Linguistics and NLP in political communication

Asad Sayeed and Ellen Breitholtz

July 2024

Organizers

Asad Sayeed

Affiliation University of Gothenburg

Email asad.sayeed@gu.se

Web site `https://scholar.google.com/citations?user=mcQ2aj8AAAAJ`

Ellen Breitholtz

Affiliation University of Gothenburg

Email ellen.breitholtz

Web site `https://www.gu.se/en/about/find-staff/ellenbreitholtz`

Proposal information

Title Linguistics and NLP in political communication

Category Language and computation

Contents information

Abstract

The world has seen a drastic growth in the use of artificial intelligence techniques for manipulation public and political life. At the same time, research on political language use has not stood still. This course combines perspectives from linguistic theory, natural language processing, and political communication to provide students with the intellectual tools to apply methods from these fields into an integrated program of opinion research.

Students will learn about opinion research techniques and scientific practices; theories of discourse based on Aristotelian notions of topos and their relationship to the personae of the speaker; slurs, dogwhistles, disinformation in the context of speech acts and communicative utility; games of ambiguous and manipulative

communication; computational language modeling for the detection of semantic drift and political consolidation; and recent developments in the automated analysis of social media.

Students will emerge with the basic knowledge to conduct contemporary multidisciplinary research into political communication.

Motivation and description

In a world with AI-accelerated political messaging via social media (Schippers, 2020), it is imperative that today's students in linguistics and Natural Language Processing (NLP) obtain a grasp of the state of the art in political communications studies and in the possibility for interdisciplinary research and development it provides (Németh, 2023). While the NLP field already has workshops devoted to political content, such as bias and hate speech, a theoretical perspective on how and why political ideas are communicated and distributed and what consequence it has for the current technological landscape remains largely absent.

At the same time, the study of political communications (PolComm) inherently relies on assumptions about how language is understood and instead focuses on the effects of political language. In some sense, linguistics — even pragmatics, considered as linguistics' way of discussing language *in situ* — is focused on the message itself, but PolComm tends to assume the process of understanding while ultimately focusing on mass behaviour. It is extremely difficult to find formal analyses of the form familiar to linguists of various stripes in the PolComm literature. What is missing is the connection between the participants in the process of political messaging (the speaker or "emitter" of political messages, the "targets" or listeners to that message, and the message itself), and the goals of politics and statecraft.

This course will bring the experience we have developed after several years of collaboration with media studies and political science researchers to the ESSLLI student population. We will outline the basic concepts of political communications research, theories of message, framing, identity, and political effect. We will discuss theories of political meaning from a semantic and pragmatic perspective, particularly theories of framing, implicatures, and speech acts from a political perspective.

We will then move onto experimental techniques in opinion research, particularly discussing survey techniques, media analysis and annotation approaches, and the ethical aspects of studying human beings through their political opinions. We will discuss current debates in representing political speech acts, the role of implicatures and personae, and concrete examples of politically manipulative speech in terms of the utility gained by speaker and hearer.

Finally, this course will discuss recent advances in natural language processing and the opportunities they represent for studying political communications, such as the role of meaning change detection to identify the development of communities of political discourse.

Political communication

Blumler (2015) describes political communication as a very multifaceted social phenomenon involving many levels and organs of society, particularly poltiicans and media, while constantly evolving inside normative boundaries. Consequently, PolComm is divided into many theoretical subdisciplines. For example, agenda-setting and priming theory examine how media coverage instigates voters to evaluate politicians and parties based on the issue being covered. Framing theory, however, discusses the way in which media coverage shapes citizen perceptions of how to evaluate the issues themselves when considering who and what to support.

According to Blumler, more recent streams of research react to the development of digital media and online culture. For example, digital media has opened further opportunities to study how politically-affected communities voice their concerns to the politicians and the rest of the public.

This and other perspectives will be covered in our course. The political dogwhistle is of particular interest, as it has served as a case study for game-theoretic approaches to political communication.

Slurs, dogwhistles, and other manipulative language

Politicians use complex and indirect language to achieve political aims. A lively literature has arisen to describe what it is that polticians seek to accomplish with this type of language. Slurs and dogwhistles can be seen as having a conventionalized component – a common and "innocent" meaning – and an unconventional component (potentially offensive to a group), with the main distinction being that in a dogwhistle, the offensive component is deniable or only apparently to an "in-group". Henderson and McCready (2021) argue against a conventional implicature account, instead relying on a Rational Speech Act (RSA) framework that is formally game-theoretic (Yoon et al., 2016).

Breitholtz and Cooper (2021) argue instead for a dialogical approach to dogwhistles that allow a return to inference as the source of dogwhistle effects using a novel representational approach. Noble et al. (2020) can be seen as a more general account of ideological communication using the Aristotelian notions of persona and topos to develop a formal account using Type Theory with Records (TTR).

In the limit, Sayeed et al. (2024) attempt to extend accounts of dogwhistles away from in-the-moment game-theoretic decisions to a larger theoretical concept of political utility over time, inviting opportunities to generalize to other PolComm phenomena. Extensions and generalizations of these ideas will be a topic of discussion in the course.

Methods in opinion research

Research methods in political communications and opinion research is a very large topic with many broadly available materials that draw on social science research in general (Boyle and Schmierbach, 2023). This course will cover the basic techniques with which there may be already broad familiarity, such as survey design and sampling, but also consider how to apply ideas from cognitive science and linguistic research, such as cloze tests and word replacement studies (McCarthy and Navigli, 2009) to PolComm research.

We will also discuss ethical questions about research into public opinion (Dong and Lian, 2021), especially in the context where such research may be gamified or exploited to further manipulate the public.

Computational approaches

There is an entire industry of content moderation, hate speech control, and debiasing in artificial intelligence and NLP, and this is arguably NLP applied to political communications. However, this type of work does not directly use NLP as a research method into the political incentives and effects to employ particular types of communication. Currently, the most developed literature once again concerns the matter of dogwhistles and deceptive meaning in media, particularly social media. Mendelsohn et al. (2023) create a typology of English dogwhistles from political speeches in the USA and explore the limitations of large language models (LLMs) in identifying them. Hertzberg et al. (2022) use a collection of Swedish dogwhistles and interpretations of the terms elected from sampled Swedish citizens to identify contextual vector spaces that distinguish "in-group" dogwhistle interpretations from their conventional meanings. Boholm and Sayeed (2023) characterize Swedish dogwhistles in terms of linguistic divergences between online social media groups that can be extracted from fine-tuning LLMs with data from more or less radicalized communities.

We will discuss the opportunity for these techniques to generalize to languages other than English (the overwhelming bulk of PolComm work, computational and otherwise, concerns the English-language USA political environment). We will also discuss how to extend these techniques to reinforcing theoretical connections between political incentives, effects, and expressions of persuasive phenomena beyond slurs and dogwhistles in public life.

Tentative outline

We expect each bullet point to take up 20-40 minutes each based on the available time per day.

Day 1 Political communication basics

- Foundational theories: media effects, media-state relationship
- "Modern" theories: voice and identity, holistic approaches
- Normative approaches and comparative analysis.

Day 2 Meaning in politics

- Framing and conceptual structures
- Implicatures in politics
- Pragmatics and speech acts in political contexts

Day 3 Experimental techniques

- Survey technique, panel design, statistical analysis
- Textual analysis of media (incl. social media)
- Ethical challenges in eliciting political opinion
- GROUP ACTIVITY: design hypothetical survey questions to assess participant interpretations of a controversial political keyphrase.

Day 4 Expectations and incentives

- Representation of political speech acts; rational approaches
- Enthymemes, topoi, and personae
- Slurs and dogwhistles from the perspective of communicative utility

Day 5 Computational approaches

- Identifying and managing online data resources (e.g., social media, archives)
- Vector-space and LLM approaches to representing political speech acts.
- Using measures of language change to track the development of political meaning.

Expected level and prerequisites

Expected level advanced

Prerequisites Students are expected to come from a variety of backgrounds but should have a grasp of basic concepts in semantics, pragmatics, and introductory NLP.

References

Blumler, J. G. (2015). Core theories of political communication: Foundational and freshly minted. *Communication Theory*, 25(4):426–438.

Boholm, M. and Sayeed, A. (2023). Political dogwhistles and community divergence in semantic change. In Tahmasebi, N., Montariol, S., Dubossarsky, H., Kutuzov, A., Hengchen, S., Alfter, D., Periti, F., and Cassotti, P., editors, *Proceedings of the 4th Workshop on Computational Approaches to Historical Language Change*, pages 53–65, Singapore. Association for Computational Linguistics.

Boyle, M. and Schmierbach, M. (2023). *Applied communication research methods: Getting started as a researcher*. Routledge.

Breitholtz, E. and Cooper, R. (2021). Dogwhistles as inferences in interaction. In Howes, C., Dobnik, S., Breitholtz, E., and Chatzikyriakidis, S., editors, *Proceedings of the Reasoning and Interaction Conference (ReInAct 2021)*, pages 40–46, Gothenburg, Sweden. Association for Computational Linguistics.

Dong, X. and Lian, Y. (2021). A review of social media-based public opinion analyses: Challenges and recommendations. *Technology in Society*, 67:101724.

Henderson, R. and McCready, E. (2021). Dogwhistles: Persona and ideology. In *Semantics and Linguistic Theory*, pages 703–719.

Hertzberg, N., Cooper, R., Lindgren, E., Rönnerstrand, B., Rettenegger, G., Breitholtz, E., and Sayeed, A. (2022). Distributional properties of political dogwhistle representations in Swedish BERT. In Narang, K., Mostafazadeh Davani, A., Mathias, L., Vidgen, B., and Talat, Z., editors, *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)*, pages 170–175, Seattle, Washington (Hybrid). Association for Computational Linguistics.

McCarthy, D. and Navigli, R. (2009). The english lexical substitution task. *Language resources and evaluation*, 43:139–159.

Mendelsohn, J., Le Bras, R., Choi, Y., and Sap, M. (2023). From dogwhistles to bullhorns: Unveiling coded rhetoric with language models. In Rogers, A., Boyd-Graber, J., and Okazaki, N., editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15162–15180, Toronto, Canada. Association for Computational Linguistics.

Németh, R. (2023). A scoping review on the use of natural language processing in research on political polarization: trends and research prospects. *Journal of computational social science*, 6(1):289–313.

Noble, B., Breitholtz, E., and Cooper, R. (2020). Personae under uncertainty: The case of topoi. In Howes, C., Chatzikyriakidis, S., Ek, A., and Somashekarappa, V., editors, *Proceedings of the Probability and Meaning Conference (PaM 2020)*, pages 8–16, Gothenburg. Association for Computational Linguistics.

Sayeed, A., Breitholtz, E., Cooper, R., Lindgren, E., Rettenegger, G., and Rönnerstrand, B. (2024). The utility of (political) dogwhistles – a life cycle perspective. *Journal of Language and Politics*.

Schippers, B. (2020). Artificial intelligence and democratic politics. *Political Insight*, 11(1):32–35.

Yoon, E. J., Tessler, M. H., Goodman, N. D., and Frank, M. C. (2016). Talking with tact: Polite language as a balance between kindness and informativity. In *Proceedings of the 38th annual conference of the cognitive science society*, pages 2771–2776. Cognitive Science Society.

Information required of course proposal

Appeal outside main discipline

This is an inherently interdisciplinary topic taught by researchers in linguistics and NLP that now have extensive experience collaborating with theorists and practitioners of political communications studies. We expect to draw in a broad audience.

Experience of proposers

Asad Sayeed has taught four previous successful courses on diverse topics at ESSLLI: in 2013 (opinion mining), 2014 (cognitive workload), 2017 (distributional semantics), and 2019 (quantifier scope and incrementality). Asad is the principal investigator of the Gothenburg Research Initiative for Politically Emergent Systems (GRIPES) supported by the Marianne and Marcus Wallenberg Foundation (2020-2025); this is a collaboration with researchers from linguistics and political communication. He received his Ph.D. in 2011 in computer science from the University of Maryland, College Park, with a thesis on linguistic approaches to opinion mining.

Ellen Breitholtz participated in the FADLI (2017) and TYTLES (2015) ESSLLI workshops and recently published the monograph *Enthymemes and Topoi in Dialogue - the use of common sense reasoning in conversation* (Brill, 2020). Ellen is a participating researcher in the GRIPES project as well as having led a previous Swedish government-funded project on interaction in dialogue. She received her Ph.D. in linguistics at the University of Gothenburg.

Both of them have participated in organizing major conferences and workshops, such as the International Conference on Computational Semantics (IWCS).

Evidence of being excellent lecturers

Both Asad Sayeed and Ellen Breitholtz are Associate Professors in the Department of Linguistics, Philosophy, and Theory of Science at the University of Gothenburg, and both of them have many years of experience as teachers and lecturers at the university level.

Asad taught computer science laboratory courses at the University of Maryland, College Park as a Ph.D. student, then advanced seminar courses in linguistics and NLP at Saarland University, as a postdoc. At the University of Gothenburg, he teaches a regular cycle of courses in NLP and machine learning, and he is current Program Chair of the Master of Language Technology program.

Ellen regularly teaches courses in undergraduate linguistics, particularly pragmatics, at the University of Gothenburg. She has also fulfilled the role of director of studies for the linguistics program.