Separated by a common language: How Breitbart and The New York Times produce different meanings from common words

Federico Mor
UK Parliament, UK

Erin J Nash
Durham University, UK

Fergus Green
Utrecht University, The Netherlands and London School of Economics, UK

Abstract
We build on the work by Peled and Bonotti to illuminate the impact of linguistic relativity on democratic debate. Peled and Bonotti’s focus is on multilingual societies, and their worry is that ‘unconscious epistemic effects’ can undermine political reasoning between interlocutors who do not share the same native tongue. Our article makes two contributions. First, we argue that Peled and Bonotti’s concerns about linguistic relativity are just as relevant to monolingual discourse. We use machine learning to provide novel evidence of the linguistic discrepancies between two ideologically distant groups that speak the same language: readers of Breitbart and of The New York Times. We suggest that intralinguistic relativity can be at least as harmful to successful public deliberation and political negotiation as interlinguistic relativity. Second, we endorse the building of metalinguistic awareness to address problematic kinds of linguistic relativity and argue that the method of discourse analysis we use in this article is a good way to build that awareness.

Keywords
discourse analysis, ideology, linguistic relativity, machine learning, political disagreement

Received: 28th July 2020; Revised version received: 30th March 2021; Accepted: 1st April 2021

Corresponding author:
Fergus Green, London School of Economics and Political Science, Houghton Street, London, UK, WC2A 2AE.
Email: R.F.Green@lse.ac.uk
Introduction

The question of how language shapes thought has preoccupied scholars for decades, if not centuries. In an article in the *American Political Science Review*, Peled and Bonotti (2016) argue that ‘linguistic relativity’ creates political challenges for multilingual communities. Linguistic relativity, they explain, is the phenomenon whereby the language we use affects our cognition and the way we understand the world. Their article explores the consequences of linguistic relativity for political theorising and democratic deliberation in societies where more than one language is spoken. They argue that linguistic discrepancies can undermine progress in moral and political scholarship, and act as a barrier to successful public deliberation and political negotiation.

Peled and Bonotti’s solution to this challenge is twofold. First, they argue that states need to ensure that all citizens and residents are taught a common civic language from an early age. Second, they advocate building higher order awareness and appreciation of language, which they call ‘political metalinguistic awareness’.

In this article, we provide novel evidence that shows that linguistic relativity effects occur intralinguistically too. While Peled and Bonotti recognise the possibility of regional and class differences within language groups, our analysis reveals another distinct source of intralinguistic relativity that crosses regional and socio-economic class divides: political ideologies. We obtain this evidence by comparing the language used by two politically distant US news outlets: *The New York Times* (hereafter *NYT*) and *Breitbart*. We then explain why Peled and Bonotti’s assertion that ‘the presence of intralinguistic relativity effects . . . should not be seen as a threat to an otherwise epistemically coordinated demos’ is much too optimistic. Crucially, Peled and Bonotti overlook the way in which intralinguistic relativity effects can be exploited and used strategically for partisan political ends, a tactic we call ‘hacking’.

If we are correct, then Peled and Bonotti’s first strategy to address linguistic relativity – ensuring all citizens learn a common language – is less likely to be an effective strategy. On the contrary, we argue that their second strategy – building political metalinguistic awareness – is more promising and important than they themselves recognise. We also suggest that the methodology we present here can be used by scholars and educators to help build such awareness.

The article proceeds as follows. We first introduce the theoretical context for our investigation, covering scholarship on political disagreement and ideology. Then, we provide an account of our methodology and outline our results, focusing on three key topics. In the discussion section, we compare inter- and intralinguistic relativity, explain how linguistic relativity effects can be strategically ‘hacked’, and describe how metalinguistic awareness could be built.

Theoretical context

Our wider theoretical aim in this article is to place the literature on linguistic relativity underpinning Peled and Bonotti’s article into conversation with work in political science that suggests the potential for a virulent and harmful form of intralinguistic relativity, which in turn has implications for the likely effectiveness of the strategies they suggest.

Political disagreement

Peled and Bonotti’s (2016: 800–806) article critiques Rawlsian liberalism on the basis that the kind of deliberative reason central to Rawls’ theory is undermined by the interpretive
pluralism to which linguistic diversity gives rise. The theoretical springboard for our dis-
cussion of intralinguistic relativity can be traced to a different critique of Rawlsian-style liberal-
ism by Chantal Mouffe (1993, 2000, 2005). Mouffe argues that Rawls (1971, 1993) is misguided in thinking that political disagreements are resolvable through public reason given that a certain type of antagonism is inherent to politics. According to Mouffe, political identities necessarily imply us-versus-them boundaries, often hierarchically ordered, and these identities are constructed in the course of people’s varying life experiences. This multiplicity of binary political identities, argues Mouffe, gives rise to the ever-present possibility of antagonistic relations and disintegrative conflict in political life, forever preclud-
ing the securing of a consensus. This pluralism, in Mouffe’s view, is a good thing because it underpins and illuminates the contingency of seemingly natural political orders and opens them up to counter-hegemonic struggles (Laclau and Mouffe, 1985), such as those carried out by the progressive social movements of the last quarter of the 20th century (Mouffe, 1993). Language and discourse play a crucial role in this antagonistic politics, since they are themselves sites of (counter-)hegemonic struggles over the meaning of political concepts and ideals that are otherwise empty signifiers (Laclau and Mouffe, 1985; Mouffe, 1993).

However, critics have expressed concern that Mouffe’s rejection of rationalistic foun-
dations for democracy may be too strong, undermining her own attempt to defend a dem-
ocratic project committed to the ideals of liberty and equality (Sparks, 1994: 148–150). Thus, whereas Mouffe’s work emphasises the emancipatory potential of political and discursive contestation, in this article we foreground a darker side: the potential for politi-

cal elites to foment and exacerbate political divisions for strategic purposes through their
control over news media. Specifically, we highlight how news media use language to
frame issues so as to construct and activate particular political ideologies among their
audience, mobilising them for partisan political ends. This process, which is elaborated
below, is well understood. Our contribution is to show how a fundamental part of that
process relies on producing and exploiting (‘hacking’, in our terms) intralinguistic rela-
tivity and to consider what this implies for strategies to improve democratic debate.

Political ideology

Individuals and groups hold ‘distinctive idiosyncratic worldviews that meaningfully shape their political thought and political behaviour’, which are best understood in terms of systems of ideas known as ideologies (Leader Maynard, 2017: 298). We follow an emerging consensus among scholars of ideology in understanding ‘political ideology’ in a broad and non-pejorative way (see Leader Maynard, 2017: 300).

To understand how ideologies are stabilised and how they change, it is useful to under-
stand ideology in terms of interactions between two levels: the individual level and the group level. At the individual level, ideologies can be understood as networks of interconnected mental elements (beliefs, concepts, images, etc.) that are emotionally valenced and often central to the holder’s identity (Homer-Dixon et al., 2013: 344–345, drawing on Thagard, 2006). Thus, the meaning of political phenomena for individuals (1) ‘emerges from the interconnections between multiple cognitive elements, structures, and processes as much as from their relationship to entities in the material and social worlds’ (Homer-Dixon et al., 2013: 344, citing Markus and Hamedani, 2007) and (2) is shaped by the emotional valences of these mental elements, and not simply by logical reasoning (Homer-Dixon et al., 2013: 345).
At the social level, scholars emphasise how ideology is socially constructed through discourse and practices, and is intertwined with multiple forms of power (Van Dijk, 1998). The discursive link is particularly important for our purposes. Ideologies are marked by certain rhetorical repertoires that individuals draw on for the purposes of argument, legitimation, or other speech acts, and these come in turn to shape the content of groups’ ideologies (Van Dijk, 1998, 2013). It is, moreover, widely recognised that political actors ‘deploy discourse strategically in political contests, creatively shaping existing political language to suit their needs’ (Homer-Dixon et al., 2013: 340–341 and references there cited), including mobilising their supporters, that is, motivating them to act collectively for a political purpose (Tarrow, 2013). Finally, discursive-ideological power is underpinned by other forms of power, such as control over communications technologies and news media (Homer-Dixon et al., 2013: 341; Fairclough, 2001; Simonds, 1989).

Framing

One key way in which news media deploy discourse for strategic-ideological purposes is through issue framing. A frame is a feature of a message (social level) or of a representation about a message (individual level), while framing is the dynamic process of advocating or applying a frame to an issue (Chong and Druckman, 2007; Nelson, 2011: 191). Of particular interest to us here are strategic forms of framing by which political actors seek to simplify an issue by highlighting a certain aspect of it (and obscuring others), typically in a way that is valenced and lays blame (or praise) and that favours (or disfavours) a particular solution (Nelson, 2011: 192–193). Social scientists use the discursive analysis of issue frames in the news media to study the role of those media in society, including as a basis for drawing inferences about the ideology promoted by particular media outlets.

Framing, ideology, and political mobilisation in NYT and Breitbart

In our study, we consider NYT and Breitbart because NYT is the most ideologically distant mainstream news outlet from Breitbart in the United States (Budak et al., 2016). NYT is a highly influential, mainstream, legacy newspaper that leans moderately liberal (left) on the standard liberal–conservative spectrum (Budak et al., 2016). Breitbart.com, founded in 2007 by Andrew Breitbart and taken over by Steve Bannon in 2012, exemplifies a ‘hyperpartisan’ (Rae, 2021) form of polemical right-wing news, establishing itself as a ‘a rallying point for the ‘alt-right’, a loosely constructed ‘movement’ that brings together white supremacism, anti-feminism, anti-Islamic, and anti-Semitic extremism’ (Davis, 2019: 244, citing Posner, 2016).

The rise of right-wing partisan and hyperpartisan news sites such as Breitbart has played a key role in reinvigorating the US ‘culture wars discourse’ of the 1980s and 1990s, that is, embedding and activating a ‘more-or-less standardised set of simple binary constructs [. . .] to caricature and reframe complex issues as a struggle between a virtuous “we” and demonic “they”’ (Davis, 2019, 243). The discourse is ‘rhetorical and polemical, designed to dramatise perceptions of conflict and division for political gain [. . .] and operates precisely as a “war of words” intended to underpin populist political strategies’ (Davis, 2019: 244). In this regard, Breitbart forms part of a network of right-wing media outlets, think tanks, republican party operatives, and conservative donors and functions to energise and mobilise grass-roots supporters towards partisan ends (Meagher, 2012). For example, Breitbart was especially prominent in engaging audiences in the 2016 US election, receiving more
audience online engagement than Fox News, while left-leaning audiences gravitated towards traditional media such as NYT (Benkler et al., 2017).

The relationship of this theoretical discussion to our study can be summarised in terms of two propositions. First, since ideologically distant news media deploy discourse to frame public issues in divergent ways, using emotionally valenced binary constructs, we have a theoretically motivated reason to expect that, for certain politically significant words, those same words will tend to be associated with different sets of words across the two media sources. Second, since individuals interpret the meaning of political phenomena via cognitive schemata that are structured according to networks of emotionally valenced mental elements, we have a theoretically motivated reason to expect that readers of these diverse news media will come to interpret the meaning of such common words quite differently, and that these differences will affect their political behaviour, including their engagement in public debates. Our study tests the first proposition using the method discussed in the next section. Our subsequent discussion assumes that the second proposition generally holds, considers the implications for democratic debate, and evaluates the proposals of Peled and Bonotti in that light.

Data and methodology

To test the proposition that ideologically distant news media construct divergent meaning from common words, we used an unsupervised machine-learning algorithm to create two separate word models, one trained with text from NYT and the other with text from Breitbart.\footnote{Our web-scraping code pulled text from all the articles appearing on the first page of results when searching by date, every day from 1 January 2016 to 20 June 2018.} In order to obtain articles covering comparable topics, we restricted the daily article search to the ‘politics’ category on NYT and what was then called the ‘big government’ category on Breitbart,\footnote{which was a much wider category than the label suggests, covering all things political.} In order to obtain articles covering comparable topics, we restricted the daily article search to the ‘politics’ category on NYT and what was then called the ‘big government’ category on Breitbart,\footnote{which was a much wider category than the label suggests, covering all things political.}

It is sometimes assumed that algorithms produce inherently objective results, because they are free from human bias. That assumption has been rightly criticised from a range of quarters (see, for example, O’Neil, 2016). Rather than being free from bias, machine learning faithfully reproduces the biases in the data that the algorithms are trained with.

While this feature of machine learning is usually a weakness, it is in fact a strength for our purposes. Our method relies precisely on machine learning to impartially reflect back at us the biases embedded in the language used by Breitbart and NYT. Pääkkönen and Ylikoski (2020: 1) support this view, arguing that ‘unsupervised machine learning methods might make hermeneutic interpretive text analysis more objective in the social sciences’ by ‘allowing interpreters to discover unanticipated information in larger and more diverse corpora and by improving the transparency of the interpretive process’. They are clear that these methods do not ‘eliminate the researchers’ judgments from the process of producing evidence for social scientific theories’, and we agree that the researchers’ interpretive engagement with the original text and model outputs remains key. Rather:

[T]he sense in which unsupervised methods can improve objectivity is by providing researchers with the resources to justify to others that their interpretations are correct. This kind of objectivity seeks to reduce suspicions in collective debate that interpretations are the products of arbitrary processes influenced by the researchers’ idiosyncratic decisions or starting points. (Pääkkönen and Ylikoski, 2020: 1)
In this vein, we argue that our method allows us to identify words that have different meanings in the two publications in a manner that is more objective and comprehensive than what a human reader—even a competent, fair and time-rich one—could do. However, the subsequent interpretive steps (from deciding which differences are truly meaningful to hypothesising which factors might explain them) inevitably involve the researchers’ judgement. To be clear, the two word models that we created do not literally tell us what words mean. Rather, they inform us as to what other words are most similar to a given word. This works well because concepts acquire their meaning in relation to other concepts, as we explained in our theoretical section.

When a word is used in similar ways by both Breitbart and NYT, we would expect the two word models to produce similar results. The match does not have to be perfect, but there should be a good deal of semantic overlap between the two sets of most similar words. What makes words ‘most similar’ in the eyes of the algorithm is appearing in the most similar context. When words share similar contexts, they are embedded close to each other in the model. An intuitive way to conceive of this type of word similarity is to think about the options served to us by word predictors when typing. Some of the options are most similar in the sense that they are expected to appear in the same context (what we typed up to that point). These are the concepts that travel well intralinguistically. On the contrary, we would expect to detect intralinguistic relativity if the models produce divergent results for the same word, revealing a different terminological network.

As a sense-check, we expected to see substantial convergence between the models with respect to basic, uncontentious words. The reasoning is that it would not make sense for basic words to mean different things in Breitbart and NYT. As English language publications, they ought to share at least some core syntax and vocabulary. The presence of differences vis-à-vis basic words would make us doubt that other differences do indeed track real instances of linguistic relativity and are not spurious.

As anticipated, we found significant convergence between the models in relation to basic words as well as some potentially contentious or politically charged words. For example, the word ‘will’ returns ‘won’t’ and other modal verbs in both models. ‘Sustainable’ means ‘long-term’, ‘innovative’, and ‘robust’ in both publications. ‘African Americans’ is most similar to ‘Blacks’, ‘Latinos’, and ‘Whites’. ‘America’ is ‘our nation’ and ‘our country’, while ‘Iraq’ is most similar to ‘Egypt’, ‘Afghanistan’, ‘Yemen’, and ‘Syria’.

**Results**

We now present findings focusing on three salient and vigorously contested topics within US public debate: abortion and contraception; the economy and the role of government; and climate change. These or similar topics were cited as priorities among a majority of voters in both 2016 and 2018, according to surveys from the Pew Research Center, and are ‘traditional culture wars themes’ (Davis, 2019: 244–245). Even where there is cross-ideological agreement about the importance of an issue, such as the economy, we still expect ideologically distant news media to promote divergent agendas, with liberal media being more open to the role of government and right-wing media being more hostile to it and supportive of markets as a solution to social problems (Davis, 2019: 244–245). This makes these topics likely loci of intralinguistic relativity, and we contend that our findings evidence it.

For each topic, we present a table of topic-related words or phrases. The table shows how the same word or phrase is embedded in divergent terminological networks (its ‘five
Mor et al.

most similar words”) in each publication, and our interpretation maps this linguistic divergence onto divergent conceptual networks, which as noted earlier form the basis of ideologies. We quote passages from both publications to illustrate our interpretation of the tables, all of which passages come from the text that was used to train the algorithm.

**Topic 1: Abortion and contraception**

Our first divergent topic is abortion and contraception (Table 1). ‘Babies’ in *NYT* are about ‘boxes’, ‘bottles’, and ‘drinks’ — middle-class parental preoccupations and policies, perhaps, of which the following passage may be an example:

> For years, local officials in South Korea, which has one of the world’s lowest birthrates, have tried ever more inventive plans to encourage women to have babies. They have offered generous maternity-leave policies, cash allowances and even boxes of beef and baby clothes to families with newborns.7

On the contrary, the babies *Breitbart* talks about are the ‘unborn children’, as this passage illustrates:

> The state legislature is close to passing a Pain-Capable Unborn Child Protection bill patterned after the one that was passed by the U.S. House. The measure, currently in conference, is based on scientific evidence that unborn babies are able to feel the pain of the abortion procedure at least as early as in the twentieth week of pregnancy, or five months.8

It is striking that all five of the words most similar to ‘babies’ are related to the abortion debate in *Breitbart*, and none in *NYT*. Also striking in *Breitbart* is that the closest association with ‘abortion rights’ is ‘NARAL’, an abortion rights advocacy group.9 This may indicate a deliberate conflation of the issue (the rights) with the advocacy group itself, which *Breitbart* consistently describes in variations on the theme of ‘political lobbying group for the abortion industry’ and ‘abortion lobbying giant’.10 By the same logic, ‘abortion rights’ are synonymous with the ‘abortion industry’:

> Fearful of being blamed by Democrats and the liberal media for causing another government shutdown, GOP leaders opted to fund the abortion business for most of 2016. [. . .]
Ryan’s office adds that the videos produced by Center for Medical Progress that exposed Planned Parenthood’s apparent practices of harvesting the body parts of aborted babies for sale and altering the positions of babies during abortions in order to harvest intact organs, ‘should offend everyone regardless of political persuasion’.11

Finally, all five most similar words are related to abortion, indicating that abortion rights belong to a category of their own in Breitbart’s discourse. That is not the case in NYT, where abortion rights are a subset within a broader category of rights and are therefore most similar to other rights within the broad category such as ‘gay rights’, ‘gun rights’, and ‘civil rights’. The passage below illustrates how abortion rights feature in the same context as other rights, including healthcare and gay rights:

She is for abortion rights; he is firmly opposed to them. She wants to expand health care; he wants to blow up the Affordable Care Act. She wants to raise taxes on the rich; he wants to impose a flat tax that would lower them. She is an eager supporter of gay marriage; he views its legalization as an infringement on religious liberty.12

In NYT, ‘family planning’ and ‘birth control’ are entwined with ‘charitable’ and state efforts to reach deprived segments of the population, as the associations with ‘social welfare’, ‘federal funds’, ‘food stamps’, and ‘low income’ indicate. It is worth noting how making children is something to be encouraged among certain segments of the population via ‘generous maternity-leave policies’ (see earlier quote) and discouraged in others via birth control, where the activity is described as ‘popping out babies’ and equated with welfare fraud, drug problems, and domestic violence:

A single night of gambling or whatever can produce life-altering bad choices. Moreover, the forces of social disruption are visible on every street: the slackers taking advantage of the disability programs, the people popping out babies, the drug users, the spouse abusers. Voters in these places could use some help. But these Americans, like most Americans, vote on the basis of their vision of what makes a great nation.13

At the same time, ‘birth control’ in NYT is also just a normal part of ‘medical care’ and ‘health plans’. In Breitbart though, the focus remains principally on ‘babies’, ‘abortions’, and ‘abortion providers’, although it also includes ‘taxpayer funding’ (which as a phrase, with its focus on the taxpayer, tends to be used within more critical discourse than the more neutral ‘federal funds’).

**Topic 2: The economy and the role of government**

Table 2 shows results for the second topic, the economy and the role of government. A number of patterns and themes can be recognised and may warrant further investigation. First, there is evidence of the small-state, low-tax versus big-state agendas (as per Davis, 2019: 244–245). For example, ‘working families’ are associated with ‘lower taxes’ and ‘cut taxes’ in Breitbart, whereas in NYT working families feature predominantly as working-class voters, and perhaps for that reason are closely associated with ‘incumbents’ and ‘elected officials’, whose re-election depends on these voters.

The NYT column also indicates the liberal embrace of state spending and intervention through closely associating general words or phrases with specific government programmes and policy priorities – associations that are wholly missing from Breitbart’s
Table 2. The economy and the role of government.

<table>
<thead>
<tr>
<th>Word</th>
<th>NYT</th>
<th>Breitbart</th>
</tr>
</thead>
<tbody>
<tr>
<td>working_families</td>
<td>working-class_voters, incumbents, labor, retirement_age, elected_officials</td>
<td>lower_taxes, small_businesses, cut_taxes, create_jobs, our_economy</td>
</tr>
<tr>
<td>taxpayer_dollars</td>
<td>nursing_homes, capital_gains, prescription_drugs, medicines, utilities</td>
<td>tax_dollars, taxpayer_funds, taxpayer_money, taxpayers, funds</td>
</tr>
<tr>
<td>federal_government</td>
<td>program, agency, cuts, local_governments, medicare</td>
<td>government, federal_funds, taxpayers, federal, federal_law</td>
</tr>
<tr>
<td>labor_market</td>
<td>renewable_energy, productivity, social_welfare, job_creation, public_education</td>
<td>wages, job_market, u.s._economy, americans_wages, workforce</td>
</tr>
<tr>
<td>Economies</td>
<td>populations, societies, sectors, regions, democracies</td>
<td>markets, industries, exports, plants, currency</td>
</tr>
<tr>
<td>Financial</td>
<td>domestic, industry, regulatory, fossil_fuel, corporate</td>
<td>banking, consumer, lending, lucrative, assets</td>
</tr>
<tr>
<td>Wealth</td>
<td>capitalism, inequality, vast, tolerance, poverty</td>
<td>middle_class, economies, coal, industrial, profits</td>
</tr>
<tr>
<td>Regulators</td>
<td>private_companies, contractors, entities, tech_companies, firms</td>
<td>contracts, federal_agencies, subsidies, government_agencies, treasury</td>
</tr>
<tr>
<td>Regulated</td>
<td>for-profit, regulating, public_sector, restricting, decades-old</td>
<td>eliminated, restricted, mandated, administered, subsidized</td>
</tr>
</tbody>
</table>


column. For instance, the ‘federal government’ and ‘taxpayer dollars’ are closely associated with healthcare in NYT through associations with ‘medicines’, ‘prescription drugs’, and ‘Medicare’, as well as ‘nursing homes’ and ‘program[s]’ and ‘agenc[ies]’ in general. Similarly, ‘labor market’ and ‘financial’ are each associated with a climate change–related phrase, ‘renewable energy’ and ‘fossil fuel’, respectively. No such associations exist in Breitbart, where the associations remain close substitutes for the examined word. For example, Breitbart’s associations for ‘taxpayer dollars’ are all strict synonyms, such as ‘tax dollars’ or ‘taxpayer money’ and ‘taxpayer funds’. Even when some associations are not direct synonyms, the focus remains strictly economic (‘our economy’) and translates narrowly into issues such as ‘small businesses’ and ‘creat[ing] jobs’.

The publications’ respective views on state intervention are well captured by their attitudes to regulation. To be ‘regulated’ is to be ‘eliminated’ or ‘subsidized’ in Breitbart, and the associations focus exclusively on those doing the regulating and its effects. By contrast, the majority of NYT’s associations concern the regulated, that is, the for-profit private sector in general, and in particular ‘tech companies’ and the financial sector. The two passages below may serve as an illustration of Breitbart’s discourse in this area:

Social justice warriors (SJW) like Senator Elizabeth Warren (D-MA) have praised negative interest rates as part of an effort to convert banks into regulated public utilities that they promise will achieve broader lending by drastically restricting profits. But the real goal is giving Congress the ability to maximize its meddling in 17 percent of America’s GDP.14

This is classic anti-gun leftism. Gun control is failing, so pass more of it by targeting an already highly regulated group of people – Federal Firearms License holders (FFLs) – who cannot afford to go against the strong arm of the government for fear of losing their license.15
Second, some of the Breitbart associations may also reflect the mercantilist approach to international affairs adopted by Donald Trump (and Steve Bannon both as Breitbart editor and Trump advisor). This hypothesis might explain the divergence in the meaning of ‘economies’, which in Breitbart means ‘plants’, ‘exports’, ‘markets’, and ‘currency’, while in NYT it is synonymous with ‘societies’, ‘populations’, and ‘democracies’. This passage from Breitbart is perhaps an example of this discourse:

The European Union [has] consistently failed to negotiate a trade agreement with the United States, so its member-states will likely be hit when planned tariffs on steel and aluminium, designed to protect American workers from competitors in low-wage economies, come into force.¹⁶

Third, NYT gives ‘wealth’ a clear distributional and ethical dimension: it is ‘vast’ and also unequal (‘inequality’), and therefore also associated with ‘poverty’. The following passage may illustrate NYT’s discourse:

The report lays out a stark narrative about the American economy as it exists today. Inequality, it maintains, is a function not of economic laws but of the preferences awarded to the powerful to extract rents – to exploit people who have little choice – especially on necessary goods like housing and health care. [. . .]. The economy has stalled because too much wealth is being generated in nonproductive activity, hoarded to preserve for the rich all the things government no longer provides.¹⁷

In Breitbart too, wealth is associated with a word that has a potential distributional dimension, but one that is narrow, geographical, and issue-specific: ‘coal’. That association may reflect Breitbart’s championing of, and appeal to, the formerly prosperous coal communities and voters.

**Topic 3: Climate/climate change**

‘Climate’ and ‘climate change’ (Table 3) is our third topic. In NYT, ‘climate change’ is most similar to words like ‘immigration’, ‘encryption’, ‘trade’, ‘policy’, and ‘issues like’. Similarly, ‘climate’ is associated with ‘governance’, ‘institutional’, and ‘strengthening’. One possible explanation for these results is that within NYT, climate change is understood to be one of a number of politically charged policy areas alongside immigration, encryption, and trade. Furthermore, words like ‘governance’ and ‘institutional’ suggest technocratic solutions.

<table>
<thead>
<tr>
<th>Word</th>
<th>Five most similar words in:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>climate_change</td>
<td>immigration, policy, issues_like, encryption, trade</td>
<td>global_warming, climate, inequality, science, co2</td>
</tr>
<tr>
<td>Climate</td>
<td>governance, strengthening, dynamic, institutional, broader</td>
<td>climate_change, warming, global_warming, scientific, co2</td>
</tr>
</tbody>
</table>

By contrast, the language *Breitbart* uses to talk about climate change – such as ‘global warming’, ‘science’, and ‘CO₂’ – suggests that its journalists view, or at least portray, climate change as a largely scientific issue rather than one of public policy. On first blush, this might be surprising; *Breitbart* is a well-known forum for various shades of climate contrarianism and denial (see, for example, Bloomfield and Tillery, 2019). However, the fact that ‘climate change’ is most similar to ‘global warming’ and ‘science’ does not mean that *Breitbart*’s editors and journalists accept the reality of anthropogenic climate change; rather, it is more likely that these three terms appear similar because *Breitbart* deploys them in the same disparaging context and casts doubt on all three. The passage below illustrates this:

Global food prices are down for the fourth year running. […] According to the Malthusian alarmists who dominate the world’s scientific establishment this just isn’t supposed to happen. […] As I mention in Watermelons, this belief that there are too many people for the planet to feed can be traced at least as far back as the 2nd century Carthaginian priest Tertullian […] But that was 1800 years ago when scientific knowledge was harder to come by. Since then, you might have thought, intelligent people would know better. […] Far worse, though, than the injustice of these posturing pillocks being taken seriously and showered with awards is the damage their junk-science fearmongering does to real people.18

**Other findings for future research**

Before moving to a discussion of the implications of our findings, we briefly note here some other results and thoughts that may be the subject of further interest or research. First, we posit that word associations involving different syntactic categories may be especially revealing, since we expect (and find) that a word’s most similar words tend to belong to the same category. Table 4 gives the example of an adjective whose most similar words include two entirely unrelated nouns. The association between an adjective and a noun is likely the result of the noun being systematically presented in the adjective’s colours. And so, the disastrous policy *par excellence* in *Breitbart* is the ‘trans-pacific partnership’, and the disastrous person is ‘his predecessor’ (i.e. Barack Obama, at the time of the Trump presidency). For *NYT*, it is the near future (‘forecast’, ‘looming’) that is anticipated to be disastrous.

Second, this type of analysis can also reveal a strong focus on a particular aspect or dimension of a multifaceted word. Table 5 offers some examples. For instance, ‘left’ in *NYT* is first and foremost the past participle of ‘to leave’ (and as such it is most similar to other past participles of related verbs), whereas in *Breitbart* ‘left’ is the political left. This difference may also reflect who the ‘us’ and the ‘them’ are in these publications. In *Breitbart*, Washington is not just the capital, but also the ‘swamp’. Terrorists and other

<table>
<thead>
<tr>
<th>Table 4. Disastrous.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
</tr>
<tr>
<td>NYT</td>
</tr>
<tr>
<td>disastrous</td>
</tr>
</tbody>
</table>

hostile actors are ‘targets’ in *Breitbart*, while in *NYT* it is websites and tech companies. Terrorists only have one religion in *Breitbart*.

**Discussion**

*Comparing inter- and intralinguistic relativity*

These results provide novel evidence that linguistic relativity affects not only communication between languages, but also within them. The divergences outlined above reveal the ‘situated webs of linguistic labels that encode and connote contextualized notions – social, political, and otherwise’ (Peled and Bonotti, 2016: 807–808).

Peled and Bonotti argue that this phenomenon can sometimes act as a barrier to successful public deliberation and political negotiation. They explain how this can be so with reference to differences between Danish and American speakers regarding ‘family values’ (‘familieværdier’), which for Danish speakers implies support for parental leave and free childcare, connotations unlikely to be heard by an American ear:

Thus, when engaged in public reasoning in English with regard to the state provision of free childcare, many English and Danish speakers will be using at some point the English term ‘family values’ in order to support or reject that measure or to specify how it should be implemented in more detail. Regardless of their specific position on the issue (e.g. some of them may be socially conservative, others more liberal), native English speakers will be ‘nudged’ to associate ‘family values’ with what is normally considered as the traditional nuclear family. Similarly, [...] Danish speakers will be ‘nudged’ to think that ‘family values’ almost implicitly involve the state provision of free childcare and they will not draw any distinction between married and unmarried, or same-sex, couples when using that term, because that distinction is not central to the semantic baggage of the term familieværdier. (Peled and Bonotti, 2016: 804–805)

Similarly, when readers of *Breitbart* and *NYT* read ‘climate change’, ‘babies’, or ‘the economy’, it evokes connotations that are unlikely to be heard by the ears of readers of the other publication. They are each being nudged to think in a certain way, just like

<table>
<thead>
<tr>
<th>Word</th>
<th>NYT</th>
<th>Breitbart</th>
</tr>
</thead>
<tbody>
<tr>
<td>left</td>
<td>lost, leaving, kept, gained, gone</td>
<td>spotlight, far_left, establishment, political_class, other_side</td>
</tr>
<tr>
<td>washington</td>
<td>capitol_hill, budget_director, white_house, hill, west_wing</td>
<td>washington_d.c., d.c., washington_dc, swamp, both_parties</td>
</tr>
<tr>
<td>targets</td>
<td>websites, arrangements, tech_companies, applications, methods</td>
<td>militants, terrorist_organizations, weapons, extremists, regimes</td>
</tr>
<tr>
<td>terrorist</td>
<td>smuggling, weapons, terrorists, foreign_intelligence, foreign</td>
<td>terrorist_group, jihadi, jihadist, terror, islamist</td>
</tr>
<tr>
<td>terrorist_groups</td>
<td>militant_groups, boko_haram, rebel_groups, rakhine, militant_group</td>
<td>terrorist_organizations, al-qaeda, al_qaed, hezbollah, jihadists</td>
</tr>
</tbody>
</table>

Danish and English speakers are when the term ‘family values’ is used. And if the discrepant semantic baggage of ‘family values’ can have significant effects on public debates and political negotiations between members of mismatched linguistic groups, then so too can the discrepancies that we have highlighted. Thus, democratic deliberation between members of different ideological groups can be frustrated by linguistic relativity, even when people share the same native tongue.

As Peled and Bonotti (2016: 804) remark, many native speakers ‘have no particular knowledge of the specific semantic baggage’ of words in their own language, and are probably unaware, or at least less mindful, of the potential for meaning to get ‘lost in translation’ when no translation is taking place. Because of its unconscious nature, its subtlety, and the difficulty in revealing and addressing it, intralinguistic relativity can be at least as problematic as when native political translators mislead their non-native interlocutors in lingua franca deliberations (Doerr, 2018).

**Hacking intralinguistic relativity**

Doerr’s work on political translators also highlights an issue that Peled and Bonotti overlook: linguistic relativity effects can be strategically exploited, or hacked, by those who wish to manipulate political debate in the service of a political agenda. This phenomenon is distinct from, but related to, other mechanisms that distort debate such as the spreading of mis/disinformation and propaganda, the selective sharing of evidence, or the creation of echo chambers (Benkler et al., 2018; O’Connor and Weatherall, 2019).

The contrasting presentations of our three topics illustrate this phenomenon well, revealing how different effects of intralinguistic relativity can be put to use. One such effect is to cement the casting of an issue as primarily a controversy over facts or over values (cf. Pielke Jr, 2007). For instance, by consistently presenting climate change as a ‘scientific’ issue, Breitbart is better able to dismiss it through the use of a complementary distortive tactic: mis/disinformation. Through such mis/disinformation, conservative elites have manufactured and maintained a public epistemic controversy over basic descriptive facts about the existence and key causes of climate change that have been accepted by the overwhelming majority of published climate scientists since at least the early 1990s (Oreskes, 2004; Shwed and Bearman, 2010). Thus, the linguistic cementing of climate change as a factual question contributes to widening ideological gaps between different groups, frustrating policymaking progress on climate change (Kitcher, 2011; Nash, 2018).

A second effect that is exploited is the reinforcement of linkages between issues (or between issue positions), especially in a way that strengthens readers’ sense of group identity, that is, their loyalty to an in-group and resentment of an out-group (cf. Iyengar et al., 2019). For instance, the algorithm found ‘climate change’ to be similar to ‘inequality’ in Breitbart (Table 3). One plausible explanation for this result is that both climate change and inequality are portrayed by Breitbart as issues of the ‘progressive left’. To some extent, NYT does the same, for example, when it frames climate change as a technocratic issue (as highlighted earlier) that its middle-class and upper-middle-class, well-educated readers – many of whom have managerial roles in society – are uniquely positioned to solve.19 These political identities, once internalised, can be used as heuristics to filter future information on such topics, as individuals evaluate claims and evidence in ways that reflect the beliefs and opinions of their political in-group (Kahan, 2017).
A third effect that is hacked is the attachment of associations to a concept that trigger strong emotions. This is illustrated by the way Breitbart frames abortion to provoke outrage. As we saw earlier, abortion is closely associated with an ‘abortion industry’ which is portrayed as systematically harvesting babies’ organs (see earlier quote). Such strong emotions can overwhelm any other mode of cognitive engagement with an issue (Haidt, 2012), and they can mobilise people to take political action in ways that reasoned argument often fails to do (Jasper, 2011). We also see this in Breitbart’s framing of government as a predatory, threatening actor, which works to delegitimise government’s role in the production and distribution of economic value.

Hacking can be most powerful when different effects are combined. Consider, for example, how the emotional valence and issue-linkages associated with one concept (e.g. abortion) can be exploited to mobilise people in support of a wider ideological agenda. This is precisely what the American right has done in the case of abortion and various other morally charged social issues (not to mention racial divisions): it successfully aligned (White) religious and socially conservative voters and movements with corporate conservatives, and mobilised them to support the republican party’s neoliberal economic policy agenda (Hacker and Pierson, 2010: 181; Haney-López, 2014; Schlozman, 2015; Schlozman and Rosenfeld, 2019). In our study, we see how such a strategy can be supported by linguistic hacking, given that one of the five most similar words to ‘family planning’ and ‘contraception’ in Breitbart is ‘taxpayer funding’ (Table 1).

Responses to intralinguistic relativity

Intralinguistic relativity will always be a feature of politics. In what follows, we are not suggesting that it is always problematic, that we should aim to eliminate it or that addressing it will resolve all political disagreements. But we do agree with Peled and Bonotti that certain types or instances of linguistic relativity can act as a barrier to successful democratic deliberation and political negotiation, that such barriers are not inevitable, and that removing them is likely to have both procedural and substantive benefits.

Peled and Bonotti’s first strategy for overcoming the problems associated with linguistic relativity – teaching a common language – clearly cannot address intralinguistic relativity. On the contrary, their second strategy – building political metalinguistic awareness – is more promising and important than they seem to appreciate. Greater political metalinguistic awareness can facilitate deeper insight into how people from different ideological groups understand the concepts doing work in political discussions. It can also provide self-insight into the ways in which we are all being subtly ideologically influenced by those few within society who hold the power to shape political discourse and ideology, and the way we are potentially being deployed to secure ends within elite political projects.

While working out the details of a strategy for building political metalinguistic awareness is beyond the scope of this article, we suggest one potential element: an application that builds on the method used in this article to compare the language of different sources and enable users to identify and understand discrepancies. Such a tool could raise awareness of the potential for meaning to get lost in translation within one’s own language. It might not lead to the bursting of epistemic bubbles and echo chambers so much as to increase their occupants’ awareness of the fact that they and others are embedded within them and provide a sense of what language might best facilitate communications with members of other groups (Boyd, 2019).
Conclusion

In this article, we have argued that Peled and Bonotti’s analysis of, and concerns about, the influence of linguistic relativity on political thinking in linguistically diverse societies are just as relevant in monolingual discourse. In fact, we suggest that such intralinguistic relativity could be more harmful to successful public deliberation and political negotiation because people are less likely to turn their mind to it, because it is more subtle and difficult to reveal, and because it can be hacked for instrumental purposes.

We endorse Peled and Bonotti’s strategy of building metalinguistic awareness. Greater awareness would help ideologically different groups to hear and understand one another better. We suggest that the method we have used in this article – highlighting linguistic discrepancies between groups that speak the same tongue – is a good way to build this awareness.

Linguistic relativity is inevitable, and not all of it is harmful. More research is needed to understand the cases and instances that genuinely nourish democratic debate. But the deliberate and strategic hacking of linguistic relativity only contributes to widening gaps within our divided societies. That is why researching and tackling this phenomenon is an urgent task, lest the barriers to successful public deliberation and political negotiation keep mounting.

Acknowledgements

The authors thank Matteo Bonotti and Yael Peled for their very helpful feedback on an earlier draft and their generous engagement with the paper. They also thank Sally Davies, Diane Bolet, and Nicole Vincent for helpful conversations that informed this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Federico Mor https://orcid.org/0000-0001-9469-9833

Notes

1. To create the models, we used a free Python library called ‘gensim’, by Radim Rehurek (https://radimrehurek.com/gensim/models/word2vec.html). We used the ‘Word2vec’ model, based on work by Mikolov et al. (2013) and available on Google’s website (https://code.google.com/archive/p/word2vec/).
2. The end date reflects the time when the web scraping was carried out, rather than any meaningful end point.
4. The address was https://www.breitbart.com/big-government, although Breitbart have changed their website categories since then.
5. Note that some of a word’s most similar words can be antonyms. This is because while these words have opposite meanings, they perform the same function and appear in the same context.
8. See Breitbart (2016).
9. See: https://www.prochoiceamerica.org/
10. For example: Berry (2018).
11. See Berry (2016).
17. See Lewis-Kraus (2016).
19. The NYT’s media kit describes the newspaper’s audience as follows: ‘The NYT reaches a deeply engaged
and highly influential reader. The NYT Weekday ranks #1 with Opinion Leaders, reaching 57% of this elite group. Our top management audience spends more than an hour (72 average minutes) with the
Sunday paper [. . .]. We deliver an audience of readers who shape society’. https://nytmediakit.com/index.
php?p=newspaper [accessed 12 Sep 2019].

References


**Author biographies**

**Federico Mor** was a Senior Researcher at the House of Commons Library, an impartial research service for UK Members of Parliament.

**Erin J Nash** is a Philosopher working at the interface of social philosophy and the philosophy of science. Her research explores issues with misinformation, especially about science, and the ethics and politics of communication.

**Fergus Green** is a Researcher whose work straddles political theory, political economy and public policy. Much of his research explores normative issues in the design and reform of public policy and political/legal institutions, with a focus on climate change and low-carbon energy transitions.