



# Public Discourse and Autocratization: Infringing on Autonomy, Sabotaging Accountability

Seraphine F. Maerz,  
Carsten Q. Schneider

February 2021

Working Paper

SERIES 2021:112

THE VARIETIES OF DEMOCRACY INSTITUTE



UNIVERSITY OF GOTHENBURG  
DEPT OF POLITICAL SCIENCE

**Varieties of Democracy (V-Dem)** is a new approach to conceptualization and measurement of democracy. The headquarters – the V-Dem Institute – is based at the University of Gothenburg with 20 staff. The project includes a worldwide team with 5 Principal Investigators, 19 Project Managers, 33 Regional Managers, 134 Country Coordinators, Research Assistants, and 3,200 Country Experts. The V-Dem project is one of the largest ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute  
Department of Political Science  
University of Gothenburg  
Sprängkullsgatan 19, Box 711  
405 30 Gothenburg  
Sweden  
E-mail: [contact@v-dem.net](mailto:contact@v-dem.net)

V-Dem Working Papers are available in electronic format at [www.v-dem.net](http://www.v-dem.net).

Copyright ©2021 by authors. All rights reserved.

# Public discourse and autocratization: Infringing on autonomy, sabotaging accountability\*

**Seraphine F. Maerz**

V-Dem Institute, Department of Political Science, University of Gothenburg

**Carsten Q. Schneider**

Department of Political Science, Central European University, Vienna, Austria

---

\*Both authors contributed equally. **Corresponding author:** Seraphine F. Maerz (seraphine.maerz@v-dem.net). We recognize support by the Knut and Alice Wallenberg Foundation to Wallenberg Academy Fellow Staffan I. Lindberg, Grant 2018.0144, as well as by internal grants from the Vice-Chancellor's office of the University of Gothenburg. We are grateful for helpful comments by the V-Dem Institute's researcher group and Philipp Lutscher.

## Abstract

Ever since its existence, democracy is periodically diagnosed to be in crisis. When such crises are analyzed, reference has usually been made to the malfunctioning of core democratic *institutions* and the *behavior* of actors. We believe that this fails to capture a crucial element of the current crisis. It is *words*, not (only) deeds that present the contemporary challenge to liberal democracy even before such challenges materialize in institutional change. There are strong theoretical and empirical reasons to take into account the public discourse of leading political figures when studying democratic decline. In this article, we conceptualize illiberal and authoritarian public rhetoric as language practices that harm democracy. Such a practice-orientated approach allows for fine-grained assessments of autocratizing tendencies in democracies that goes beyond an exclusive focus on democratic institutions. By using a corpus of 4 506 speeches of 25 heads of government from 22 countries and data on government disinformation and party identity, we empirically illustrate with dictionary analysis and logistic regression that illiberal and authoritarian public rhetoric are tightly linked to autocratization and should no longer be ignored.

## 1 Introduction

“Language changes always manifest social changes – but language changes (or changes in language behaviour) can also trigger social changes” (Wodak, 1989, p. XV).

In this article, we argue that the public discourse<sup>1</sup> of core political leaders should be taken into account when analyzing the state of democracy. Through their signalling via speeches, political leaders can either support or undermine the political structures in place. Empirically and logically, the norm is that the values incorporated in political institutions and public speeches coincide and thus reinforce each other (Linz, 1978; Levitsky and Ziblatt, 2018; Dawson, 2018). In liberal democracies, leaders tend to choose a language that is in line with the normative commitment this form of government entails (Maerz and Schneider, 2019). Likewise, in autocracies of various forms, leaders in their speeches usually reflect and embody the essence of what their regime stands for and is based on (Maerz, 2019).

We are interested in deviations from this norm. We aim at detecting instances in which the public rhetoric of central political leaders is out of sync with the formal-institutional norm system of a political regime because they provide clues for likely changes to the nature of that regime. We hold that such periods of incongruence cannot last forever. One option is that institutions prevail and the political leader is removed or at least changes the tone - through elections in democracies and by other, often violent means, in non-democracies. Another scenario is that the political leader succeeds in bringing the institutional system in line with the values expressed in his or her public speeches.

By bringing public discourse (back) into the study of political regimes and with a particular focus on democracies in crisis, we investigate into the question of how illiberal and authoritarian public rhetoric relate to autocratization.<sup>2</sup> We conceptualize illiberal public rhetoric as a practice of infringing on autonomy and dignity by exceedingly emphasizing nationalist, paternalist, and traditionalist values in official speeches or statements. Authoritarian public rhetoric is the practice of sabotaging accountability and disabling voice by officially promoting disinformation or anti-pluralism (Glasius, 2018). One advantage of this two-fold framework is that it captures not only those autocratizers who harm democracy with misanthropic speeches - Hungary’s current prime minister Viktor Orbán and his anti-immigration discourse being a prime example -, but also accounts for the ruthless spread of disinformation, such as, for instance, former US president Donald Trump.

We believe that our focus on public rhetoric is timely chosen. Few observers contest that the calamitous events on Capitol Hill on January 6, 2021, demonstrate the effect of public speeches of leading politicians on shifting behavioral norms. Surprisingly, though, theories on regime transitions attribute little to no role on how political actors speak. Liberal democracy is deemed to be

---

<sup>1</sup>We define public discourse broadly as the entity of official communication put forward by a government and its agents. We use public discourse, public communication, and public rhetoric interchangeably.

<sup>2</sup>Autocratization, the reverse concept of democratization, is an overarching term for any substantial decline of democratic regime attributes in democracies or autocracies (Lührmann and Lindberg, 2019).

in crisis - again (Maerz et al., 2020b; Kirsch and Welzel, 2018; Schmitter, 2015).<sup>3</sup> And according to some, symptoms of crisis are as old as liberal democracy and probably an inherent feature of this form of political regime (O'Donnell, 2007a). What seems to distinguish the current crisis is that it is driven from within and atop. In most instances, the driving forces of autocratization are not primarily foreign actors, nor fringe opposition groups at the margin of society. Instead, liberal democracy is questioned by actors in government. Another feature of the current crisis is that it has been proceeding slowly and in small steps and it therefore has long been left undiagnosed and untreated. Early symptoms of the malaise have been ignored or mis-interpreted. In this article, we argue that listening to what rulers say can detect crises of democracy in the making. Public rhetoric expresses declarations of intent. In the past, when democracy crises were diagnosed, reference has usually been made to the malfunctioning of core democratic *institutions* and the actual *behavior* of different actors, who - purposefully or inadvertently - undermine democracy with their (non-) action (Levitsky and Ziblatt, 2018; Merkel and Kneip, 2018, p. 23). This, so we argue, leaves a blank spot in the democracy crisis literature: the use of public *rhetoric* as both a symptom and a driver of the democratic malaise. Neglecting the positioning of leading politicians towards the foundational principles and practices of democracy is particularly troublesome in analyzing the current crisis. One of its distinct features is that in many cases democracy is verbally attacked from its helm rather than by opposition parties or fringe groups.

Our claim that public speeches matter for the future of any political regime is, of course, far too plausible to have gone unnoticed. It already features in the classic literature on democratic demise (Linz, 1978) and is rediscovered in contemporary approaches (e.g. Schedler, 2019). One reason for why this analytic approach has, so far, not been employed more broadly in large-N comparative regime analyses seems to be of practical nature: until recently, gathering and then analyzing huge amounts of text has presented itself with prohibitively high costs. The development of text-as-data approaches and computational solutions for harvesting unprecedented amounts of texts from the web have dramatically reduced these costs (Grimmer and Stewart, 2013; Lucas et al., 2015; Maerz and Puschmann, 2020). While these new tools cannot - and should not - replace human input and interpretation, they can, however, radically augment a researcher's capacity to analyze large quantities of text, a property that we capitalize on with our approach.

We proceed as follows: In section 2 we present various strands of literature in which public discourse takes center stage in analyzing social reality, in general, and political processes, in particular. In section 3, we outline how we conceptualize public rhetoric. The empirical illustrations of our argument are discussed in section 4 before we conclude by providing an outlook on the role of public rhetoric as a additional perspective in the political regime change literature.

---

<sup>3</sup>For a contrarian view, see Merkel (2014).

## 2 Bringing public discourse (back) in the study of regime types and their change

Language matters. This commonsensical insight has made inroads into many different approaches in the social sciences. Discursive institutionalism (Schmidt, 2008), conversation analysis (Hutchby and Wooffitt, 2008), sociolinguistic analysis (Wodak, 2009), discourse-historical approach (Wodak, 2001), (critical) discourse analysis (Wodak and Meyer, 2009; Maynard, 2013), thematic analysis, and content analysis (e.g. Mayring, 2010) are just some of the most prominent among those. Also in the more positivistically inclined realm of the social sciences, the spoken word is becoming the center of analytic attention, as the boom of text-as-data studies testifies (Grimmer and Stewart, 2013; Lucas et al., 2015). To this, one can add traditional fields like that of political communication and media studies and it becomes clear that there is a widely shared understanding that not just institutions and observable behavior matter for analyzing politics. We argue that the study of political regimes and their evolution and change should not be an exception and, instead, start to embrace public discourse in their analytic tool kit.

Critics of an increased focus on rhetoric might object that what is said matters less than what is done, that actors might simply pay lip service, and that language largely remains without consequences in the "real" world. To this Monroe and Schrodt (2008, p. 351) reply that "[t]ext is arguably the most pervasive – and certainly the most persistent – artifact of political behavior." And "[...] what political actors say, more than the behaviour they exhibit, provides evidence of their true inner states" (Benoit, 2019, p. 1).

An analytic focus on actors is not new in the regime change literature. It was pushed into the background with the shift from research on regime transitions to their consolidation and the concomitant shift in focus away from actors to institutions and their behavior-regulating norms (Schmitter and Guillhot, 2000). But even before the surge of third-wave democracy studies, influential writings on the breakdown of democracy stipulated that actors and their publicly expressed attitude towards democracy should play a prominent analytic role. In his seminal book, Linz (1978) argues that, while it is normal in democracies that political competition usually entails depicting the opponent as a representative of narrow or misguided interests, it is the "[...] style, intensity, and fairness in conducting these actions [that] mark the distinction between loyal and disloyal oppositions" (Linz, 1978, p.31).

Forces loyal to democracy are crucial for its survival, according to Linz. Distinguishing between loyal, semi-loyal, and disloyal is not easy and for this task Linz attributes high importance to the public rhetoric of actors.<sup>4</sup> For him, it suffices to rhetorically cross the boundaries of democratic norms to become semi- or disloyal and thus put in peril the survival of a democratic regime.

---

<sup>4</sup>Loyal forces are defined by, among other things: "1) An unambiguous public commitment to achievement of power only by electoral means and a readiness to surrender it unconditionally to the other participants with the same commitment. [...] 4) An unambiguous rejection of the rhetoric of violence to mobilize supporters in order to achieve power, to retain it beyond the constitutional mandate, or to destroy opponents [...]. The defense of democracy must be carried out [...] without arousing popular passion and political vigilantism." Further, "[...] a source of perception leading to questions about the loyalty of the party to the system, is a willingness to encourage, tolerate, cover up, treat leniently, excuse, or justify the actions of other participants that go beyond the limits of peaceful, legitimate patterns of politics in a democracy" (Linz, 1978, p.32).

Quite obviously, attentive observers of democracies currently deemed to be declining, such as in Brazil, Poland, or the US,<sup>5</sup> find numerous instances of public discourses by heads of government in these countries that fit the bill of semi-loyal or even disloyal discourse that Linz spelled out over 40 years ago. Despite its analytic relevance for studying regime change, public rhetoric is still largely neglected by the current literature. Our argument is that this shortcoming should be rectified. Only slowly do actors and their public discourse find their way back in as an analytic category. Levitsky and Ziblatt (2018), in their influential book on how democracies die, follow Linz in attributing importance to how politicians in democracies speak about core rules of the democratic game. Their checklist for authoritarian behavior contains various entries that refer to what political leaders say rather than to what they do (Levitsky and Ziblatt, 2018, p.23-24).

Schedler (2019) explicitly spells out the need to embrace the language dimension in order to understand current processes in political regimes. He argues that there is a mismatch between what predominant theories of regime change postulate as relevant factors, on the one hand, and the widespread impressions of what is currently going wrong in democracies such as the US under Trump, Bolsonaro's Brazil, or Poland under Kaczynski, on the other.<sup>6</sup> We concur with Schedler in concluding that current developments in several democracies make us to acknowledge that political actors, political language, and normative commitments are relevant factors for whether or not democracies die. The democratic regime literature's excessive focus on structures and institutions, though, has made the discipline overlook the long-acknowledged "[...] possibility that deviant actors may disrupt structural equilibria always exists" (Schedler, 2019, p. 437).

All in all, with Schedler (2019, p. 452) we see the need for "bringing language back in" to the analyses of political regimes and their transformations.<sup>7</sup>

### 3 Conceptualizing illiberal and authoritarian public rhetoric

We theorize public rhetoric to be a part of the *everyday practices* that fill institutions with life, and shape the character of a political regime (Glasius, 2018). *Illiberal* and *authoritarian* public rhetoric are two different but conceptually and empirically strongly overlapping and mutually not exclusive categories, as depicted in the Venn diagram in Figure 3.

The illiberal category comprises all practices that infringe on autonomy and dignity (Glasius, 2018, p. 531).<sup>8</sup> The rhetoric manifestations of these practices are speeches or official statements

---

<sup>5</sup>Hungary, often also mentioned in such lists, is omitted because according to most regime measures it no longer is classified as a liberal democracy.

<sup>6</sup>Many of our democratic worries about Donald Trump seem to arise from factors that do not bother us in the comparative study of democracy: his unconstrained discourse, his violation of linguistic norms, and his programmatic amorality." (Schedler, 2019, p. 434). And further: "[...] our perceptions of the threats Trump poses to US democracy do not match our main theories of democratic instability" (Schedler, 2019, p. 435).

<sup>7</sup>As further examples for this nascent approach, Schedler (2019, p. 452) lists studies by Aalberg et al. (2017); Hawkins and Kaltwasser (2017); McCoy et al. (2018).

<sup>8</sup>Regarding our use of the terms liberal and illiberal we emphasize that we do not refer here to any aspects of the economic left-right but strictly limit ourselves to the narrow definition of political liberalism and illiberalism, as conceptualized here. Hence, by liberal we do not mean economic (neo)liberalism, such as the ideology of free markets. Illiberal neoliberal regimes are, thus, not an oxymoron but describe regimes that propagate illiberal political values



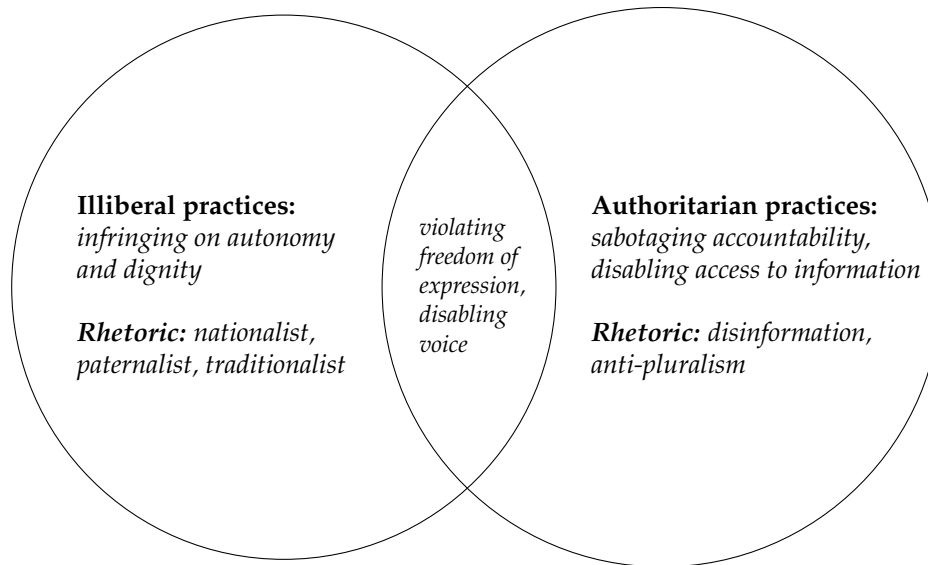


Figure 1: Illiberal and authoritarian public rhetoric (adjusted from Glasius, 2018, p. 531)

by the government and its agents which exceedingly emphasize nationalist, paternalist, and traditionalist values.<sup>9</sup> One example for such an illiberal public rhetoric is Hungary's Viktor Orbán.<sup>10</sup> According to Orbán, the future of Hungary, Europe, and of Western civilization is endangered by mass migration of Muslims, helped by sleep-walking European leaders under the influence of some semi-hidden hands of world conspirators. We claim that it is unlikely to remain without consequences for the functioning of democracy in a country, if its first representative publicly speaks the way the Hungarian Prime Minister does - even more so, if he speaks like this over several years, on virtually any occasion, and essentially every channel of communication.

The authoritarian category contains all practices that sabotage accountability<sup>11</sup> and disable access to information (Glasius, 2018, p. 531). The rhetoric manifestations of these practices are official disinformation campaigns and other anti-pluralists communication strategies of regimes to flood, manipulate, and control the public sphere of their countries. Such manipulations of how citizens (can) talk and think about politics are typical phenomena to be observed in strictly authoritarian regimes (Dukalskis, 2017). Yet, in the age of "fake news", official attempts of disabling access to reliable information are increasingly observed also in Western liberal democracies.

One prominent, yet not the only (and most sophisticated) example of authoritarian language practices in contemporary democracies is former US president Trump and his thousands of misleading and false claims on a wide array of topics - as counted by major news outlets such as The Washington Post (2020) or The Guardian (2020). During the Covid-19 pandemic, for in-

while pursuing neoliberal economic and social policies.

<sup>9</sup>As our empirical illustrations show, there is a qualitative difference between conservative values expressed in public speeches (e.g. a more or less balanced reference to traditional values while also signalling tolerance for alternative views) and, for example, an excessive emphasis of nationalism in illiberal public rhetoric.

<sup>10</sup>For some illustrations, we refer the reader to the Appendix.

<sup>11</sup>For various forms of accountability vital to democracy, see O'Donnell (2007b) or Lührmann et al. (2020).

stance, Trump repeatedly downplayed the severity of the virus or spread disinformation about testing and (unproven) treatments such as his idea of injecting antibacterial cleaning agents to cure Covid-19 infections (Timm, 2020b,a; Paz, 2020).<sup>12</sup> Beside this and other serious and irresponsible disinformation on Covid-19 matters by Trump, his unsubstantiated allegations of electoral fraud after Joe Biden's victory in the 2020 presidential elections signaled yet another peak in his falsehoods. As Kreuz and Windsor (2021) explain, throughout his presidency Trump stood out with his unique speaking style - yet, it was particularly during the two weeks before the Capitol riot in early January 2021 that his language style changed to more inflammatory rhetoric. His public call to action to "[f]ight like hell, and if you don't fight like hell, you're not going to have a country anymore" (cf. Blake, 2021) was followed by the storming of the Capitol by his supporters.

The examples of Orbán and Trump illustrate the difference between illiberal and authoritarian public rhetoric and show why it makes sense to analytically distinguish between the two practices. Illiberal practices are typically motivated by a more or less coherent nationalist and deeply conservative ideology. Authoritarian practices, in contrast, are characterized by arbitrary falsehoods. Yet, as Figure 3 illustrates, the two categories also overlap. Orbán's conspiracy beliefs on who "really" runs the EU, for example, are similar to Trump's false accounts. While Trump is known for his contradictory statements on all kind of political and economic topics (Kruse and Weiland, 2016), he also disseminates highly nationalist views (Mudde, 2020). Political leaders might rely on just one of the two language practices. Yet, if an overall nationalist, paternalist, and traditionalist language is combined with official disinformation campaigns and anti-pluralist rhetoric, the population is substantially hindered from accessing reliable information and freely engaging with public discourses. Hence, as Figure 3 shows, the intersection of illiberal and authoritarian practices concerns violations of freedom of expression and attempts of disabling voice.

Liberal and democratic public rhetoric are the opposite poles to illiberal and authoritarian public rhetoric. They, too, represent two different but heavily overlapping categories. Liberal public rhetoric is characterized by an inclusive, egalitarian, and non-discriminatory language. Democratic public rhetoric is reliable, fact-based, and emphasizes pluralist values.<sup>13</sup>

In sum, the everyday practices in political regimes also consist of public rhetoric by political leaders. Such rhetoric comes in two distinct yet overlapping forms: liberal-illiberal and democratic-authoritarian. Capturing these elements of political regimes fills a gap in the regime change literature and has several advantages. First, overlooking rhetoric has been wrong all along. For democracy, a reason-based and respectful dialogue of leaders, both with and in the public, is a *sine qua non* (Habermas, 1984).<sup>14</sup> Second, by analyzing public speeches as *not yet institutionalized practices* of a political regime, changes in the regime do not have to be thought of as *regime change* only, but can also be conceptualized and operationalized as more *gradual regime transformations*.

---

<sup>12</sup>For a detailed documentation of official disinformation campaigns during the Covid-19 pandemic see Edgell et al. (2020a).

<sup>13</sup>While in this article we mainly conceptualize and empirically analyze how illiberal and authoritarian public rhetoric relates to autocratization, future research might further explore the relationship between liberal and democratic public discourse and democratization.

<sup>14</sup>For empirical efforts to capture the quality of discourses, see Steenbergen et al. (2003).

Radical and often quite dramatic and visible events are often just the endpoint of prior silent erosion. Third, and related, public rhetoric should be established not as a direct indicator of, but as a crucial proximate factor for political regimes and their transformations. Similar arguments have been made about proximate causes being so proximate that they are considered essential parts of the measurement strategies for gauging the likely development of political regimes. Consider the role of coup attempt in the democratic consolidation literature. They are not a direct indicator of failed democratic consolidation. It depends on the reaction to the coup attempt: do democratic actors and institutions resist or succumb? If the latter, the coup attempt is an indicator of democratic breakdown; if the former, it is considered an indicator of the consolidation of democracy (O'Donnell, 1996; Schedler, 2001). Public rhetoric plays a similar analytic role in our framework for capturing present reality in political regimes and anticipating their likely development: do actors and institutions resist anti-democratic rhetoric or do they succumb? If the former, democracy prevails, if the latter, it fails.<sup>15</sup>

#### 4 Empirical illustrations

In the first part of this section, we empirically show how *illiberal public rhetoric* manifests itself in official speeches of political leaders. Based on a dictionary application, we develop the Illiberal Speech Index (ISI). The second part of this section looks into the other language practice we conceptualized above: *authoritarian public rhetoric*. We test with a regression analysis how authoritarian public rhetoric - measured as official disinformation campaigns on social media and anti-pluralist rhetoric of ruling parties - relates to the autocratization of regimes.

##### 4.1 *Illiberal public rhetoric: Nationalist, paternalist, and traditionalist speeches*

We analyze the language of political leaders as the most crucial supporters, or underminers, of political institutions and the norms and values they enshrine. For measuring illiberal public rhetoric, we zoom in on the public speeches of heads of government as the actors with most power to defend, or undermine, a political regime. Especially in non-democratic settings, the impact of public rhetoric by heads of government is exacerbated because access to the broader public is controlled by them, making it more unlikely that citizens are exposed to rhetoric countering that of the government (Dukalskis, 2017). Such trends are also under way in those still-democracies that are currently regressing towards more authoritarian forms of ruling.

To capture recurrent and persistent patterns in public rhetoric, we collect a large and heterogeneous pool of public speeches. The amount and diversity of speeches is crucial for our claim that public discourses have an impact on political regimes. As Jaeger and Maier (cited in Wodak, 2009, p. 200) put it: "When analyzing power effects of discourse, it is important to distinguish between the effects of a text and the effects of a discourse. A single text has minimal effects [...]. In

---

<sup>15</sup>The obvious difference between coups and public rhetoric is that the former are short and high-salience events whereas the other stretches over long periods of time with no clear beginning or peak. This is why coups cause sudden and corrosive public rhetoric more likely slow regime deaths.

contrast, a discourse, with its recurring contents, symbols, and strategies, leads to the emergence and solidification of 'knowledge' and therefore has sustained effects."

Table 1 summarizes our corpus of speeches by displaying the speaker, country, status of the regime, the number of speeches scraped, the dates of the first and the last speech in our data, and the source from where the speeches were scraped. The corpus consists of 4 506 speeches from 25 heads of government of 22 countries during the period of 1999 to 2020. With our aim of investigating the relationship between illiberal public rhetoric and democratic decline, we selected many speeches from heads of government in political regimes that are experiencing autocratization (11 speakers in 9 regimes). Some of these regimes already transitioned into autocracy before or during the time periods covered by our corpus and are further regressing, such as Serbia, Turkey, Venezuela, and Hungary. Other regimes are declining democracies. Overall, we label the status of these regimes "regressing". In order to calibrate our illiberal speech index, we include several heads of government of clearly democratic (e.g. Sweden) and autocratic (e.g. North Korea) regimes that do not experience an episode of autocratization.<sup>16</sup>

The speeches of each speaker in our corpus are of varying number and length. The majority of them date back not more than a few years,<sup>17</sup> enabling us to assess illiberal public rhetoric during the most current crisis of democracy. The corpus consists of a mix of speeches for each speaker in front of highly diverse audiences (national or international) and for different occasions. This heterogeneity of the speeches is crucial for the validity of our dictionary analysis because it helps to prevent context-specific bias (Maerz, 2019). We applied a broad understanding of political speeches that includes political statements such as those typically done during (joint) press conferences and interviews or other more spontaneously produced documents.<sup>18</sup> Speeches are collected from the official websites of the heads of government<sup>19</sup> by using web-scraping techniques in *R*. The majority of the speeches is in English language (occasionally report-like translations). The speeches of Jair Bolsonaro and Recep Tayyip Erdogan were posted in a different language, which is why we translated them with the help of machine translation tools.<sup>20</sup> All texts with mixed contents - e.g. comprising statements of other politicians - were filtered accordingly. We pre-processed the speech corpus only minimally for the dictionary application (letters converted to lower case, punctuation removed, and misleading collocations erased).<sup>21</sup>

---

<sup>16</sup>Our regime classifications and the operationalization of autocratization are based on the Episodes of Regime Transformation (ERT) dataset of Edgell et al. (2020b), as also further explained in Section 4.2.

<sup>17</sup>The official usage of the Internet by political regimes to legitimize their rule is only a recent trend, cf. Maerz (2016).

<sup>18</sup>The material of the Polish prime ministers Beata Szydlo and Mateusz Morawiecki consists of comparatively short direct quotations which are regularly published in the news section of the government's official websites.

<sup>19</sup>One exception is the collection of speeches given by Donald Trump which we scraped from factba.se.

<sup>20</sup>de Vries et al. (2018) show that for bag-of-words text mining approaches the outcome of machine-translated and human-translated texts highly overlap.

<sup>21</sup>All analyses were done in *R*; replication files are available upon request.

Speaker	Country	Status	Speeches	First	Last	Source
Angela Merkel	Germany	Democracy	69	2014-02	2020-07	bundesregierung.de
Erna Solberg	Norway	Democracy	243	2013-11	2020-10	regjeringen.no
Jacinda Ardern	New Zealand	Democracy	66	2017-11	2020-07	beehive.govt.nz
Justin Trudeau	Canada	Democracy	151	2015-12	2020-10	pm.gc.ca
Leo Varadkar	Ireland	Democracy	216	2017-06	2020-06	gov.ie
Stefan Löfven	Sweden	Democracy	79	2014-10	2020-10	government.se
Alexander Vučić	Serbia	Regressing	18	2019-07	2020-10	predsednik.rs
Andrej Babis	Czech Republic	Regressing	30	2018-01	2020-09	vlada.cz
Andrzej Duda	Poland	Regressing	18	2015-11	2019-11	prezydent.pl/
Beata Szydło	Poland	Regressing	24	2015-11	2017-09	premier.gov.pl
Donald Trump	USA	Regressing	286	2017-02	2020-09	factba.se
Jair Bolsonaro <sup>a</sup>	Brazil	Regressing	291	2019-01	2020-10	gov.br
Mateusz Morawiecki	Poland	Regressing	39	2017-12	2020-08	premier.gov.pl
Narendra Modi	India	Regressing	83	2014-11	2020-10	pmindia.gov.in
Nicolas Maduro Moros	Venezuela	Regressing	8	2015-07	2020-09	pmppre.gob.ve
Recep Tayyip Erdogan <sup>a</sup>	Turkey	Regressing	745	2014-08	2020-10	tccb.gov.tr
Viktor Orbán	Hungary	Regressing	598	2014-06	2020-10	kormany.hu
Dmitry Medvedev	Russia	Autocracy	131	2008-05	2012-04	kremlin.ru
Emomali Rahmon	Tajikistan	Autocracy	102	2004-09	2020-09	president.tj
Ilham Aliyev	Azerbaijan	Autocracy	613	2010-04	2020-09	president.az
Kim Jong Un <sup>b</sup>	North Korea	Autocracy	24	2012-04	2016-09	naenara.com.kp
M.b.R. Al Maktoum	UAE	Autocracy	94	1999-04	2019-12	sheikhmohammed.ae
Paul Biya	Cameroon	Autocracy	56	2013-01	2020-01	prc.cm
Vladimir Putin <sup>c</sup>	Russia	Autocracy	433	2000-01	2020-09	kremlin.ru
Yoweri Museveni	Uganda	Autocracy	89	2011-05	2020-10	statehouse.go.ug
<b>Overall corpus:</b>			<b>4506</b>			

<sup>a</sup> The speeches of Jair Bolsonaro and Recep Tayyip Erdogan are machine translated using the Google translation API.

<sup>b</sup> The material of North Korea consists of official letters rather than speeches.

<sup>c</sup> The speeches cover Putin's four presidential terms.

Table 1: Overview of speakers and speech corpus

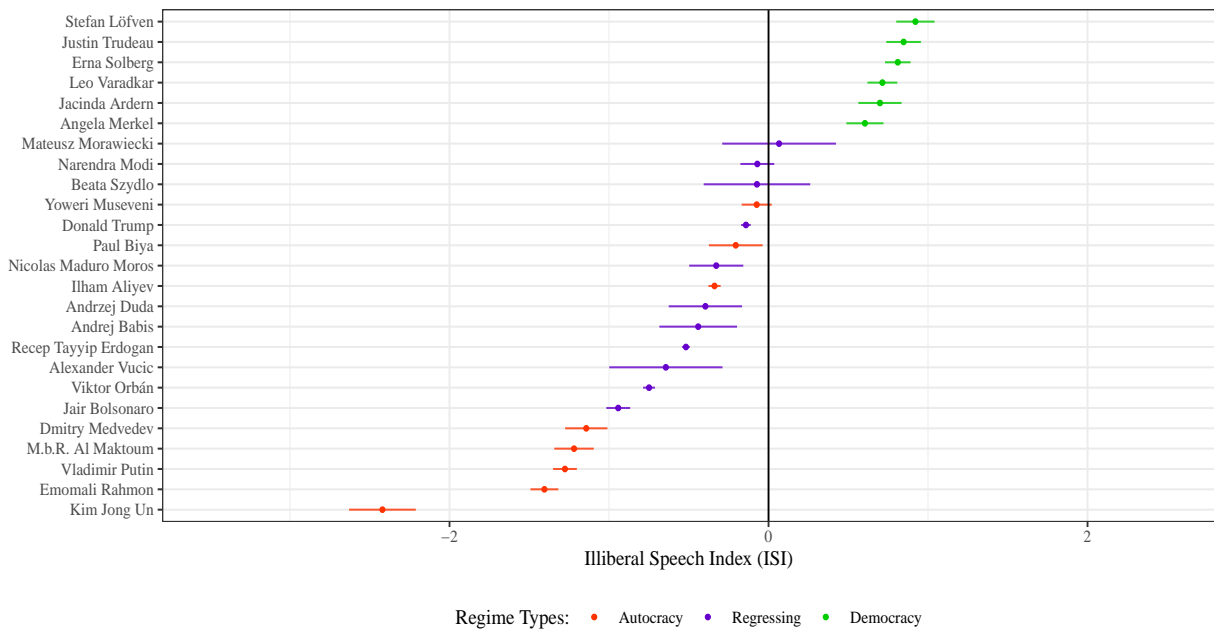
The core of our argument consists in the claim that speeches can be ranked on a scale of liberalism. The speech index that we propose has as its endpoints in purely illiberal and liberal speeches. We use a dictionary approach to classify each speech and, based on that, each speaker of these speeches on our Illiberal Speech Index (ISI) on public rhetoric. The dictionary approach requires a collection of liberal and illiberal terms. Those terms are provided by researchers. Each speech is attributed a value on the speech index based on the relative frequency with which liberal and illiberal terms from the dictionary occur in it. If liberal and illiberal terms are in a balance, the speech receives a score of 0 on our scale. If liberal terms predominate, the score turns positive and approaches the purely liberal speech. If illiberal terms occur more frequently, the index score turns negative, approaching the ideal-typical illiberal speech.

Our dictionary for the analysis is adopted from Maerz and Schneider (2019).<sup>22</sup> The dictionary includes 129 terms that are clearly associated with either *liberal* or *illiberal public rhetoric*, as conceptualized above. Examples of dictionary terms that measure liberal public rhetoric are free-, liberal-, tolera- and words such as discriminat-, authoritarian-, and repressi-, typically used to insist on more civil and political rights. In contrast to this, illiberalism is measured with key words such as anarch-, chaos, destabili-, or patriot- and pride. Maerz and Schneider (2019) com-

<sup>22</sup>See the Appendix for the list of dictionary terms. We slightly adjusted the dictionary and deleted six terms which we identified as ambiguous in a qualitative key-words-in-context check in parts of our corpus. While the dictionary was comprehensively validated by Maerz and Schneider (2019), we deem such additional qualitative checks and re-validation as crucial every time dictionaries are re-used for other corpora of texts (Maerz and Puschmann, 2020).

prehensively test the validity of the dictionary with qualitative hand-coding, unsupervised topic modeling, network analysis, and by illustrating its criterion validity. To illustrate the *relative proportional difference* between the speakers in the analysis, we make use of Lowe et al. (2011)'s spatial model of logit scaling. For a more detailed and technical description of the statistical procedure for creating the index, see Maerz and Schneider (2019).

Figure 2 displays the Illiberal Speech Index (ISI) for the cases listed in Table 1 as the result of our dictionary analysis. From top to bottom, speakers are ranked from highest (most liberal) to lowest (most illiberal) scores on our speech index. The markers indicate the average score across each speaker's speeches, with 95 percent confidence intervals expressing the uncertainty attached to each point estimate. The vertical line at index value 0 separates liberal (right) from illiberal (left) speakers. The marker color separates speakers in democratic (green), regressing (purple), and autocratic (red) regimes, based on data from Edgell et al. (2020b).



The lines represent 95% confidence intervals.

Figure 2: The Illiberal Speech Index (ISI): Scaling the ideological content of all speeches per speaker

As Figure 2 shows, all heads of government from democratic, non-regressing regimes score as clearly liberal speakers on our index. In contrast to this, the speakers of stable autocratic regimes are at the (far) illiberal end of the scale. Together, these classifications are in line with what one would expect based on common sense and case expertise, fostering the face validity of our index. While three leaders of regressing regimes (Beata Szydlo, Mateusz Morawiecki, and Narendra Modi) cannot be classified on our scale with statistical certainty,<sup>23</sup> all other regressing regimes are

<sup>23</sup>Their point estimates as well as the point estimate of Uganda's Museveni have a confidence interval that crosses the 0 value of our scale, that is, statistically the content of their speeches cannot be unequivocally classified as liberal or illiberal.

ruled by illiberal speakers. For some of the illiberal speakers in regressing regimes we were able to collect higher numbers of speeches (e.g. Jair Bolsonaro, Viktor Orbán, Recep Tayyip Erdogan, Donald Trump, cf. Table 1) which is why their confidence interval is comparatively small, lending further reliability to their positions on the illiberal side of the scale.

The data shown in Figure 2 show that for the cases included in our analysis there is a stark qualitative difference between the liberal public rhetoric used in functioning democracies such as Sweden, Norway, or Germany, on the one hand, and the illiberal language of heads of government in declining regimes such as Turkey under Erdogan or Hungary under Orbán. Figure 2 also reveals that for several regressing democracies, for instance the US under Trump or Brazil under Bolsonaro, their illiberal public rhetoric shows remarkable similarities to that in fully-fledged autocracies like Azerbaijan or Russia.

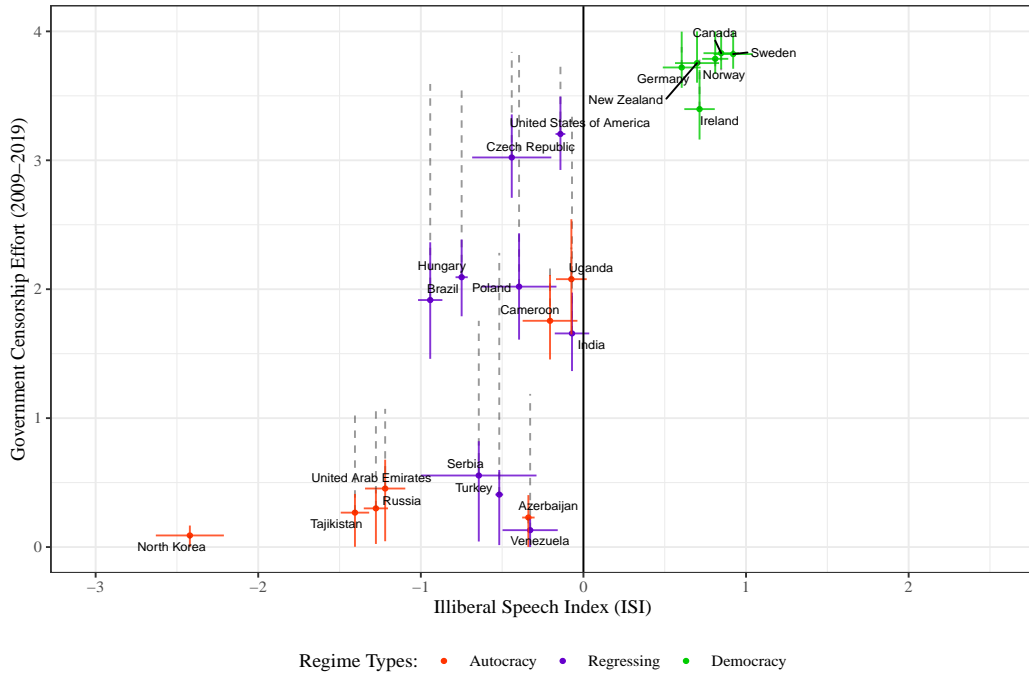
To further illustrate the capability of the Illiberal Speech Index to signal ongoing regime transformation processes already before it comes to a breakdown of democracy, we compare our index to some of V-Dem's "early warning" indicators of autocratization (Maerz et al., 2020b; Coppedge et al., 2020a). We use V-Dem here since other political regime measures, such as Polity or Freedom House, are highly aggregated constructs. This makes their indices notoriously insensitive for analytically relevant gradual transformations that occur under the hood and only in some aspects that matter for the overall assessment of the nature and future of a political regime. This is often problematic, and in particular so in the context of our research. We do not expect that public rhetoric alone and in a matter of only a few years will change the institutional make-up of a political regime *tout court*. It is much more plausible to expect the effect of illiberal speech in a democracy to have corrosive effects on only parts of what makes up the essence of democracy, as several of V-Dem's disaggregated subindicators show. Those effects on sub-components of a regime might add up over time and eventually lead to changes big enough to be also registered by Freedom House or Polity scores.

As Maerz et al. (2020b) illustrate, harbingers of democratic decline are not, as one might expect, the decay of electoral institutions. Instead, democracy starts dying via declines in freedom of the media, civil liberties, freedom of academic and cultural expression, and other liberal principles. Our expectation therefore is: formally democratic regimes in which heads of government speak with an illiberal tongue are also those regimes in which V-Dem's early warning indicators foresee democratic decline.<sup>24</sup>

Figure 3 ("Government Censorship Effort") and Figure 4 ("Freedom of Academic and Cultural Expression") show precisely this: the speakers we identify as clearly illiberal on our scale (left of the vertical line) are all leaders of regimes that experienced dramatic drops in terms of media freedom and other liberties during the last decade. This is shown by the purple-colored cases and their long grey dashed lines leading South, with Turkey being the case declining most on both of V-Dem's early warning indicators. In contrast, among democracies (green markers) with liberal

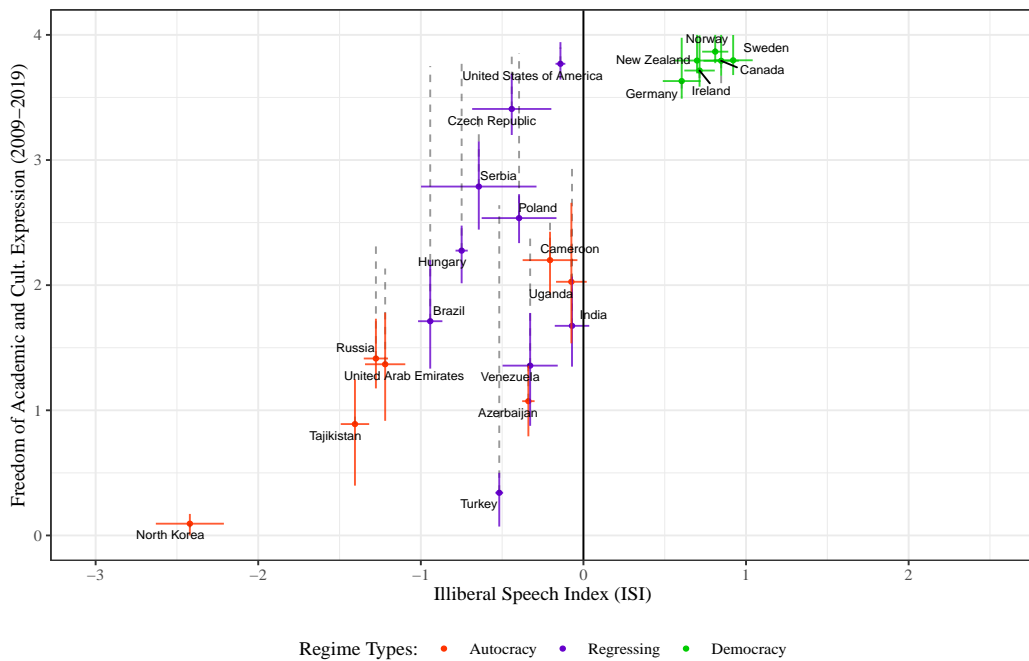
---

<sup>24</sup>For illustrative purposes, we chose the ordinal versions of all V-Dem variables used. The section on cross-validation in the Appendix includes also plots with the V-Dem indicators on "Media Bias" and the repression of civil society organisations ("CSO Repression").



The vertical dashed lines signify the change of the Government Censorship score in 2019 compared to 2009. The horizontal colored lines represent the 95% confidence intervals of the ISI, the vertical colored lines the Bayesian 68% highest posterior density intervals of the Government Censorship score.

Figure 3: Government Censorship Effort and the Illiberal Speech Index (ISI)



The vertical dashed lines signify the change of the Freedom of Academic and Cult. Expression score in 2019 compared to 2009. The horizontal colored lines represent the 95% confidence intervals of the ISI, the vertical colored lines the Bayesian 68% highest posterior density intervals of the Freedom of Academic and Cult. Expression score.

Figure 4: Freedom of Academic and Cult. Expression and the Illiberal Speech Index (ISI)



speakers (right of the vertical line), no decay in liberal institutions is detectable. By identifying the liberal rather than the electoral component as the weak spot of democracy, these comparative illustrations further validate our Illiberal Speech Index and also provide additional support for our basic assumption that what matters most when analyzing the current crisis of democracy is a focus on its liberal essence.

#### 4.2 *Authoritarian public rhetoric: Government disinformation and anti-pluralism*

Following Glasius (2018) and our conceptualization above, we stipulate that there are two analytically distinct yet overlapping dimensions of public rhetoric that are relevant as early warning signs for democratic decay: illiberal and authoritarian rhetoric. The former, we have just analyzed using a machine-based approach to texts. Measuring authoritarian public rhetoric across countries is a methodologically challenging matter, though. Identifying official government disinformation campaigns and anti-pluralist speeches in multiple cases is more difficult than the detection of an illiberal discourse in the form of nationalist, paternalist, and traditionalist language style. Unfortunately, the dictionary analysis we have used for this is too coarse and not suitable for tracking deceitful lies, "fake news", or a subliminal advertising of anti-pluralist values in official speeches. Novel deep learning techniques to automatically detect disinformation are still limited in terms of their real-world implementation (Islam et al., 2020). We therefore rely on human-coded data of authoritarian public rhetoric and apply logistic regression analysis<sup>25</sup> in order to assess the impact of authoritarian public discourse on the autocratization of political regimes.

Our dependent variable in the logistic regression is a binary variable that indicates whether a country is experiencing autocratization (1) or not (0). For the operationalization of *Autocratization*, we rely on the new Episodes of Regime Transformation (ERT) dataset (Edgell et al., 2020b; Maerz et al., 2020a). The ERT detects episodes of autocratization based on a sustained and substantial decline of V-Dem's expert-coded Electoral Democracy Index (EDI).<sup>26</sup> The EDI assesses the degree to which a political regime respects the principles of polyarchy by Dahl (1971). It is based on over forty indicators - including several measurements on liberal components of democracy, such as media freedom and civil liberties.

Our main independent variable - authoritarian public rhetoric - is operationalized with two indicators: government disinformation on social media (*Disinformation*) and *Anti-pluralist rhetoric* of ruling parties. For the operationalization of government disinformation on social media we rely on V-Dem's indicator *v2smgovdom* of the Digital Society Survey (Mechkova et al., 2019). It assesses the extent to which a government or its agents use social media to disseminate misleading viewpoints or false information to influence its own population. For measuring anti-pluralist rhetoric of ruling parties we use the novel V-Party index on Anti-pluralism (Lührmann et al., 2020). It captures the extent to which ruling parties are committing to non-democratic norms already

---

<sup>25</sup>Due to its varying coverage, the Illiberal Speech Index is (not yet) available as country-year measures and therefore not included in this regression analysis.

<sup>26</sup>Sustained and substantial means, for example, a decline of more than 10 percent of the index value over the duration of the episode. For further details on the ERT parameters see Edgell et al. (2020b).

before coming to power. While these expert-coded indicators on government disinformation and party rhetoric can merely serve as proxies of actual analyses of public speeches as we have done above, they are still useful pointers to how authoritarian public rhetoric covaries with regime autocratization.

As control variables we include several aspects identified in the literature as crucial factors of autocratization (e.g. Boese et al., 2020). Military coups can be serious threats to democracy (Marinov and Goemans, 2014). We control for the occurrence of one or several coups per country and year as a binary *Coup* variable based on two coup datasets included in V-Dem (*e\_pt\_coups* and *e\_coups*). Following the theory of Linz (1990) on the perils of presidentialism due to less constraints on the executive and by making use of V-Dem’s fine-graded data, we further control for *Judicial constraints on the executive* (*v2x\_jucon*) and *Legislative constraints on the executive* (*v2xlg\_legcon*) as expected safeguards against autocratization. In addition, we account for diffusion effects (e.g. Brinks and Coppedge, 2006) by including a measurement on average *Regional democracy levels*, constructed with the help of V-Dem’s EDI. We also control for economic factors by including a variable on *GDP growth* from the Maddison Project (Bold and Luiten van Zanden, 2020). Apart from that, the analysis contains a time index (*Timetrend*) to control for time trends which might affect autocratization and a variable on *Internet* use (ITU, 2020). The latter is included as an interacting variable with *Disinformation* since we expect that with increasing Internet use the effects of social media disinformation by regimes might also change.

Based on this data, the analysis covers 1 365 cases of country years from 2001 to 2019. With our main interest in analyzing how authoritarian public rhetoric relates to democratic decline, the analysis includes only democracies and regressing democracies, but not autocracies. All independent and control variables except for *Anti-pluralist rhetoric*,<sup>27</sup> *Coup*, and *Timetrend* are lagged by one year. Due to the binary nature of our dependent variable, we use the R package *bife* to estimate a logistic model with country-fixed effects.<sup>28</sup> As illustrated in the robustness checks in the Appendix, we cross-checked the outcome of the *bife* model with a generalized linear model (*glm*) - showing the same results - transformed some of the variables due to skewed distributions (*Disinformation*, *Regional democracy levels*, and *Internet* have been logged) and made use of *bife*’s post-estimation routine to substantially reduce the incidental parameter bias problem present in non-linear fixed-effects models (Stammann et al., 2020). We also run alternative models based on all political regimes and all but closed autocracies.<sup>29</sup> The Appendix includes all regression tables and further diagnostics, such as heat map (Esarey et al., 2016) and separation plot (Greenhill et al., 2011) to specifically assess the robustness of binary-dependent variable models.<sup>30</sup>

---

<sup>27</sup>The V-Party dataset assesses party identities before elections; we filled the variable for the entire election cycle of a ruling party.

<sup>28</sup>We include the *Timetrend* variable instead of year-fixed effects to not overly restrict the model.

<sup>29</sup>The main difference of these alternative models compared to our main model on autocratization in democracies only is that *Disinformation* is not showing as statistically significant variable, which could be due to the highly frequent use of disinformation in autocracies - independent of regressing or not.

<sup>30</sup>Replication files are available upon request.

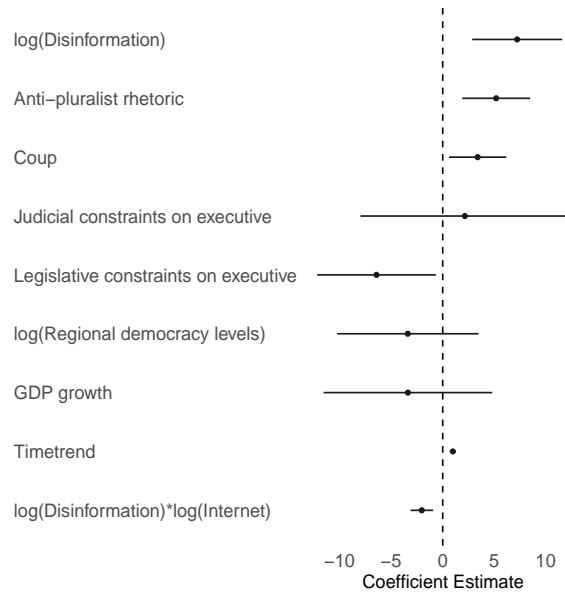


Figure 5: Coefficient estimates and 95% confidence intervals from logistic regression with country-fixed effects predicting autocratization (regression table with full details in Appendix).

Figure 5 summarizes the results of the analysis by illustrating the coefficient estimates and 95% confidence intervals of our binary logistic regression model with country-fixed effects. The coefficients for our two main independent variables - *Disinformation* and *Anti-pluralist rhetoric* - are positively related to autocratization on a statistically significant level. As the regression table (Table 3) in the Appendix shows, this is true for both a bivariate and multivariate model with  $p < 0.01$ . Overall, these findings are in support of our argument that authoritarian public rhetoric is an important factor of autocratization.

Apart from the two independent variables, Figure 5 shows that the occurrence of coups is also positively related to autocratization. *Legislative constraints on the executive* is negatively related to autocratization, indicating that if fully-functioning, such constraints can indeed be one potential safeguard from autocratization. While the other control variables do not show up as statistically significant coefficients in our model, the *Timetrend* variable is slightly positively related to autocratization, suggesting that during the time period observed (2001-2019), autocratization has accelerated. Lastly, the interaction between *Disinformation* and *Internet* is - on a comparatively small level - negatively related to autocratization. This finding confirms our expectation that with increasing Internet use among the population, the effects of official disinformation campaigns on social media might change: as Internet diffusion increases, the effect of government disinformation via social media on autocratization weakens. The population can refer to alternative online sources of information to fact-check and reject officially propagated "fake news".

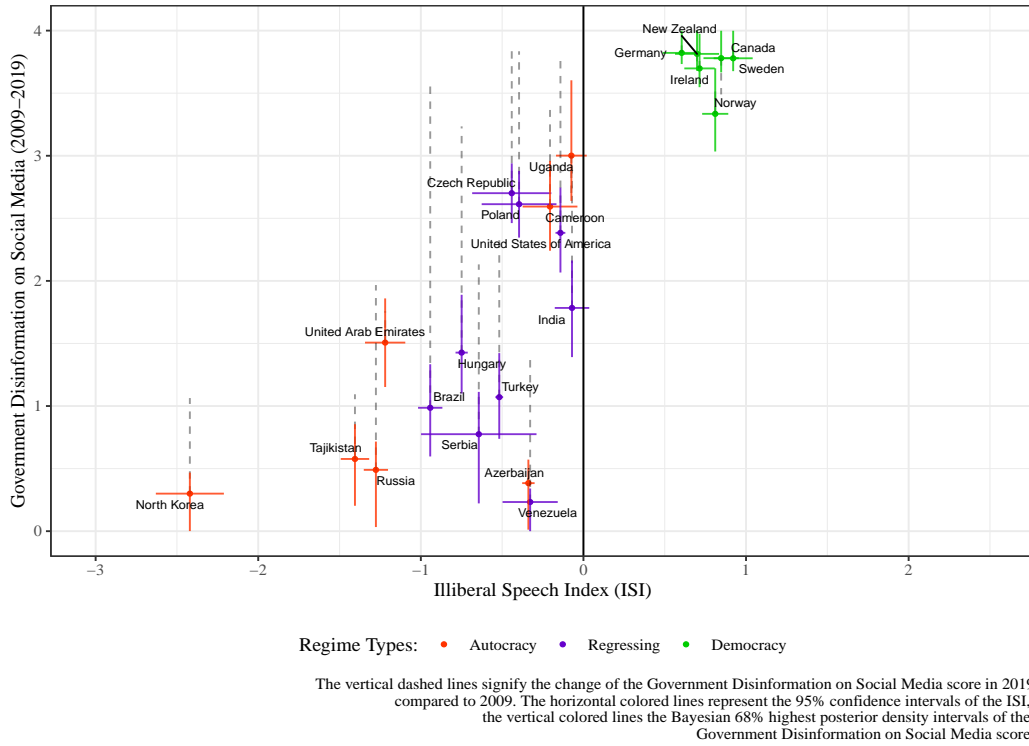


Figure 6: Disinformation and the Illiberal Speech Index (ISI)

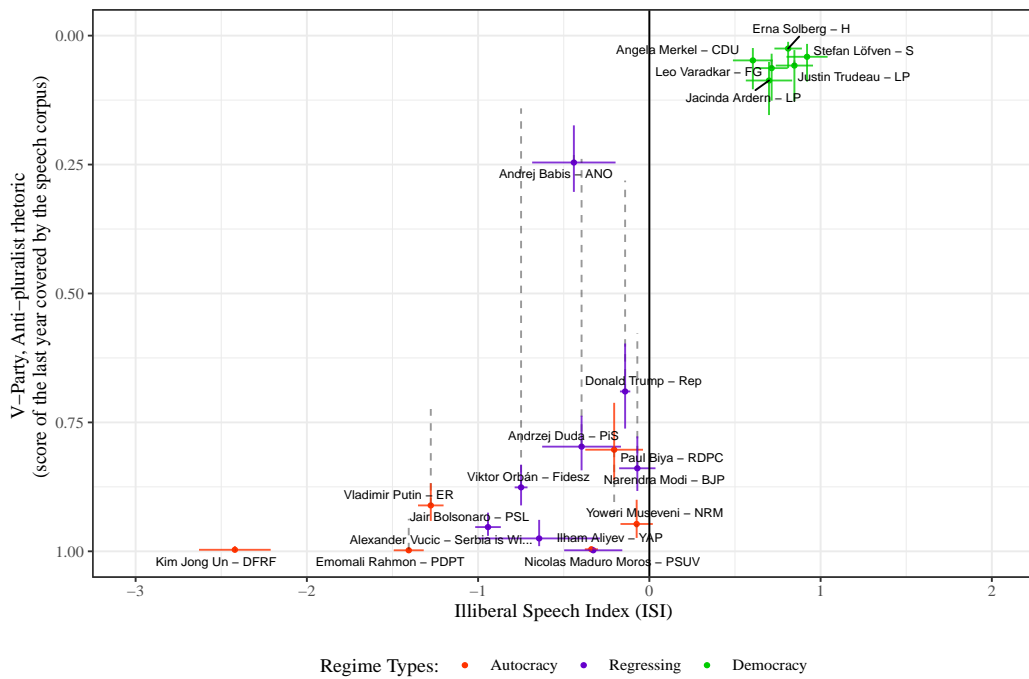


Figure 7: Anti-pluralist rhetoric and the Illiberal Speech Index (ISI)

## 5 The relationship between illiberal and authoritarian public rhetoric

We conceptualized illiberal and authoritarian public rhetoric as two different but strongly overlapping and mutually not exclusive categories. How does that unfold empirically? In Figure 6 and Figure 7, we illustrate the relationship between authoritarian public rhetoric - measured based on the two variables government disinformation on social media (*Disinformation*) and *Anti-pluralist rhetoric* of ruling parties - and the Illiberal Speech Index, in both plots displayed on the x-axis. The y-axis of Figure 6 shows the extent to which a government spreads disinformation on social media on a flipped Likert scale from 0 (extremely often) to 4 (never).<sup>31</sup> The horizontal colored lines and point estimates are the variable scores and confidence intervals for 2019. The vertical grey dashed lines show how much these scores changed compared to 2009. The y-axis of Figure 7 shows the extent to which ruling parties make use of anti-pluralist rhetoric (interval, from low to high, 0-1). Since this V-Party index (Lührmann et al., 2020) is based on election cycles instead of country years, the horizontal colored lines and point estimates are here the scores and confidence intervals of the last year covered by our speech corpus. The vertical grey dashed lines signify the change of the last compared to the first score covered by the speech corpus.

Figure 6 illustrates that for most cases covered in our analyses, autocratization comes along with both an illiberal and authoritarian public rhetoric: all case markers in violet are left of the vertical line and below the *Disinformation* score of 3. While several of the regressing regimes, such as Brazil, Hungary, Poland, Czech Republic, and the US did not exhibit any meaningful extent of official disinformation in 2009, their levels substantially dropped during the last decade. Regarding the US under Trump or Bolsonaro's Brazil, this trend has further exacerbated in 2020 as Edgell et al. (2020a) show with their data on pandemic backsliding and official disinformation campaigns during the Covid-19 pandemic.

Figure 7 points even more explicitly to the strong connection between illiberal and authoritarian public rhetoric: nearly all heads of government in regressing regimes that score as illiberal speakers on the ISI,<sup>32</sup> also score as parties with high levels of anti-pluralist rhetoric.<sup>33</sup> Furthermore, some cases, such as Hungary's Fidesz or Poland's PiS and the Republicans under Trump, show a remarkable downtrend during the time covered by our speech corpus. In sum, both figures further strengthen the face validity of the ISI index and also cross-validate our measures of illiberal and authoritarian public rhetoric in general since stable democracies typically score high and autocracies low on both dimensions. The figures also illustrate that illiberal and authoritarian public rhetoric seem to be two sides of the same coin: worrisome language practices that harm democracy.

---

<sup>31</sup>For illustrative purposes and in contrast to the regression analysis, we use the ordinal version of the V-Dem variable *v2smgovdom* here.

<sup>32</sup>India's confidence interval is crossing the zero point.

<sup>33</sup>While Czech Republic's ruling party, the ANO (Action of Dissatisfied Citizens), scores high on V-Party's Populism index, it is not classified as particularly non-pluralist by the dataset (Lührmann et al., 2020).

## 6 Conclusion

This paper has focused on the rhetoric of political leaders and its role in regime transformation. Long being ignored as a crucial aspect of the survival prospects of liberal democracies, public discourse is slowly finding its way (back) into the analysis of political regimes - a development triggered by the emergence of both modern text-as-data techniques and blatantly obvious examples of democratic decline via a deteriorating public rhetoric of core political figures. Distinguishing two forms of anti-democratic discourses - illiberal and authoritarian - we have shown that both of them are closely linked to democratic decline. In other words, it does not remain without consequences for the political regime as a whole if its main representatives signal via their rhetoric that they have abandoned the core values of that regime.

We see several avenues for further research. One such avenue could treat our speech index as the independent variable for a whole set of phenomena of interest. For instance, does lasting illiberal rhetoric by the head of government in liberal democracies lead to an erosion of citizens' support for liberal values? Or, does such rhetoric affect the level of trust in liberal institutions among citizens? Another avenue forward could treat our speech index as the dependent variable and investigate the causes of why such rhetoric is adopted in some countries and not others. Is this driven by latent normative commitments among citizens? Or, is it triggered by major economic and social crises? A third, related, road ahead could be to unpack the causal mechanism that links anti-democratic rhetoric to democratic decline. Does illiberal discourse shift the realm of the thinkable (and sayable) first among political elites and then among citizens, vice versa, or simultaneously? Is anti-democratic discourse closely tied to the speaker or does it outlast a change in leadership? How do the frequency and duration of anti-democratic discourse influence its effect on citizens' perceptions? These and similar questions are crucial for those whose goal it is to counter-act the consequences of illiberal and authoritarian discourse.

One key aspect of future research on public rhetoric and regime change ought to be the role of the (new) media. The more heterogeneous the messaging and information is to which citizens are exposed, the lower the impact of government-led discourses should be. Inversely, the more governments in backsliding democracies manage to get the media under their control, the more impact their illiberal and authoritarian discourse will have on the survival rate of liberal democracy. If true, this does not bode well for the prospect of democracy in backsliding regimes, for backsliding consists precisely of the step-wise breakdown of democratic institutions and each such step tends to make the next step down the ladder toward autocracy more likely and easier to achieve.

All possible future avenues will benefit from more data. Extending our corpus of speeches both in time and space is definitely one of the next steps. Fortunately, such an extension of this already massive data set comes at relatively low costs. Once more extensive use is made of machine-based translations into English of the original speech texts, no insurmountable obstacle exists to harvesting speeches of all heads of government around the globe who make their speeches available in electronic form, which an overwhelming majority does.

## References

- Aalberg, T., Esser, F., Reinemann, C., Strömbäck, J., and de Vreese, C. H. (2017). *Populist political communication in Europe*. Routledge, New York.
- Benoit, K. (2019). Text as data: An overview. In Curini, L. and Franzese, R., editors, *The SAGE Handbook of Research Methods in Political Science and International Relations*. SAGE Publishing, London.
- Blake, A. (2021). What Trump said before his supporters stormed the Capitol, annotated. <https://www.wsj.com/articles/what-trump-said-to-supporters-on-jan-6-before-their-capitol-riot-11610498173>.
- Boese, V. A., Edgell, A. B., Hellmeier, S., Maerz, S. F., and Lindberg, S. I. (2020). Deterring Dictatorship: Explaining Democratic Resilience since 1900. *V-Dem Working Paper Series*, 101.
- Bold, J. and Luiten van Zanden, J. (2020). The Maddison Project: Maddison style estimates of the evolution of the world economy. A new 2020 update. *Maddison Project Working Paper*, (15).
- Brinks, D. and Coppedge, M. (2006). Diffusion is no illusion: Neighbor emulation in the third wave of democracy. *Comparative Political Studies*, 39(4):463–489.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K. M., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skanning, S.-E., Staton, J. K., Wilson, S., Cornell, A., Alizada, N., Gastaldi, L., Gjerløw, H., Hindle, G., Ilchenko, N., Maxwell, L., Mechkova, V., Medzihorsky, J., von Römer, J., Sundström, A., Tzelgov, E., Wang, Y.-t., Wig, T., and Ziblatt, D. (2020a). V-Dem Dataset V10, Varieties of Democracy (V-Dem) Project. <https://www.v-dem.net/en/data/data-version-10/>.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K. M., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skanning, S.-E., Staton, J. K., Wilson, S., Cornell, A., Alizada, N., Gastaldi, L., Gjerløw, H., Hindle, G., Ilchenko, N., Maxwell, L., Mechkova, V., Medzihorsky, J., von Römer, J., Sundström, A., Tzelgov, E., Wang, Y.-t., Wig, T., and Ziblatt, D. (2020b). V-Dem Codebook V10, Varieties of Democracy (V-Dem) Project. <https://www.v-dem.net/en/data/data-version-10/>.
- Dahl, R. A. (1971). *Polyarchy: participation and opposition*. Yale University Press, New Haven.
- Dawson, J. (2018). 'Everyday Democracy': an ethnographic methodology for the evaluation of (de-) democratisation. *East European Politics*, 34(3):297–316.
- de Vries, E., Schoonvelde, M., and Schumacher, G. (2018). No Longer Lost in Translation. Evidence that Google Translate Works for Comparative Bag-of-Words Text Applications, Political Analysis, Online First.
- Dukalskis, A. (2017). *The Authoritarian Public Sphere - Legitimation and Autocratic Power in North Korea, Burma, and China*. Routledge, London, New York.
- Edgell, A. B., Lührmann, A., Maerz, S. F., Lachapelle, J., Grahn, S., Alijla, A., Boese, V. A., Fernandes, T., Gafuri, A., Hirndorf, D., Howell, C., Ilchenko, N., Yuko Kasuya, Medzihorsky, J.,

- Khawaja, A. S., Shenga, C., Tiulegenov, M., Tung, H. H., Wilson, M. C., and Lindberg, S. I. (2020a). *Pandemic Backsliding: Democracy During Covid-19 (PanDem)*, Version 5. Varieties of Democracy (V-Dem) Institute.
- Edgell, A. B., Maerz, S. F., Maxwell, L., Morgan, R., Medzihorsky, J., Wilson, M. C., Boese, V. A., Hellmeier, S., Lachapelle, J., Lindenfors, P., Lührmann, A., and Lindberg, S. I. (2020b). *Episodes of regime transformation dataset and codebook, v2.1*. <https://github.com/vdemoinstitute/ERT>.
- Esarey, J., Pierce, A., and Esarey, J. (2016). *Society for Political Methodology Assessing Fit Quality and Testing for Misspecification in Binary-Dependent Variable Models* Published by : Oxford University Press on behalf of the Society for Political Methodology. 20(4):480–500.
- Glasius, M. (2018). What authoritarianism is and is not: A practice perspective. *International Affairs*, 94(3):515–533.
- Greenhill, B., Ward, M. D., and Sacks, A. (2011). The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models. *American Journal of Political Science*, 55(4):991–1002.
- Grimmer, J. and Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 21(03):267–297.
- Habermas, J. (1984). *The theory of communicative action*. Beacon Press.
- Hawkins, K. A. and Kaltwasser, C. R. (2017). The ideational approach to populism. *Latin American Research Review*, 52(4):513–528.
- Hutchby, I. and Wooffitt, R. (2008). *Conversation Analysis*. Polity Press, Cambridge.
- Islam, M. R., Liu, S., Wang, X., and Xu, G. (2020). Deep learning for misinformation detection on online social networks: a survey and new perspectives. *Social Network Analysis and Mining*, 10(1):1–20.
- ITU (2020). Percentage of individuals using the Internet. <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>.
- Kirsch, H. and Welzel, C. (2018). *Democracy Misunderstood: Authoritarian Notions of Democracy around the Globe*. *Social Forces, Online First*.
- Kreuz, R. J. and Windsor, L. C. (2021). How Trump’s language shifted in the weeks leading up to the Capitol riot – 2 linguists explain. <https://theconversation.com/how-trumps-language-shifted-in-the-weeks-leading-up-to-the-capitol-riot-2-linguists-explain-152483>.
- Kruse, M. and Weiland, N. (2016). Donald Trump’s Greatest Self-Contradictions. <https://www.politico.com/magazine/story/2016/05/donald-trump-2016-contradictions-213869>.
- Levitsky, S. and Ziblatt, D. (2018). *How Democracies Die - What History Reveals About our Future*. Penguin Random House, UK.
- Linz, J. J. (1978). *The Breakdown of Democratic Regimes: Crisis, Breakdown and Reequilibration*. Johns Hopkins University Press.
- Linz, J. J. (1990). The perils of presidentialism. *Journal of Democracy*, 1(1):51–69.



- Lowe, W., Benoit, K., Mikhaylov, S., and Laver, M. (2011). Scaling policy preferences from coded political texts. *Legislative Studies Quarterly*, 36:123–155.
- Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., and Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, 23:254–277.
- Lührmann, A. and Lindberg, S. I. (2019). A third wave of autocratization is here: what is new about it? *Democratization*, 26(7).
- Lührmann, A., Marquardt, K. L., and Mechkova, V. (2020). Constraining Governments: New Indices of Vertical, Horizontal, and Diagonal Accountability. *American Political Science Review*, 114(3):811–820.
- Lührmann, A., Düpont, N., Higashijima, M., Kavasoglu, Y. B., Marquardt, K. L., Bernhard, M., Döring, H., Hicken, A., Laebens, M., Lindberg, S. I., Medzihorsky, J., Neundorf, A., Reuter, O. J., Ruth-Lovell, S., Weghorst, K. R., Wiesehomeier, N., Wright, J., Alizada, N., Bederke, P., Gastaldi, L., Grahn, S., Hindle, G., Ilchenko, N., von Römer, J., Wilson, S., Pemstein, D., and Seim, B. (2020). Varieties of Party Identity and Organization (V-Party) Dataset V1. Varieties of Democracy (V-Dem) Project .
- Maerz, S. F. (2016). The electronic face of authoritarianism: E-government as a tool for gaining legitimacy in competitive and non-competitive regimes. *Government Information Quarterly*, 33(4):727–735.
- Maerz, S. F. (2019). Simulating Pluralism: The Language of Democracy in Hegemonic Authoritarianism. *Political Research Exchange*, 1(1).
- Maerz, S. F., Edgell, A. B., Krusell, J., Maxwell, L., and Hellmeier, S. (2020a). ERT - An R package to load, explore, and work with the Episodes of Regime Transformation (ERT) dataset. <https://github.com/vdeminstitute/ERT>.
- Maerz, S. F., Lührmann, A., Hellmeier, S., Grahn, S., and Lindberg, S. I. (2020b). State of the world 2019: Autocratization Surges - Resistance Grows. *Democratization*, 27(6):909–927.
- Maerz, S. F. and Puschmann, C. (2020). Text as Data for Conflict Research : A Literature Survey. In Deutschmann E., Lorenz J., Nardin L., Natalini D., W. A., editor, *Computational Conflict Research. Computational Social Sciences*, pages 43–65. Springer, Cham.
- Maerz, S. F. and Schneider, C. Q. (2019). Comparing Public Communication in Democracies and Autocracies - Automated Text Analyses of Speeches by Heads of Government -. *Quality and Quantity, Online First*.
- Marinov, N. and Goemans, H. (2014). Coups and democracy. *British Journal of Political Science*, 44(4):799–825.
- Maynard, J. L. (2013). A map of the field of ideological analysis. *Journal of Political Ideologies*, 18(3):299–327.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Studium Paedagogik. Beltz, Weinheim.
- McCoy, J., Rahman, T., and Somer, M. (2018). Polarization and the global crisis of democracy: common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, 62(1):16–42.

- Mechkova, V., Pemstein, D., Seim, B., and Wilson, S. (2019). Introducing the Digital Society Project. *DSP Working Paper*, (1).
- Merkel, W. (2014). Is There a Crisis of Democracy? *Democratic Theory*, 1(2):11–25.
- Merkel, W. and Kneip, S. (2018). *Democracy and Crisis*. Springer International Publishing.
- Monroe, B. L. and Schrodtt, P. A. (2008). Introduction to the special issue: The statistical analysis of political text. *Political Analysis*, 16(4):351–355.
- Mudde, C. (2020). Trump is pushing nationalist myths. But Democrats indulge lavish patriotism, too. <https://www.theguardian.com/commentisfree/2020/sep/20/donald-trump-nationalist-myths-democrats-patriotism>.
- O'Donnell, G. A. (1996). Illusions and Conceptual Flaws. *Journal of Democracy*, 7(4):160–168.
- O'Donnell, G. A. (2007a). The Perpetual Crises of Democracy. *Journal of Democracy*, 18(1):5–11.
- O'Donnell, G. O. (2007b). Notes on Various Accountabilities and Their Interrelations. In *Dissonances. Democratic Critiques of Democracy*, pages 99–109. University of Notre Dame Press, Notre Dame, IN.
- Paz, C. (2020). All the President's Lies About the Coronavirus. <https://www.theatlantic.com/politics/archive/2020/11/trumps-lies-about-coronavirus/608647/>.
- Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y.-t., Medzihorsky, J., Krusell, J., Miri, F., and von Römer, J. (2020). The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data. *V-Dem Working Paper*, (21).
- Schedler, A. (2001). Measuring Democratic Consolidation. *Studies in Comparative International Development*, 36(1):66–92.
- Schedler, A. (2019). The breaching experiment: Donald Trump and the normative foundations of democracy. *Zeitschrift für Vergleichende Politikwissenschaft*, pages 433–460.
- Schmidt, V. A. (2008). Discursive Institutionalism: The Explanatory Power of Ideas and Discourse. *Annual Review of Political Science*, 11(1):303–326.
- Schmitter, P. C. (2015). Crisis and transition, but not decline. *Journal of Democracy*, 26(1):32–44.
- Schmitter, P. C. and Guilhot, N. (2000). From transition to consolidation: Extending the concept of democratization and the practice of democracy. In Dobry, M., editor, *Democratic and Capitalist Transitions in Eastern Europe*, pages 131–146. Springer, Dordrecht.
- Stammann, A., Czarnowska, D., Heiss, F., and McFadden, D. (2020). The R package bife. <https://cran.r-project.org/web/packages/bife/bife.pdf>.
- Steenbergen, M. R., Bächtiger, A., Spörndli, M., and Steiner, J. (2003). Measuring Political Deliberation: A Discourse Quality Index. *Comparative European Politics*, 1(1):21–48.
- The Guardian (2020). Seven of Donald Trump's most misleading coronavirus claims. <https://www.theguardian.com/us-news/2020/mar/28/trump-coronavirus-misleading-claims>.

- The Washington Post (2020). President Trump has made more than 20,000 false or misleading claims. <https://www.washingtonpost.com/politics/2020/07/13/president-trump-has-made-more-than-20000-false-or-misleading-claims/>.
- Timm, J. C. (2020a). 'It's irresponsible and it's dangerous': Experts rip Trump's idea of injecting disinfectant to treat COVID-19. <https://www.nbcnews.com/politics/donald-trump/it-s-irresponsible-it-s-dangerous-experts-rip-trump-s-n1191246>.
- Timm, J. C. (2020b). Trump versus the truth: The most outrageous falsehoods of his presidency. <https://www.nbcnews.com/politics/donald-trump/trump-versus-truth-most-outrageous-falsehoods-his-presidency-n1252580>.
- Wodak, R. (1989). *Language, Power and Ideology: Studies in Political Discourse*. John Benjamins Publishing.
- Wodak, R. (2001). The Discourse-Historical Approach. In Wodak, R. and Meyer, M., editors, *Methods of Critical Discourse Analysis*, pages 63–94. Sage, London.
- Wodak, R. (2009). *The Discourse of Politics in Action. Politics as Usual*, volume 53. Palgrave Macmillan.
- Wodak, R. and Meyer, M. (2009). *The Discourse of Politics in Action. Politics as Usual*. Sage.

## Appendix

### Quotes by Viktor Orbán

"[...] The truth is that Europe is being threatened by mass migration on an unprecedented scale. Tens of millions of people could come to Europe. Today we are talking about hundreds of thousands, but next year we will be talking about millions, and this will never end. There is unlimited supply for this mass migration, because only some of those who are coming are actually from war zones. [...] And one morning we could wake up and realise that we are in the minority on our own continent. [...] they (*the migrants*) could occupy Hungary – something not unprecedented in our history – or they could introduce communism."<sup>34</sup>

" Why does the Hungarian government support the fence? It supports the fence because it works. It works! [...] A winter break. I beseech you not to fall for illusions: there will be no winter break. There will be no winter break in refugee affairs, illegal migration and the global crisis. [...] We may expect quite the opposite: rather than a break, [...] we should prepare for an increased flood and mass of people. The real problem is that there are some in Europe – I do not know if they are in the majority, but it is certain that leaders who say things like this are in the majority, which is of course not the same as the people – who believe that what is happening is a good thing, and that this is a great opportunity for their countries to change. [...] They are not worried by the experience that we – the European Christian cultural community – have so far been unable to integrate them (*Muslim communities*), and that therefore parallel societies are coming into being in a number of European countries, with declining Christian and increasing Muslim ratios."<sup>35</sup>

"[...] I could also say that they (*in Brussels*) have opened the door to George Soros's plan, and in the period ahead I expect to see the acceleration of that plan's implementation. [...] So we're not an immigrant country, and Hungary doesn't want to become an immigrant country; at least I as Prime Minister have been authorised by the Hungarian people to protect this country's economic, cultural and spiritual identity, its traditions and interests. [...] I will never allow any great power – any top dog, any former colonial power or their Brussels supporters beholden to them – to turn us into an immigrant country. [...] Therefore we must ensure that they withdraw this decision on the mandatory quota system: they must change it, and we must prevent this one-off decision being turned into a permanent migrant distribution mechanism. In other words we must foil the plan of George Soros's people in Brussels. Now there will be a political struggle, which may take a number of forms and in which we shall fight with all our strength."<sup>36</sup>

---

<sup>34</sup>Interview with Viktor Orbán on September 4, 2015: <https://2015-2019.kormany.hu/en/the-prime-minister/the-prime-minister-s-speeches/if-we-do-not-protect-our-borders-tens-of-millions-of-migrants-will-come>

<sup>35</sup>Prime Minister Viktor Orbán's reply in Parliament on September 25, 2015: <https://2015-2019.kormany.hu/en/the-prime-minister/the-prime-minister-s-speeches/prime-minister-viktor-orban-s-reply-in-parliament>

<sup>36</sup>Prime Minister Viktor Orbán on Kossuth Radio, September 8, 2017: <https://2015-2019.kormany.hu/en/the-prime-minister/the-prime-minister-s-speeches/prime-minister-viktor-orban-on-kossuth-radio-s-180-minutes-programme-20170908>

## Dictionary Terms

- **illiberalism vs liberalism**

- **illiberalism**

- **nationalism, paternalism**

- allah
- almighty
- anarch\*
- chaos
- christ
- christianity
- christians
- church
- danger\*
- destabili\*
- evil
- father\*
- god
- hero\*
- homeland
- illegal\*
- immigra\*
- invincible
- islam
- jesus
- migrant\*
- migration\*
- militar\*
- minaret
- motherland
- muslim\*
- negative
- patriot\*
- police
- pride
- proud
- rebell\*
- religi\*
- riot\*
- separatist\*
- soldier\*
- spiritual
- stabil\*
- subversive
- territory
- uniqueness
- unity
- unlawful

- **traditionalism**

- ancestors
- brothers
- discipline
- family\*
- forefather
- glories
- glorious
- heritage
- honor
- honour
- inherit\*
- loyalty
- majesty
- monarch\*
- moral\*
- obscene\*
- pervert\*
- pornograph\*
- recapture\*
- reliab\*
- shameful
- tradition\*

- **liberalism**

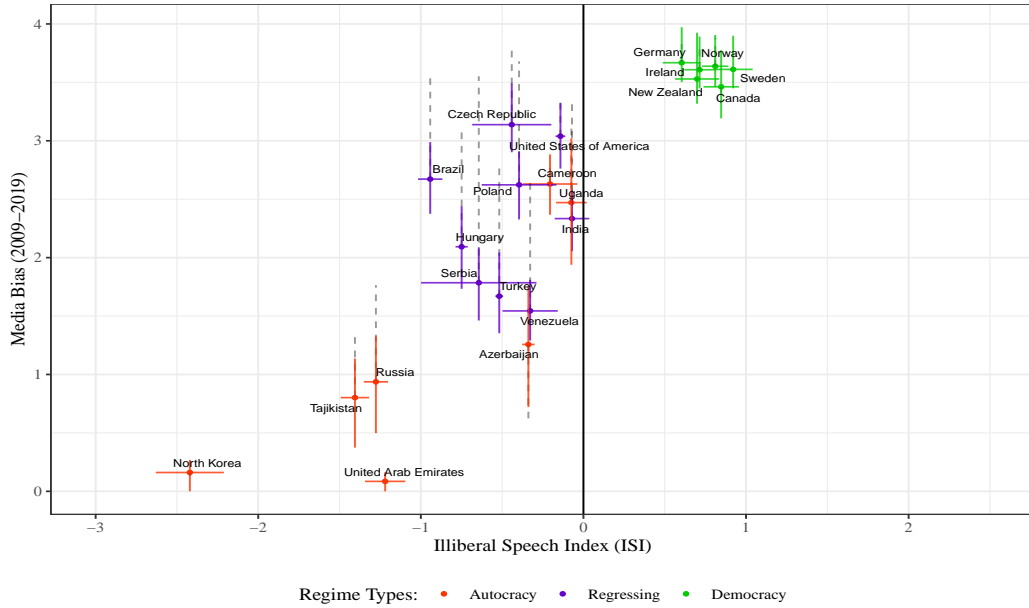
- **liberal values**

- authoritarian\*
- autocra\*
- corrupt\*
- cruel\*
- demilitarization
- dictator\*
- disarmament
- discriminat\*
- diverse
- diversity
- equal\*
- fair\*
- fascism
- free
- freedom
- harassment
- inclusion
- inclusiv\*
- innocent
- interfaith
- interreligious
- liberal\*
- mediat\*
- multicult\*
- negotiation\*
- oppression
- pluralis\*
- referendum
- repressi\*
- suppression
- tolerance
- tolerat\*
- transparen\*
- violat\*
- voluntary
- vote\*
- voting

- **woman, minorities**

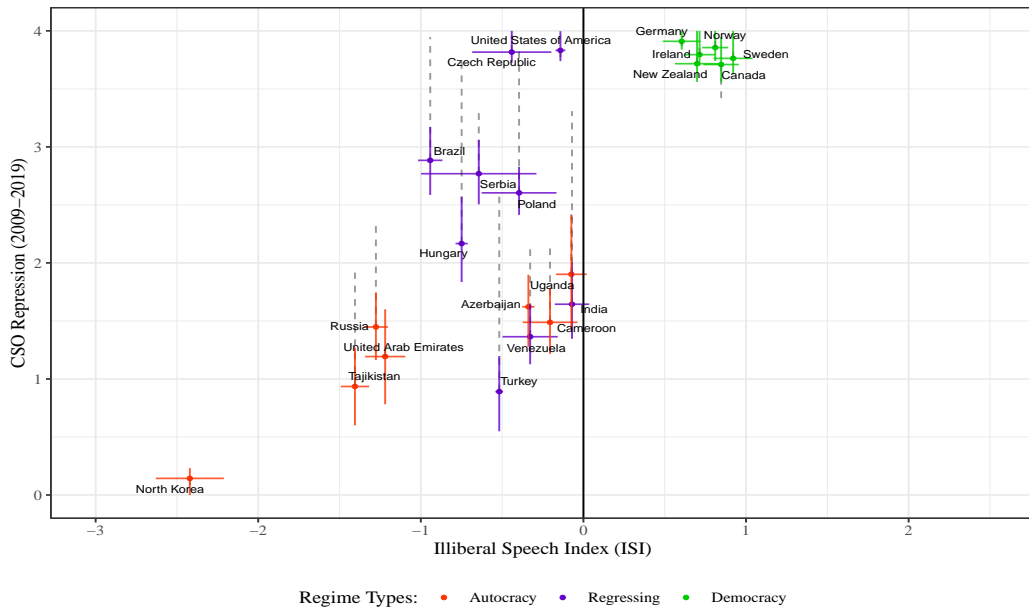
- dialect\*
- ethni\*
- gay\*
- gender
- genocide
- girls
- handicapped
- harmon\*
- indigenous
- injustice\*
- intolerance
- justice
- lesbian\*
- lgbt
- marginalize\*
- minorities
- multiethnic
- queer\*
- racist\*
- rights
- sisters
- solidarity
- transgender
- unfair\*
- voice\*
- woman
- women

Cross-validating the Illiberal Speech Index (ISI) with V-Dem's "early warning" indicators



The vertical dashed lines signify the change of the Media Bias score in 2019 compared to 2009. The horizontal colored lines represent the 95% confidence intervals of the ISI, the vertical colored lines the Bayesian 68% highest posterior density intervals of the Media Bias score.

Figure 8: Media Bias and the Illiberal Speech Index (ISI)



The vertical dashed lines signify the change of the CSO Repression score in 2019 compared to 2009. The horizontal colored lines represent the 95% confidence intervals of the ISI, the vertical colored lines the Bayesian 68% highest posterior density intervals of the CSO Repression score.

Figure 9: CSO Repression and the Illiberal Speech Index (ISI)

Regression analysis, robustness checks

	glm	bife1	bife2 <sup>a</sup>	bife3 <sup>b</sup>
Disinformation	24.07	24.07	9.08	7.23
Anti-pluralist rhetoric	7.87	7.87	7.10	5.19
Coup	2.56	2.56	4.20	3.39
Judicial constraints on executive	10.91	10.91	2.91	2.15
Legislative constraints on executive	-11.88	-11.88	-7.45	-6.43
Regional democracy levels	-17.25	-17.25	-4.04	-3.39
GDP growth	-5.37	-5.37	-4.23	-3.38
Timetrend	0.49	0.49	1.20	0.98
Disinformation*Internet	-0.59	-0.59	-2.51	-2.03

<sup>a</sup> Transformed variables (logged Disinformation, Internet, Regional democracy levels).

<sup>b</sup> Transformed variables (logged Disinformation, Internet, Regional democracy levels) and corrected.

Table 2: Comparing coefficients of Model 2 with glm and bife

	Autocratization (in democracies)	
	Model 1	Model 2
log(Disinformation)	2.54*** (0.62)	7.23** (2.23)
Anti-pluralist rhetoric	4.35*** (0.95)	5.19** (1.68)
Coup		3.39* (1.42)
Judicial constraints on executive		2.15 (5.17)
Legislative constraints on executive		-6.43* (2.94)
log(Regional democracy levels)		-3.39 (3.50)
GDP growth		-3.38 (4.18)
Timetrend		0.98*** (0.15)
log(Disinformation)*log(Internet)		-2.03*** (0.56)
Log Likelihood	-194.47	-101.14
Deviance	388.94	202.28
Num. obs.	408	408

\*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$

Table 3: Model outcomes for autocratization in democracies (main models)

	Autocratization (all regimes)	
	Model 1	Model 2
log(Disinformation)	1.45*** (0.43)	-0.59 (0.73)
Anti-pluralist rhetoric	3.26*** (0.67)	3.19*** (0.86)
Coup		2.79*** (0.65)
Judicial constraints on executive		3.31* (1.67)
Legislative constraints on executive		-2.73* (1.32)
log(Regional democracy levels)		-0.75 (1.75)
GDP growth		-1.95 (1.46)
Timetrend		0.47*** (0.06)
log(Disinformation)*log(Internet)		-0.04 (0.18)
Log Likelihood	-374.53	-308.41
Deviance	749.06	616.82
Num. obs.	719	719

\*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$

Table 4: Model outcomes for autocratization in all regimes

	Autocratization (all but closed regimes)	
	Model 1	Model 2
log(Disinformation)	1.44*** (0.44)	0.27 (0.81)
Anti-pluralist rhetoric	3.20*** (0.68)	2.68** (0.87)
Coup		3.53*** (0.77)
Judicial constraints on executive		4.28* (1.85)
Legislative constraints on executive		-5.42** (1.78)
log(Regional democracy levels)		-1.40 (1.78)
GDP growth		-0.11 (1.26)
Timetrend		0.49*** (0.06)
log(Disinformation)*log(Internet)		-0.26 (0.20)
Log Likelihood	-357.05	-288.70
Deviance	714.11	577.40
Num. obs.	683	683

\*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$

Table 5: Model outcomes for all but closed regimes



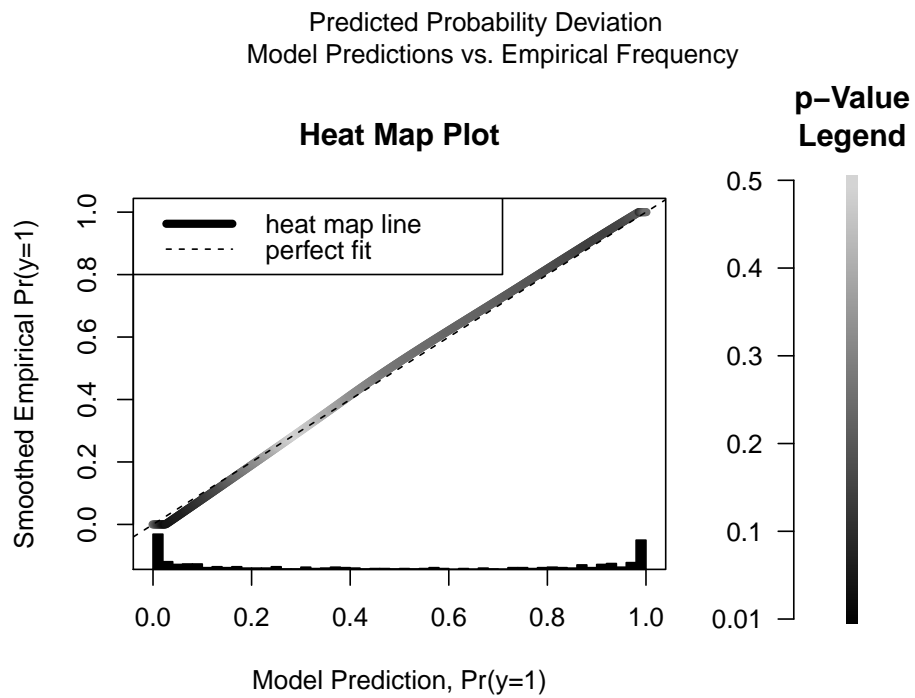


Figure 10: Heat map plot to evaluate the fit of our main binary model, cf. Esarey et al., (2016).

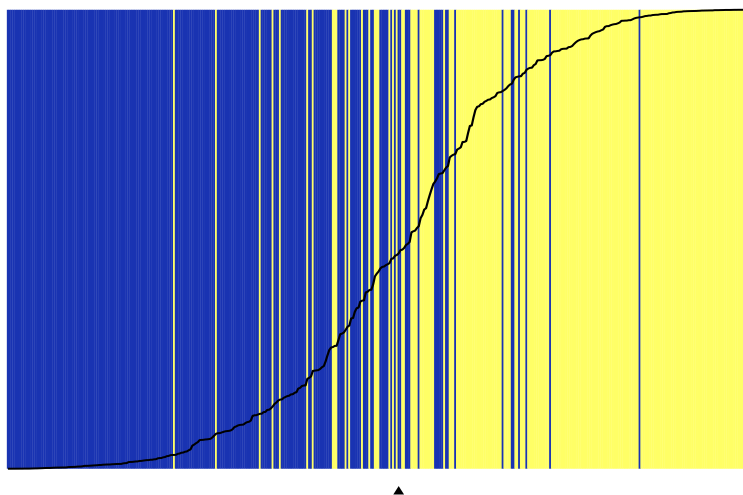


Figure 11: Separation plot to evaluate the fit of our main binary model, cf. Greenhill et al. (2011).