



UNIVERSITY OF
GOTHENBURG

HOW PEOPLE AROUND THE WORLD ASSESS DEMOCRACY COMMUNICATING ONLINE

SÖREN HOLMBERG

WORKING PAPER SERIES 2020:5

QOG THE QUALITY OF GOVERNMENT INSTITUTE
Department of Political Science
University of Gothenburg
Box 711, SE 405 30 GÖTEBORG
August 2020
ISSN 1653-8919
© 2020 by Sören Holmberg. All rights reserved.

How People Around the World assess Democracy Communicating Online
Sören Holmberg
QoG Working Paper Series 2020:5
August 2020
ISSN 1653-8919

Sören Holmberg
The Quality of Government Institute
Department of Political Science
University of Gothenburg
Soren.holmberg@pol.gu.se

The idea is simple. Why not listen in when people talk about democracy and use what is said to analyze support and eventual problems in different democratic systems. If we take advantage of available online communications it could be an unobtrusive data collection method. No surveys demanding people to answer questions are needed. Instead, online communicative behavior is anonymously tapped for information. Two obvious drawbacks are that the resultant data in some case might be less specific compared to data from surveys, and, secondly, that the online method is – for practical and ethical reasons – not useful in collecting individual level data. It is best suited for securing aggregate data on for example regional or national levels.

On the following pages, we will test to what extent data from communicative behavior on the internet can be useful in studying public support for and critique of democracy across different nations of the world. The analysis is based on online text data from the research project Linguistic Explorations of Societies (LES) at the University of Gothenburg. LES has systematically collected language data from online editorial as well as social media from some 140 language-country combinations across the world.

By using recent advances in natural language processing, LES has built a distributional semantic lexicon, which returns semantically similar terms to a given target term from a selected language, country and media type. The lexicon is based on a cumulative data sample from a vast index of online text sources collected between early 2015 and late 2017. More details on the methodology and data can be found in Dahlberg et al. (2020).

For our test, we have selected 70 nations including old established democracies as well as new emerging democracies and authoritarian regimes. More specifically, we have collected 30 semantically similar terms to the target term democracy from each nation, 15 from editorial media and 15 from social media. For some nations, the language analyzed did not differ. For example, English is the chosen language in 17 nations and Spanish in 12. Overall, the unique number of different languages among our selected 70 nations is 34 – among them Chinese, Farsi, Swahili and Afrikaans.

Two Tests and Three Hypotheses

Analyzing democracy, it is too much of a simplification to distinguish – as was commonly done in the early and mid 1900s – between democracy and dictatorship. Political reality is more complex and need more nuanced concepts than a crude dichotomy like this. Democracy is a multidimensional phenomenon best conceptualized and measured on continuous scales. We can talk about different kinds of democracies and of different degrees of democracy.

Democratic decision-making has been portrayed as a triangular drama (Barrling & Holmberg 2018). Three different kind of values have to be balanced and weighted against each other in order for a democratic system to be sustainable and legitimate in the long term. These three critical values are a) majoritarian rule by the people, b) rule of law including minority rights, and c) system capacity and efficiency to act. All three values are essential but could easily end up in conflict with each other – this is the potential triangular drama of democracy. For example, the rule of law and minority rights can hinder majority rule as well as system effectiveness. And the will of the people is not always

congruent with what system action demands. In today's – mainly American – discussion, the multiple facets of democracy are acknowledged, and it has become commonplace to talk about liberal democracy, where liberal refers to individual freedoms and minority rights, while democracy denotes elections and majority rule (Mounk 2018).

Our first test is to see to what extent people talk about things related to these three values when they talk about democracy online. For each value, we have singled out some 15 to 20 specific words that indicate that people have thought of our addressed issues related to the rule of the people, the rule of law or system capacity. If any of these target concepts are present among the 15 most frequently occurring words for a given language-country combination, it is registered as a hit on the relevant variable – rule of the people, rule of law, or system capacity. Multiple hits are possible with a score of 15 being the theoretical maximum.

Two hypotheses will be tried in this context. The first states that people in new emerging democracies will be more occupied with phenomena related to system of rule and rule of the people compared to citizens in more established democracies. The second hypothesis asserts that people in older democracies tend to be more engaged by problems dealing with the rule of law and individual liberties than people in new democracies or in authoritarian regimes. The theoretical underpinnings of both hypotheses have to do with institutional learning and agenda setting in different democratic systems (Rohrschneider 2011). System of rule, which is focused on elections, constitutions, parliament, majority rule and the like could be seen as more relevant and urgent in new democracies, whereas issues dealing with individual freedoms and minority safe-guards become more important as a democracy grows older and more established.

Like in sports, a game is best played if all participants accept and agree to the rules. Transcribed to democratic decision-making, it is a normative and desirable ideal that system arrangements (constitutions) and procedural rules should be uncontroversial and non-politicized. Democratic design can and do differ but should in all cases and forms be seen as neutral – a kind of level playing field.

An indirect way of studying the extent to which democracy is discussed and assessed in non-partisan terms in different political systems would be to look for how often ideological and conflict dimensional concepts are used when people talk about democracy online. We hope that the results would reflect few ideological and politically conflicted words as it would suggest a more normatively desirable state of affairs.

Our second test deals with this normative question. We have chosen three ideological contested issue areas and looked for words or concepts related to them when citizens in our 70 nations address the topic of democracy. The selected conflict dimensions are left-right, religion and nation/nationalism. Our hypotheses are fairly straightforward. Ideological concepts overall will be less frequently used when talking about democracy in established democracies compared to in emerging democracies. Institutional learning should over time lead to less conflicted assessments of democratic rule in order and more settled democracies.

Editorial Versus Social Media

Across most language-country combinations, talking about democracy is very similar between editorial and social media. For our six analytical variables, the correlations between frequencies of relevant word mentions for editorial and social media are always on the semi-high positive end varying between .xx and .33. This means that most often, editorial media is a reasonable mirror or representative of what is talked about on social media.

The results in Table 1 clearly demonstrate how similar talk about democracy online is between editorial and social media. Mean frequencies of relevant words for all the six variables differ only very little between the two.

TABLE 1, MENTIONS OF DEMOCRACY-RELATED WORDS IN ONLINE SOCIAL AND EDITORIAL MEDIA ACROSS 70 LANGUAGE-COUNTRY COMBINATIONS

Variable	Social media		Editorial media		Mean difference
	Number of relevant words	Country mean	Number of relevant words	Country mean	
System of rule	155	2.2	148	2.1	0.1
Rule of law	53	0.8	67	1.0	-0.2
System capacity	53	0.8	48	0.7	0.1
Left-right ideology	119	1.7	98	1.4	0.3
Religion	62	0.9	52	0.7	0.2
Nation/nationalism	89	1.3	79	1.1	0.2

Note: LES data for 70 language-country combinations. See appendix for relevant words.

Interestingly, relevant word mentions are somewhat more common in social media for five of the six analysis variables – most notably for the left-right ideological variable. Talking about democracy, people on social media tend to refer to ideological term like socialism, liberalism, communism or conservatism somewhat more often than is the case on editorial media. The only variable of our six showing more intense editorial media discussion compared to social media is the rule of law variable. However, please note that the difference is miniscule.

Differences in Talking About System of Rule and Rule of Law in Established and Emerging Democracies

The close similarity between the relevant word mentions in editorial and social media makes it feasible to combine them in the analysis. Doing so does not affect the results and simplifies our hypotheses testing. We do not have to test twice, separately for editorial and social media.

The results in Table 2 for combined editorial and social media data contain statistics relevant for testing hypotheses number one and two. In both cases, the outcome lends some support to the hypotheses – a weak support but in the expected direction.

TABLE 2, MENTIONS OF DEMOCRACY-RELATED (CORE VALUE) WORDS IN ONLINE TEXT DATA IN ESTABLISHED DEMOCRACIES, EMERGING DEMOCRACIES, AND AUTHORITARIAN REGIMES

Variable	Relevant word mentions (percent)			Sum (percent)	Mean (number)
	0-1 mentions	2-4 mentions	5+ mentions		
System of rule					
Established democracies	7	55	38	100	4.2
Emerging democracies	9	43	48	100	5.0
Authoritarian regimes	17	50	33	100	3.6
Rule of law					
Established democracies	59	24	17	100	2.3
Emerging democracies	78	18	4	100	1.1
Authoritarian regimes	56	33	11	100	1.7
System capacity					
Established democracies	55	41	4	100	1.6
Emerging democracies	61	35	4	100	1.5
Authoritarian regimes	61	39	0	100	1.1

Note: LES data for 70 language-country combinations. 29 established democracies, 23 emerging democracies and 18 authoritarian regimes. See appendix for relevant words and for the regime classification.

People in emerging democracies talk more often about system of rule than people in established democracies (hypothesis 1). When it comes to talking about the rule of law and minority rights, the result is the opposite; more discussion in established than in emerging democracies (hypothesis 2).

Examples of emerging democracies where people most frequently talk about system of rule are Bulgaria, Mexico, Nigeria and Colombia. Similarly, examples of established democracies where people most often discuss issues related to the rule of law are Germany, Austria, Switzerland, Denmark and Sweden.

Assessing Democracy Through an Ideological Lens

Based on the theory of institutional learning, our expectation was that online talk about democracy would be less ideologically controversial in older democracies than in newer emerging democracies. Looking at the rest results in Table 3, it is very evident such a hypothesis is unequivocally *not* supported by our LES data. Contrary to the hypothesis, people in established democracies refer most often to left-right ideological concepts and religious phenomena when talking about democracy.

TABLE 3, MENTIONS OF DEMOCRACY-RELATED (IDEOLOGICALLY CONTESTED) WORDS IN ONLINE TEXT DATA IN ESTABLISHED DEMOCRACIES, EMERGING DEMOCRACIES, AND AUTHORITARIAN REGIMES

Variable	Relevant word mentions (percent)			Sum (percent)	Mean (number)
	0-1 mentions	2-4 mentions	5+ mentions		
Left-right ideology					
Established democracies	17	38	45	100	4.4
Emerging democracies	48	35	17	100	2.3
Authoritarian regimes	61	22	17	100	2.1
Religion					
Established democracies	59	41	0	100	1.6
Emerging democracies	74	22	4	100	1.1
Authoritarian regimes	50	44	6	100	1.7
Nation/nationalism					
Established democracies	41	45	14	100	2.3
Emerging democracies	22	65	13	100	2.7
Authoritarian regimes	39	61	0	100	2.1

Note: LES data for 70 language-country combinations. 29 established democracies, 23 emerging democracies and 18 authoritarian regimes. See appendix for relevant words and for the regime classification.

Some support for the hypothesis can only be found if we focus on how people online discuss democracy in relation to nationalism, globalism and multiculturalism. Here, we find some more relevant mentions among people talking online in emerging democracies compared to amongst online talkers in established democracies.

The results most evidently going against our hypothesis is that left-right ideological framing of debates on democracy turned out to be most common, not in emerging democracies but in older more established democracies. It is obvious that the classical left-right conflict dimension – historically dominant in especially Western and European democracies – is still very much relevant, and clearly so when talking about democratic rule. Among the nations where people most often discuss democracy in ideological left-right terms, we find established democracies like the United Kingdom, Ireland, the United States, Canada and Australia; curiously all English-speaking countries who used to be part of the British empire.

LES Data is Useful

The most significant conclusion from this small test using data from the LES project to study popular assessments of democracy across the world is very positive. Using online text data is challenging, particularly because it is difficult to assess its representativity in a valid and reliable way. Although more careful studies of data representativity is certainly warranted, and currently ongoing within the LES project, the preliminary findings of this paper indicate the usefulness of online text data and natural language processing methodology. On the aggregate level, data that reflect people's online communicative behavior is a possible complement or even sometimes a substitute for results from survey studies. And, importantly, data collection is unobtrusive, could potentially be less expensive than fielding large-scale surveys, and need not be affected by the increasing problem of sample loss that survey research is currently suffering from. Doing social science research the LES way has the potential to be a win-win solution.

REFERENCES

Barrling, Katarina & Holmberg, Sören (Eds., 2018). *Demokratins framtid*. Stockholm: Sveriges riksdag.

Dahlberg, Stefan, et al. (2020). "The LES Online Distributional Semantic Lexicon". *Working Paper, Linguistic Explorations of Societies*.

Lürmann, Anna et al. (2020). *Autocratization Surges-Resistance Grows*. Gothenburg: Democracy Report. Varieties of Democracy (V-Dem) Institute.

Mounk, Yascha (2018). *The People vs. Democracy*. Cambridge, MA: Harvard University Press.

Rohrschneider, Robert (1999). *Learning Democracy*. Oxford: Oxford University Press.

APPENDIX

TABLE A1, SPECIFIC DEMOCRACY-RELATED WORDS ASSOCIATED WITH RELEVANT VARIABLES

System of rule	Rule of law	System capacity
tyranny	minority rights	prosperity
oligarchy	human rights	welfare state/welfare society
rule of the people	legal system	market economy
dictatorship	constitution	planned economy
governance	freedom of speech	stability
government	freedom of opinion	meritocracy
parliament	freedom of the press	corruption
coalition	freedom of religion	poverty
opposition	separation of powers	mafia
equality	legitimacy	bureaucracy
majority	justice	health
minority	impunity	modernity
community		cohesion
monarchy		immigration
social partnership		oppression
authority		
leadership		
republic		
elections		
federalism		
Left-right ideology	Religion	Nation/nationalism
liberalism	islamism	globalism
conservatism	secularism	nationalism

socialism	laicity	multiculturalism
communism	religion	nation
capitalism	religious party	racism
ideology		populism
plutocracy		pluralism
imperialism		nationalist party
fascism		zionism
social democracy		

Note: Democracy-related words found in LES project data.

TABLE A2, CLASSIFICATION OF REGIME TYPES AND LIBERAL DEMOCRACY RANK SCORE

Established democracies	Rank	Emerging democracies	Rank	Authoritarian regimes	Rank
Australia	14	Argentina	44	Algeria	148
Austria	30	Brazil	60	China	174
Canada	21	Bulgaria	67	Cuba	159
Chile	75	Colombia	62	Egypt	145
Costa Rica	8	Ecuador	63	Iran	141
Czech Republic	38	Ghana	45	Kyrgyzstan	101
Denmark	1	Hungary	85	Mali	104
Estonia	2	India	90	Morocco	115
Finland	11	Indonesia	65	Pakistan	126
Germany	20	Kenya	97	Philippines	118
Greece	24	Malaysia	98	Russia	156
Ireland	12	Mexico	68	Thailand	139
Italy	22	Namibia	51	Turkey	153
Latvia	34	Nigeria	96	Ukraine	107
Lithuania	27	Paraguay	80	Venezuela	163
Netherlands	10	Peru	40	Vietnam	146
New Zealand	10	Poland	64	Zambia	111
Norway	5	Romania	77	Zimbabwe	134
Portugal	7	Senegal	52		

Spain	9	Singapore	92
Slovakia	31	South Africa	48
Slovenia	33	Sri Lanka	70
Sweden	3	Tunisia	41
Switzerland	4		
Taiwan	37		
United Kingdom	13		
United States	36		
Uruguay	19		

***Note:** The classification is based on country ranks on liberal democracy in the Varieties of Democracy Annual Democracy Report 2020.*