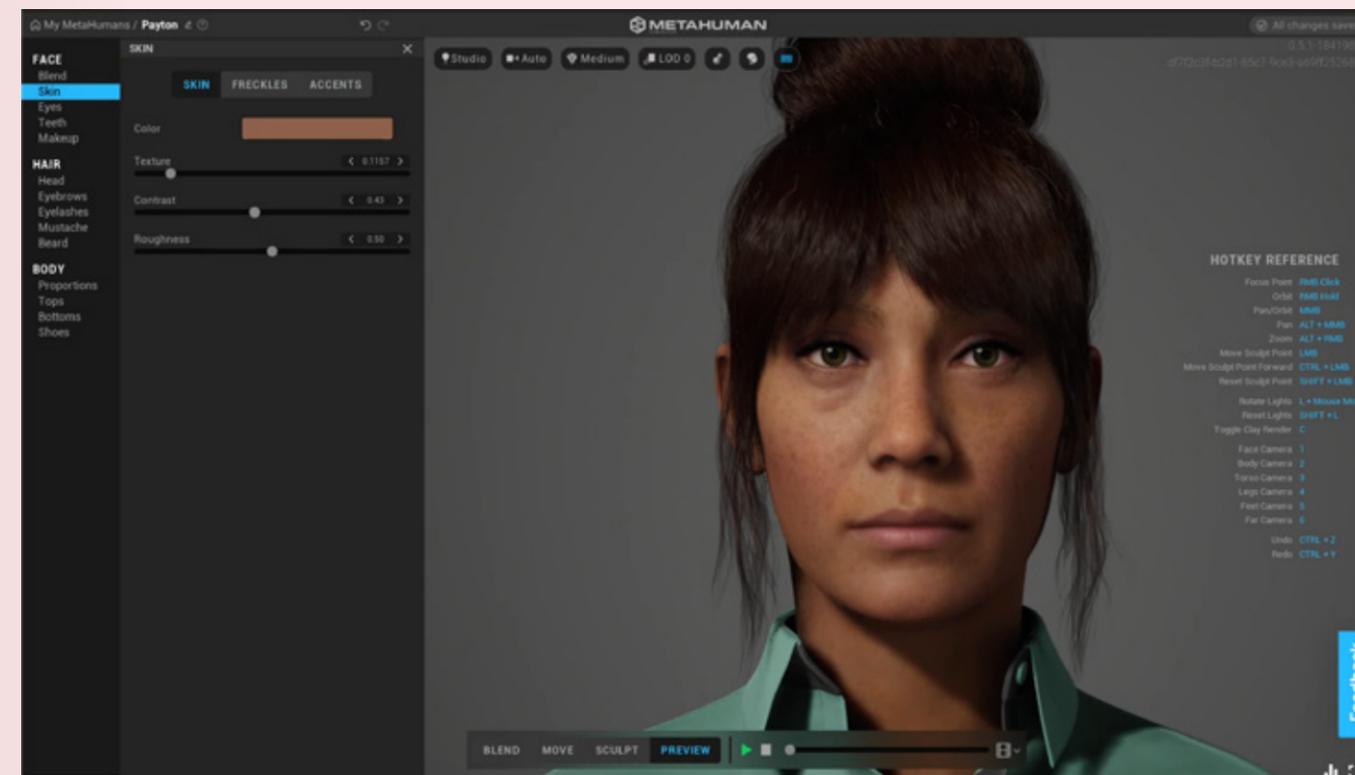
dio messages? The answer is “Yes”: the modality of messages does play a role for the positive effect of emotions on charisma. The effect of emotionally valent words (pleasant words) as well as inclusive pronouns (“we”, “us”) was significantly weaker or not even present in audio messages. Emotionality strongly impacted on charisma in written messages. We interpret this result by suggesting that perhaps other features of the speech “overwrite” the effect of emotional content in audio messages. These other features could be related to speakers’ voice such as their individual pitch. The possibility to recognize the identity of politicians in the audio message might also play a role.

Factors like recognizability of speakers point to the shortcomings of our study. The

major limitation is that participants potentially could have recognized politicians in audio messages which led them to assign specific charisma ratings based on who politicians were and less based on the words they used. This problem of recognition is present in all audio(-visual) stimuli which are taken from natural settings like our Brexit speeches. To address this problem, researchers have used recordings of actors performing scripted speeches and presented these stimuli to participants to be rated. There is a downside: the stimuli become less naturalistic. And even worse, we still have too many unknown factors based on actors’ individualistic voices and individualistic appearance. How can we know whether individual difference in charisma is based on their different voices or the different color of their eyes, for example? We could only make such inferences if we found the actor’s identical twin or even their clone who look exactly the same but have for example blue eyes instead of green. Or a higher or lower voice but otherwise completely similar looks. →

33

↓ An Avatar with identical facial features but different skin color





Judit Vári is a postdoctoral research fellow in the Cluster investigating the "Perception of Political Charisma in Low-Status Speakers."

The color of our eyes is not likely to impact on charisma but other aspects of appearance such as skin color surely are. Skin color could be one indicator—among other factors—of speakers' marginalized backgrounds on which we focus in our project "Perceptions of Political Charisma in Low-Status Speakers." Our project at the Cluster investigates under what circumstances politicians from marginalized backgrounds are perceived to be charismatic in contrast to politicians of high status. We would like to especially know what role the perceivers' background plays by asking who finds whom charismatic? We try to tackle these questions by using innovative methods that improve on the slightly unrealistic notion of using clones or triplets for research: Avatars, computer animated graphical representations. With their help when can aim to isolate specific

factors which might impact on charisma. Our Avatars have the advantage that we can manipulate just one specific factor: Avatars can have the same facial features but for example have a different skin color. Surely, using Avatars will lead to less naturalistic stimuli compared to our study on Brexit speeches. But it's a trade-off: less naturalistic stimuli and more control over stimuli. Therefore, Avatars will allow us to systematically investigate whether there are biases towards politicians from marginalized backgrounds in the perception of charisma—without twins and clones. Currently we are creating the Avatars and their voices and preparing to run our first perception experiments. So, stay tuned for the first results. /

To be continued: Watch the videos on how we change the look of our Avatars with the "MetaHuman Creator" →



Predicted Values of Charisma

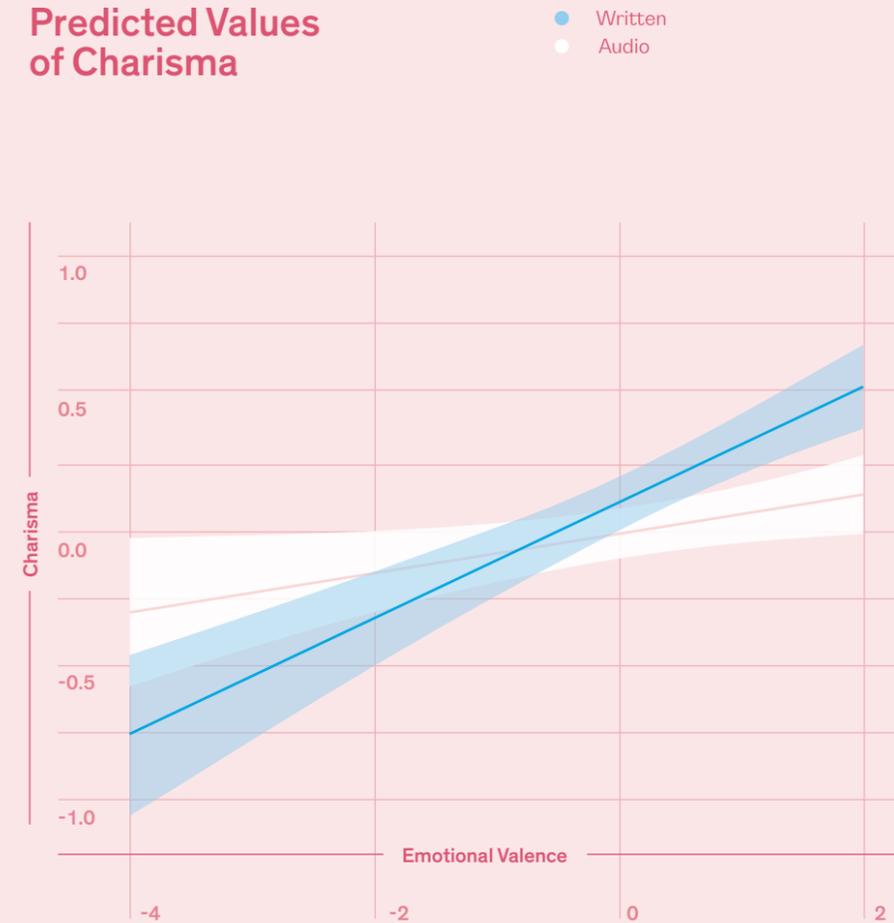


Figure 1: The more positive the emotional valence of words, the higher the charisma ratings for the text modality (blue line). Not the case for the audio modality (white/ pink line).

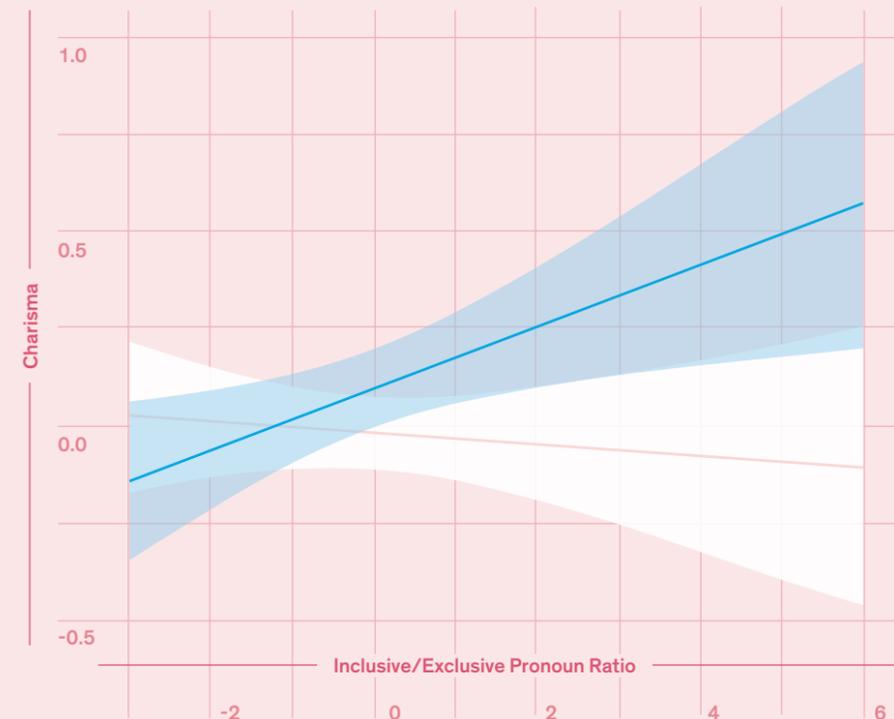


Figure 2: The more inclusive pronouns, the higher the charisma ratings for the text modality (blue line). Not the case for the audio modality (white/ pink line).

HONORS

Honors, memberships, appointments, and awards for Cluster members (Selection, October 2022—May 2023)

36

Prof. Dr. Marius R. Busemeyer was elected to the supervisory board of the Heinrich Böll Foundation and appointed to the scientific council of the Berlin think tank “Das Progressive Zentrum.” Furthermore, he was elected as President of the User Advisory Board of the Gesis Leibniz Institute of the Social Science, Cologne / Mannheim.

Prof. Dr. Claudia Diehl was elected into the “Heidelberger Akademie der Wissenschaften HAdW.”

Prof. Dr. Urs Fischbacher received the Allais Memorial Prize in Behavioral Sciences awarded by the Prague Conference on Behavioral Sciences in cooperation with Fondation Maurice Allais.

Niklas Hänze received the Christiane-Rajewsky-Award by the German Association for Peace and Conflict Studies in March 2023 for this master thesis titled “Reversing the Causal Arrow: The Nexus Between Armed Conflict and Natural Disasters.”

Dr. rer. pol. Roman Krtsch received the dissertation award by the “Förderkreis der Osnabrücker Friedensgespräche” in March 2023 for his dissertation “Spatial and Temporal Dynamics of Civil Resistance in War Contexts.”

Prof. Dr. Gabriele Spilker has been appointed member of the Research Board of the Einstein Foundation in March 2023.

Prof. Dr. Nils B. Weidmann stayed in November 2022 as Visiting Fellow at the Center for Advanced Studies at Ludwig-Maximilians-University Munich.

Niola Asri, Katharina Michaelowa, Sitakanta Panda, Sourabh B. Paul (2022)

The Pursuit of Simplicity: Can Simplifying Eligibility Criteria Improve Social Pension Targeting? *Journal of Economic Behavior & Organization* 200: 820–846. <https://doi.org/10.1016/j.jebo.2022.06.003>

Sharon Baute, Anniek de Ruijter (2022)

EU Health Solidarity in Times of Crisis: Explaining Public Preferences Towards EU Risk Pooling for Medicines. *Journal of European Public Policy* 29(8): 1183–1205. <https://doi.org/10.1080/13501763.2021.1936129>

Luna Bellani, Anselm Hager, Stephan E. Maurer (2022)

The Long Shadow of Slavery: The Persistence of Slave Owners in Southern Law-Making. *The Journal of Economic History* 82, Nr. 1: 250–283. <https://doi.org/10.1017/S0022050721000590>

Ariane Bertogg, Nevena Kulic, Susanne Strauss (2022)

Protected through Part-time Employment? Labor Market Status, Domestic Responsibilities, and the Life Satisfaction of German Women during the COVID-19 Pandemic. *Social Politics: International Studies in Gender, State & Society* 29, Nr. 4 (2022): 1236–1260. <https://doi.org/10.1093/sp/jxab048>

Ole Brüggemann, Thomas Hinz (2023)

Do Women Evaluate their Lower Earnings Still to Be Fair? Findings on the Contented Female Worker Paradox Examining the Role of Occupational Contexts in 27 European Countries. *European Sociological Review*, 3. Januar 2023, jcac073. <https://doi.org/10.1093/esr/jcac073>



↑ **Marius R. Busemeyer** (2023) Who Cares for the Future? Exploring Public Attitudes towards the Needs of Future Generations in Germany. *Journal of European Public Policy*. <https://doi.org/10.1080/13501763.2023.2165697>

Steffen Eckhard, Laurin Friedrich (2022) Linguistic Features of Public Service Encounters: How Spoken Administrative Language Affects Citizen Satisfaction. *Journal of Public Administration Research and Theory*. <https://doi.org/10.1093/jopart/muac052>

Gianna M. Eick, Marius R. Busemeyer (2023) Migration Levels and Welfare Support: Evidence from the Local Level. *Journal of European Public Policy*. <https://doi.org/10.1080/13501763.2023.2195440>

PUBLISHED

Selected publications by Cluster researchers (Published October 2022—May 2023)

→ **Urs Fischbacher, David Gramling, Jan Hausfeld, Vojtěch Zírka** (2023) Identity Breeds Inequality: Evidence From a Laboratory Experiment on Redistribution. *Journal of Public Economics* 222. <https://doi.org/10.1016/j.jpubeco.2023.104866>

Max Heermann, Sebastian Kfoos, Dirk Leuffen (2022) Who Deserves European Solidarity? How Recipient Characteristics Shaped Public Support for International Medical and Financial Aid during COVID-19. *British Journal of Political Science*, 1–22. <https://doi.org/10.1017/S0007123422000357>

Anika Lloyd-Smith, Fabian Bergmann, Laura Hund, Tanja Kupisch (2023) Can Policies Improve Language Vitality? The Sámi Languages in Sweden and Norway. *Frontiers in Psychology* 14. <https://doi.org/10.3389/fpsyg.2023.1059696>

Adrian Rinscheid, Sebastian Kfoos (2023) War and Pandemic Do Not Jeopardize Germans’ Willingness to Support Climate Measures. *Communications Earth & Environment* 4, Nr. 1. <https://doi.org/10.1038/s43247-023-00755-z>

→ **Nils B. Weidmann** (2023) Data Management for Social Scientists: From Files to Databases. Cambridge University Press. <https://doi.org/10.1017/9781108990424>



Felix Wolter, Jochen Mayerl, Henrik H. Andersen, Theresa Nieland, Justus Junkermann (2022) Overestimation of COVID-19 Vaccination Coverage in Population Surveys Due to Social Desirability Bias: Results of an Experimental Methods Study in Germany. *Socius: Sociological Research for a Dynamic World* 8: 1–8. <https://doi.org/10.1177/23780231221094749>



37

The Cluster of Excellence “The Politics of Inequality” contributes to teaching in three different master’s degree programs at the University of Konstanz. Here’s what students are saying:

Zoé Wolter, 4th semester



“What excites me about Social and Economic Data Science is its ability to uncover hidden patterns and insights from vast amounts of data. The program provides the perfect opportunity to practically apply theoretical statistical concepts and coding skills in a social science setting. My biggest 'eureka moment' so far was realizing that data science doesn't just mean fancy programming skills and creating crazy models, but that even a simple spreadsheet and simple graphs are already data science.”

Master of Science
“Social and Economic Data Science”

With its innovative interdisciplinary combination of political science and public administration, computer science, statistics, and economics, this degree program prepares students for the diverse challenges of our digitalized society. The focus is on the basic questions of measurement, representation, and causal relationships that are often overlooked in current big data debates. This enables students to analyze the wealth of data on human behavior and social interactions and use it to inform decision-making in social practice.



Conrad Winch, 4th semester



Master of Arts
“Politics and Public Administration”

This master’s degree is flexible, interdisciplinary, and international. The thematic focus on “Inequality” can be combined with specializations in international politics, political economy, conflict studies, and management research, depending on students’ interests. The program features a variety of international opportunities, including the availability of double degrees, co-curricular internships, and study abroad options.



“I chose the specialization ‘Political Economy and Inequality’ because I have had a long-lasting interest in economic effects of politics and government policy. This may involve questions such as how institutional norms might affect a foreign investor’s perception of a country as a viable market to invest in, whether the ideological traits of a government may affect the effectiveness of its policy implementation, or how political actors shape the regulation of financial markets.”

Jonas Stark, 2nd semester



“Inequality lies at the heart of sociological interest. In my last term paper for example I focused on public opinion about housing market regulation in Germany—a heavily debated topic. It was fascinating to see who actually demands (de)regulation and what drives these opinions. The master program enables me to put the great theories to the test and assess their value in numbers.”

Master of Arts
“Sociology of Inequality”

This master’s program offers a research-oriented curriculum with an emphasis on contemporary issues and problems. The focus is on topics such as migration and social integration, gender inequalities, education and the labor market, organizations and economics, culture, and social networks. Students receive research-oriented training in theory building, analytical methods, and classification and are systematically involved in ongoing sociological research.



— Cluster members

Marco Bitschnau
Postdoctoral Researcher

Christina Felfe de Ormeño
Principal Investigator

Alexandra Frasch
Managing Office,
Early Career Coordination

Max Heermann
Postdoctoral Researcher

40

Celina Kacperski
Postdoctoral
Researcher

Annalena Kampermann
Managing Office, Public Outreach
and Science Communication

Maren Lüdecke
Doctoral Researcher

Rebecca Strauch
Postdoctoral Research Fellow

Roberto Ulloa
Postdoctoral Researcher

Frank Wehinger
Managing Office, Research and
Data Management, Methods
and Data Hub, Reporting

Franziska Windisch
Doctoral Researcher

— Guests and Fellows 2023

Prateek Bhan
University of Würzburg
Guest Researcher

Alexandra Bögner
University of Salzburg
Doctoral Fellow

Dietmar Fehr
University of Heidelberg
External Senior Fellow

Elisa Funk
University of Salzburg
Doctoral Fellow

Heiko Giebler
SCRIPTS, *Freie Universität Berlin*
Guest Researcher

Swen Hutter
Freie Universität Berlin
Director of the Center for
Civil Society Research,
Guest Researcher

Dafni Kalatzi Pantera
University of Essex
Guest Researcher

Lena Maria Schaffer
University of Luzern
Guest Researcher

Murat Yildirim
University of Stavanger
Guest Researcher

ARRIVALS

Sebastian Findeisen
“Wie die KI unsere Arbeitswelt verändern wird” (MDR, 16 May 2023)

Marius Busemeyer
“Reiche Menschen unterschätzen, wie gut es ihnen geht” (ZEIT, 11 May 2023)

Daniel Thym
“Deutschland darf bestimmen, wer kommen darf” (FAZ 06 May 2023)

Florian Kunze
“SWR Aktuell—Wirtschaft: Homeoffice entscheidend bei der Jobwahl” (SWR, 20 April 2023)

Guido Schwerdt
“Ramadan: Bessere Schulnoten im Fastenmonat” (Campus & Karriere, DLF, 22 March 2023)

Rüdiger Wilhelmi
“Banken am Bodensee sagen: Geld ist sicher” (SWR, 17 March 2023)

Marius Busemeyer, Adrian Rinscheid
“Viel Geld für viel Autofahren” (ZEIT, 23 February 2023)

Thomas Hinz:
“Krieg in der Ukraine: Wie schauen Studierende in Konstanz und Kiew drauf?” (SWR, 31 January 2023)

Sebastian Koos
“Kleben für das Klima—Wie weit darf Protest gehen?” (DLF, 16 December 2022)

Gerald Schneider
“Auf Nukleardrohungen mit Defaitismus reagieren? Wider den Vulgär-Realismus in der Ukraine-Debatte” (NZZ, 21 October 2022)

*Media coverage of the Cluster & its members
(Selection, October 2022—May 2023)*

HEADLINES

41

News

10–12/
04/2024

In_equality Conference 2024
Bodenseeforum Konstanz and online

<https://inequality-conference.de/>



DATES



INSIGHTS FROM OUR RESEARCH

Subtle Messages. *Artificial Intelligence Enhances Research on Framing, but Is still Far 42 from Perfect*

(Q. Yu, A. Fliethmann)

Artificial intelligence (AI) is a helpful tool in research: It can be used to process large amounts of data and to incorporate new information and distinctions on an ongoing basis. Cluster researchers, for example, analyze how the recent refugee crisis is framed in the German media. *The findings also show that AI at this point is not a perfect tool for capturing all the linguistic means that people use to frame an issue or an opinion.*



Anselm Fliethmann is a doctoral researcher in the Cluster project "Framing Inequalities."



Qi Yu is a doctoral researcher in the Cluster project "Framing Inequalities."

In August 2022, German Finance Minister Christian Lindner (FDP) said the term "Dienstwagenprivileg" (official vehicle privilege) was a case of "leftist framing," conveying the feeling that the wealthy receive unfair, privileged treatment from the government. According to Lindner, a linguistic device is used to arouse emotions from an ideologized position. This kind of terminology, or "framing," he said, influences opinions and attitudes.

Frames in communication are "verbal or nonverbal statements that place a clear emphasis on particular considerations," in the definition by political scientist James Druckman. This means that one aspect of an issue is emphasized through a substantive identifier or linguistic device. This can be done with an intention to manipulate, as a subtle way to arouse associations or to emotionalize, for example. Frequently, however, framing also occurs unconsciously and unintentionally.

Framing is of interest to multiple academic disciplines: Linguists, for example, explore its occurrence in text corpora and the semantic interpretation of frames. Psychologists mainly try to un-

derstand the cognitive processes that lead to frames, read or heard, changing people's perception.

As linguists and political scientists in the Cluster project "Framing Inequalities," we study how refugees in Germany are framed in public discourse. We analyze the framing strategies used in print media coverage of the refugee crisis and the extent to which we can automatically recognize and analyze framing strategies with the help of artificial intelligence (AI).

Frames in a representative body of text can best be studied with the help of electronic tools. In late 2022, the launch of ChatGPT made the computer-assisted analysis of large text corpora a major public issue, both inside and outside of academia. Developing AI programs that can process human language, like ChatGPT, is part of "Natural Language Processing" (NLP), a subdiscipline of AI. Given that NLP researchers are also part of our project, we wondered: can we create a tool, a "Framing-GPT," that helps us automate the analysis of framing strategies? Is modern artificial intelligence smart enough to do this? →

44

Our answer is both “yes” and “no”. Let’s start with the “yes”: most powerful NLP algorithms are based on so-called supervised learning: the algorithm must be provided with a large training dataset, where each entity of text must be assigned to a corresponding annotation or desired output format. The algorithm is then optimized to infer the desired output value. In other words, the learning process is guided or monitored by the annotations previously assigned. This is what happened with ChatGPT, whose underlying structure was likewise built with supervised learning.

But there is also a “no.” Annotated data is very rare—an annotation is time-consuming and labor-intensive, requiring a high level of experience, knowledge, and specific training. Fortunately, not all NLP algorithms need human-curated corpora—but can still provide us with important insights about large bodies of text. In one of our studies, we use so-called word embeddings to identify the kinds of frames used in three major German newspapers: BILD Zeitung, Frankfurter Allgemeine Zeitung (FAZ),

and Süddeutsche Zeitung (SZ). Word embeddings help us measure semantic similarity, that is, how closely the meanings of individual words are related to each other. But how do you quantify “closeness” between word meanings?

Consider the three-dimensional space you are in right now. Mathematically speaking, each point in that space can be represented by unique coordinates, such as 1 meter forward, 50 cm to the left, 2 meters up (1, 0.5, 2), or by a three-dimensional vector. Word embeddings work on a similar principle: each word represents a unique point in space. In this space, semantically similar words are closer to each other. And just as we can measure the distance between two points in space—for example, between the point above (1, 0.5, 2) and our own position (e.g., 0, 0, 1.73)—we can also measure the distance between two semantic points or semantic vectors. The smaller the distance, the more similar the meaning.

This form of quantifying semantic similarity is also applied in our project. The spaces we consider, however, contain several →

↓ Stacks of newspapers at the roadside waiting for their delivery



Publishers

Prof. Dr. Marius R. Busemeyer (Speaker)
Prof. Dr. Claudia Diehl (Co-Speaker)
Prof. Dr. Gabriele Spilker (Co-Speaker)

Managing Director

Dr. Thomas Wöhler

University of Konstanz

**Cluster of Excellence
“The Politics of Inequality”**

Universitätsstraße 10
D-78464 Konstanz
T +49 7531 88-5772
cluster.inequality@uni-konstanz.de

Editors

Dr. Paul Stoop
Annalena Kampermann

Proofreading

Susanne Darabas

Translation

Carsten Bösel

Design

yella park, Aachen

Printed by

Druck & Verlag Kettler GmbH, Bönen

Paper for print version

Circleoffset Premium White (100% recycled paper, "Blue Angel" certified)

Contents may be reprinted

with permission from the editors.

Online version & subscription

https://inequality.uni.kn/magazine

The Cluster of Excellence “The Politics of Inequality. Perceptions, Participation and Policies”

is an interdisciplinary Cluster of Excellence at the University of Konstanz within the framework of the Excellence Strategy of the federal and state governments. The gap separating the poor from the rich, the worldwide rise of populism, the division of burdens in the fight against climate change, unfairly distributed access to education—many current debates are as much about inequality as they are about other issues. These topics pose highly complex questions, yet scientifically grounded answers are still few and far between. This is where we come in to investigate “The Politics of Inequality”: the political causes and consequences of inequality.

The Cluster of Excellence is grateful to the University of Konstanz and the German Research Foundation for their funding and support.

Funded by:



inequality.uni.kn



@EXCInequality

Credits

- p. 6: Philipp Uricher
- p. 10: Ines Janas
- p. 11: Ines Janas (top), Inka Reiter (bottom)
- p. 13: Unsplash / Thom Reijnders
- p. 17: Ines Janas
- p. 19: IMAGO / Milestone Media
- p. 22: IMAGO / NurPhoto
- p. 23: Ines Janas
- p. 26: IMAGO / wolterfoto
- p. 29: Ines Janas, ZU/Nicolas Bühringer (bottom right/Steffen Eckhard)
- p. 31/33: Unreal® Engine, Copyright 1998–2022, Epic Games, Inc. All rights reserved
- p. 34: Ines Janas
- p. 38: private
- p. 39: private
- p. 41: Ines Janas
- p. 43: Ines Janas
- p. 45: IMAGO / Michael Gstettenbauer
- p. 49: IMAGO / ZUMA Wire
- p. 51: private
- p. 52: María F. Sánchez (Legewie), private (Piketty), Kellogg Institute (Huber), private (Polinsky), Center for European Studies (Stephens), Ines Janas (Diehl)

i

For our research project, we created a text corpus featuring, among other data, newspaper articles from the years 2014–2018. It includes all articles that contain at least one of our query words (Flüchtling*, Geflüchtet*, I*mmigrant*, Asylbewerb*, etc.) and were published in print or online in BILD, FAZ or Süddeutsche Zeitung during this period.

hundred dimensions. While this is impossible for humans to imagine, it comes much closer to the multiple dimensions of word meanings than the three dimensions used in our example. In our research project, we construct a vector that encompasses the meanings of all our “refugee query terms”—that is, all synonyms of the German word “Flüchtling”, such as “Geflüchtete*r”, “Asylant*in”, “Migrant*in.” Then we measure the proximity of other query terms using this vector. Finally, the top 10 closest words to the query term from a topic area (e.g., security) can show us how refugees are usually portrayed, or framed, in different newspapers:

- BILD: minors, offense, criminal, jihad, violent crime, Islamist, civil war, crime suspect, shipwreck, detain
- FAZ: minors, illegal, civil war, coast guard, crime, crime rate, shipwreck, human traffickers, prison, prison sentence
- SZ: rescue mission, minors, war, civil war, illegal, minor, human traffickers, crime, protection status, shipwreck

It turns out that the refugee query term vector in BILD is semantically closer to crime-related words (see criminal, crime suspect). In the SZ, by contrast, words suggesting a humanitarian context are closer to the refugee query term vector (e.g., rescue mission, protection status). FAZ is somewhere in between.

Using word embeddings to determine semantic similarity between words is not particularly groundbreaking. In fact, among NLP researchers, this method is already considered old-fashioned. And yet, it yields quick results without prior cumbersome annotation. It also illustrates how AI can help us analyze thematic framing strategies of different newspapers.

Aside from semantic relationships, there are more subtle forms of framing, such

as the so-called grammatical framing. For example, in the headline “Erneut Dutzende Flüchtlinge ertrunken” (“Again, dozens of refugees drowned,” BILD, February 2, 2016), the word “erneut” can convey a sense of urgency and suggest the need for countermeasures—unlike the matter-of-fact reporting in the SZ of the same day: “Mindestens 35 Flüchtlinge vor türkischer Küste ertrunken” (“At least 35 refugees drowned near the Turkish coast”).

The way in which “native” language is valued is highly relevant for debates about national identity and for policies enforcing a sense of community. Distinctions between speakers of the “main” or “authentic” language and others can contribute to inequality in status.

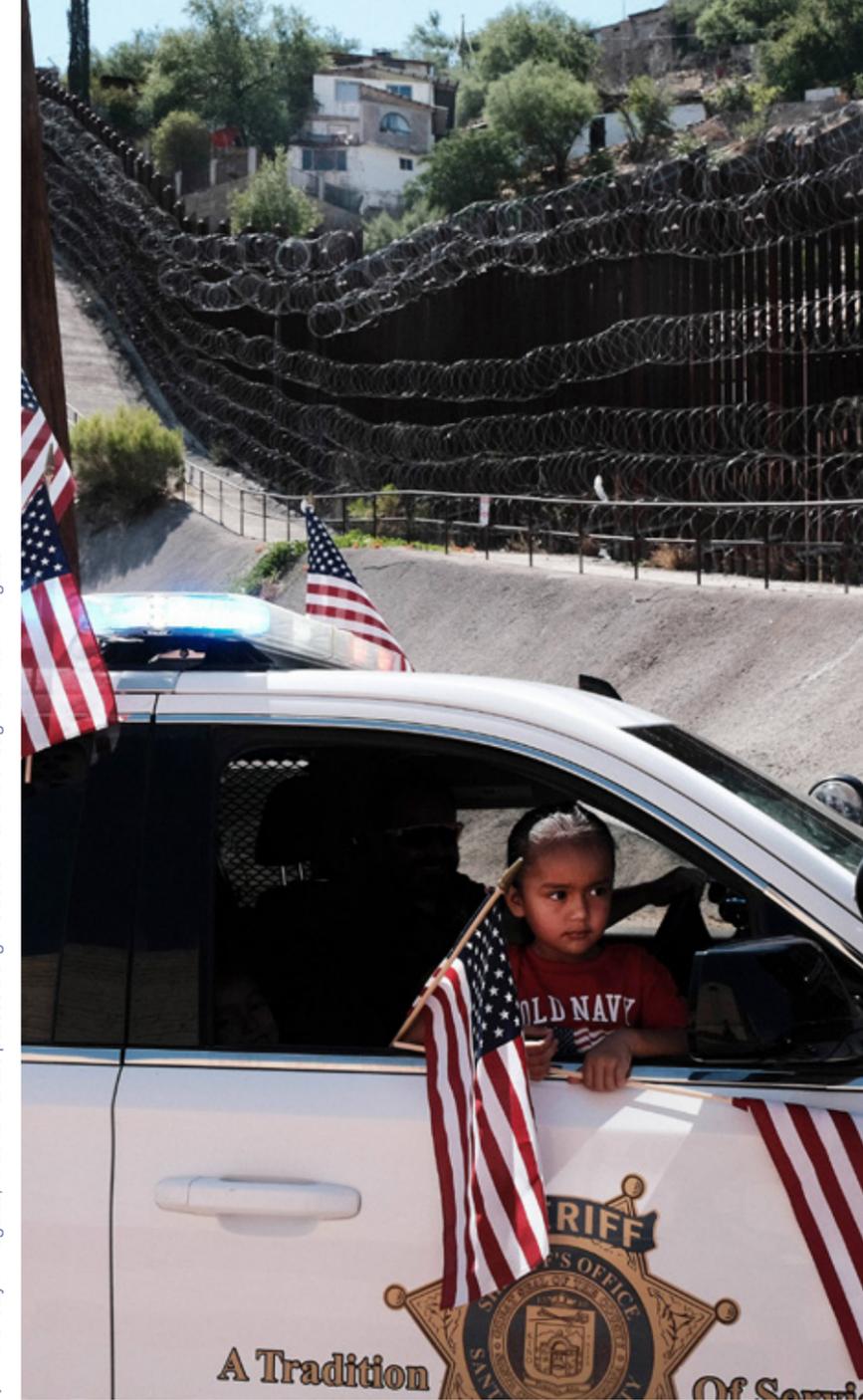
WIDE ANGLE

48 Nations, Languages, and Speakers. *Different Ways of Understanding Native Tongue and Its Role for Communities* (J. Errington)

Over the past three centuries, philosophers like Johann Gottfried Herder, anthropologists like Ernest Gellner, and political scientists like Benedict Anderson have discussed two broad ways in which languages enable membership in national societies. On one hand, competence in a native language—acquired before the onset of puberty and the loss of neuroplasticity—can be construed as a mark

of broad intragenerational and intergenerational sharedness. On the other hand, knowledge of language conventions, uninflected for background or region, is necessary for participation in modern political economies. In this article I consider the different, sometimes conflicting ways in which these two understandings of languages and their speakers have figured in a range of national ideologies and projects.

→ 4th of July in Nogales, Arizona. Parade passes along the border wall and through downtown Nogales.



Linguistic sharedness and sameness

Johann Gottfried Herder (1772) articulated the connection between individuals and nations with an account of languages as experiential bonds between members of communities and their natural environments. In a different vein, Benedict Anderson (1991) observed that one’s native language enables feelings of connectedness with national predecessors, from whom it is inherited, and national successors to whom it is bequeathed. Anderson argued that these senses of membership in “imagined” national communities were enabled by the spread of print technologies and market economies.

He further argued that monolingual nations of Europe and North America became models for those which came into being after World War II, in the so-called “Third World.” Projects of social integration and economic development, carried out within linguistically diverse populations, gave rise to what some called “Language problems of developing nations.” It was easy to see linguistic homogeneity or sameness as a prerequisite for modern nationalisms, and linguistic diversity as a mark and cause of “backwardness.” →

i

Substandard

In linguistics, “Substandard” is used to describe spoken language that does not conform to the standard language. This can include dialects or colloquial expressions, for example.

50

Ideologies of authenticity and anonymity

Images of national sharedness or sameness can be aligned with different ways of privileging national languages. Linguists Susan Gal and Kathryn Woolard link images of linguistic nativeness with ideologies of authenticity on one hand, and linguistic homogeneity or standardness with ideologies of anonymity on the other.

Ideologies of language authenticity derive support from the natural process of human development. This is because competences in a “first” language or “mother tongue,” acquired before puberty, will shape any others acquired later in life. This allows them to be treated as part of an inalienable condition shared among persons who are “fated” more broadly to membership in the same community.

Languages which are ideologically privileged by their anonymity, on the other hand, have undergone standardization: they become codified with technologies of literacy, serving as instruments of social institutions and accepted modes of public discourse. Standard national languages can be understood as “the same” for everyone, in all places, and so to belong to everybody because they belong to nobody.

In reality national languages are spoken in ways which vary significantly, and sociolinguistic research shows how speakers can

internalize a sense that their “native” habits of speech are substandard relative to anonymous norms in the nation at large. Linguists James and Lesley Milroy showed that native habits of speech associated with England’s traditional elites became associated also with anonymous literate varieties of the language. In this way a subset of a national population can be perceived to be native speakers of standard varieties of the language thanks to association with literate norms. In this way one variety of speech becomes the object of what the linguistic anthropologist Michael Silverstein calls an “ideology of the standard.” It can appear anonymous or neutral in comparison with “dialects” or “accented” speech, and for that reason associated with images of “best speakerhood.” Social hierarchies in this way can be naturalized by evaluations of other-than-standard manners of speaking as “native” but substandard, and so as inauthentic relative to national norms and communities.

Naturalizing social hierarchies

Two controversies illustrate the multiple, sometimes conflicting roles that authenticity and anonymity can play in language-centered claims about national identities.

In 2004, the distinguished American policy analyst Samuel Huntington wrote about Mexican migrants in the US in an article entitled “Who are we?” Unlike more obviously xenophobic writers, he acknowledged that children of these migrants in fact acquire English—the anonymous, neutral “language of opportunity”—at the same rate as had children of previous waves of migrants. He found problematic their continuing use of Spanish in a range of local and domestic contexts. This condition of “native bilingualism” counted for him as an obstacle to full integration into the American community. Only by “forgetting” Spanish, he argued, could they gain full entry into “the American” (not “the Americano”) dream, deeply rooted as it is in Anglophone political traditions. By his account bilingualism ideally involves an anonymous “second” or “foreign” language acquired only after the crucial period during which one acquires a native language and national membership.

Quite different is the controversy over the vision proclaimed by India’s prime minister, Narendra Modi, of Hindi as a national language. Two conflicting invocations of the values of

“authenticity” figure in this debate. To privilege Hindi for the nation as a whole, proponents of this new dispensation contrast it with the current role of English in India. They invoke “deep” historical context and Hindi’s authentic, “native” connections to a “purer” (i.e. pre-colonial) era. Opposition, on the other hand, can focus on local outcomes of this work, and the privilege it could produce for native speakers of Hindi. Even speech of those who gained full competence in “standard Hindi” as a “second language” would bear traces of their native languages, groups and regions. These linguistic artefacts of individual biographies offer an obvious resource for judging less-than-native speakership as a mark of marginal or subordinate status in the nation. In this way a project of language unity would extend Hindi political and demographic dominance into the intimacies of speakership.

These and similar conflicts, around the world, demonstrate how multiple, not entirely compatible images of languages and speakership can serve national projects and ideologies. They can vary and conflict, as do the social conditions in which they are invoked, and with the interests of those who deploy them.

Situations in which the category “native speaker” may now be losing some of its former naturalness support a similar conclusion. One involved controversy in the 1990s, for instance, about claims that English was native for members of a generation who had come of age in India speaking, reading, and writing that language. Objections were raised not on grounds of their biographies or fluency, but on historical (and cryptoracial) relations between the United Kingdom and one of its former settlement colonies. These social presuppositions or values, rather than issues of language competences, were made clear by Pieter Muyskens when he noted that we are all native speakers, but posed the question “what language are we native speakers of?”

Now a newer body of research suggests that older, print-based norms of language are losing relevance thanks to a “radical weakening or abandonment of the standard ideology” in at least ten Western European nations, as reviewed by Tore Kristiansen and Nikolas Coupland. Dominant images of native speakership have lost anonymity and authority among young persons who display a “reluctance to play by the rules of such standards,” as linguists Stefan Grondelaers and Roeland van Hout observe. Acceptance of formerly “substandard” varieties of speech is enabled by at least tacit acceptance of multiple images of speakership, and so modes of membership in a common nation.

Space does not allow for a comparison between this emerging diversity in the global North and plural nations of the global South. But it is worth conjecturing that among these latter may be found alternate models for linguistic nationalism in an era when diversity its an increasingly common feature of national societies. /



J. Joseph Errington is a Professor Emeritus of Anthropology at Yale University with a research focus on Linguistic Anthropology.

51



↑ [Thomas Piketty](#) Studiendirektor EHESS (School for Advanced Studies in the Social Sciences) / Professor an der Paris School of Economics, *Director of Studies at the EHESS (School for Advanced Studies in the Social Sciences) and Professor at the Paris School of Economics*



↓ [Evelyne Huber](#) (University of North Carolina, Chapel Hill) Morehead Alumni Professorin für Politikwissenschaft
Morehead Alumni Professor of Political Science



↑ [Joscha Legewie](#) (Harvard University) John L. Loeb Associate Professor für Sozialwissenschaften *John L. Loeb Associate Professor of the Social Sciences*

ANNOUNCEMENT

KEYNOTE SPEAKER



↑ [Maria Polinsky](#) (University of Maryland, College Park) Professorin für Linguistik *Professor of Linguistics*



↓ [John D. Stephens](#) (University of North Carolina, Chapel Hill) Gerhard E. Lenski, Jr., Distinguished Professor für Politikwissenschaft und Soziologie *Gerhard E. Lenski, Jr., Distinguished Professor of Political Science and Sociology*



↑ [Claudia Diehl](#) (University of Konstanz) / Co-Speaker of the Cluster of Excellence "The Politics of Inequality" Professorin für Mikrosociologie *Professor of Microsociology*