

https://doi.org/10.1017/9781009451727 Published online by Cambridge University Press

ISSN 2634-7334 (online)
ISSN 2634-7326 (print)

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Cambridge Elements

Elements in Forensic Linguistics
edited by
Tim Grant
Aston University
Tammy Gales
Hofstra University

DIGITAL MULTILINGUALISM AND PLATFORM GOVERNANCE

Janny H. C. Leung
The University of Hong Kong

CAMBRIDGE UNIVERSITY PRESS

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009862462

DOI: 10.1017/9781009263177

© Janny H. C. Leung 2026

This publication is in copyright. Subject to statutory exception and to the provisions of
relevant collective licensing agreements, with the exception of the Creative Commons
version the link for which is provided below, no reproduction of any part may take
place without the written permission of Cambridge University Press & Assessment.

An online version of this work is published at doi.org/10.1017/9781009263177 under
a Creative Commons Open Access license CC-BY-NC 4.0 which permits re-use,
distribution and reproduction in any medium for non-commercial purposes providing
appropriate credit to the original work is given and any changes made are indicated.
To view a copy of this license visit https://creativecommons.org/licenses/by-nc/4.0
When citing this work, please include a reference to the DOI 10.1017/9781009263177

First published 2026

A catalogue record for this publication is available from the British Library

*A Cataloging-in-Publication data record for this Element is available from
the Library of Congress*

ISBN 978-1-009-86246-2 Hardback
ISBN 978-1-009-26313-9 Paperback
ISSN 2634-7334 (online)
ISSN 2634-7326 (print)

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

For EU product safety concerns, contact us at Calle de José Abascal, 56, 1°, 28003
Madrid, Spain, or email eugpsr@cambridge.org

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Digital Multilingualism and Platform Governance

Elements in Forensic Linguistics

DOI: 10.1017/9781009263177
First published online: April 2026

Janny H. C. Leung
The University of Hong Kong

Author for correspondence: Janny H. C. Leung, hiuchi@hku.hk

Abstract: As custodians of global public discourse today, transnational tech platforms govern who may speak, to whom, and how. While they have helped document and revitalize minoritized languages and connect diasporic communities, they also make language-related decisions that can disproportionately disadvantage speakers of those languages. On platforms like Facebook, non-English users navigate a linguistic environment where content moderation is often severely under-resourced compared to that available to English speakers. They may not receive warnings about disinformation or disturbing content, may not be told about what rules apply, and may have their content wrongly removed – or violating content left untouched – because neither human moderators nor automated systems can understand their language. This Element examines forms of global linguistic justice that platforms create and reproduce, highlighting a critical yet underexplored dimension of structural inequality in contemporary platform governance. This title is also available as Open Access on Cambridge Core.

Keywords: linguistic inequalities, digital language divide, platform governance, minoritized languages, digital multilingualism

© Janny H. C. Leung 2026

ISBNs: 9781009862462 (HB), 9781009263139 (PB), 9781009263177 (OC)
ISSNs: 2634-7334 (online), 2634-7326 (print)

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Contents

Series Editor's Preface 1

Preface: Internet Governance in a Linguistically Diverse World 1

Terminological Note 3

1 Navigating Global Inequalities in the Digital Age 5

2 How Language Shapes Platform Experience 20

3 Interventions and Interpretations 44

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Series Editor's Preface

The Elements in Forensic Linguistics series from Cambridge University Press publish across five main topic areas (1) investigative and forensic text analysis; (2) the study of spoken linguistic practices in legal contexts; (3) the linguistic analysis of written legal texts; (4) interdisciplinary linguistic research bridging allied fields; and (5) explorations of the origins, development, and scope of the field in various regions. Situated in the fourth category at the intersection of fields such as linguistics, law and public policy, and information science, *Digital Multilingualism and Platform Governance* by Janny Leung explores the complex question of how powerful internet platforms manage – or more accurately *mismanage* – digital multilingualism.

According to Leung, more than half of the world's eight billion people lack strong digital support for their first language. This results in a disproportionate lack of access to vital health and security information, beneficial e-commerce opportunities, and popular social media sites. These disparities for minority language speakers are exacerbated by large tech platforms' profit-first approach to the digital linguistic market. For example, digital platforms invest less in content moderation, leaving speakers vulnerable to violence caused by hate speech; they fail to provide reliable machine translation systems, leading to serious communication errors with criminal consequences; and they fail to provide multilingual access to important legal documents, leaving millions of speakers without the necessary resources to participate in their online communities legally or safely. To address these systemic problems in platform governance, Leung provides practical ideas and realistic solutions throughout her Element, emphasizing that language is the single most powerful proxy for understanding and combating global digital injustice.

Digital Multilingualism and Platform Governance is a welcome addition to similar Elements in our series that focus on digital disinformation and linguistic injustice such as *The Language of Fake News*, *Online Child Sexual Grooming Discourse*, and *Doxxing Discourse*. We encourage future submissions addressing such important interdisciplinary questions.

Tammy Gales
Series Editor

Preface: Internet Governance in a Linguistically Diverse World

When I typed 'social media is' on Google, the autocomplete suggestions included phrases like 'toxic', 'bad for mental health', and 'cancer'.[1] Social media platforms have connected people globally and become a major source of news and

[1] Search performed on 23 April 2025.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

information for many of us; however, they are also associated with extremism and polarization, abusive and hateful speech, misinformation and disinformation, violent and graphic content. As a result, the safety of individuals and groups has been threatened, the society has become ever more polarized, and political systems destabilized. Racialized populations and other minoritized groups have complained about being censored by platforms. These are, in many ways, the known problems. What American-centric literature on platform governance rarely considers is this: if English speakers sometimes find their platform experience troubling, imagine what the equivalent experience is like for speakers of other languages, and what broader impact these platforms have on their societies.

A genocide happened around 2017 against Rohingya Muslims in Myanmar, killing approximately 10,000 people and displacing more than 700,000. According to the United Nations Human Rights Council's investigation,[2] Facebook, used by many Burmese people as their primary gateway to the internet, was slow and ineffective in curbing hate speech that fuelled mass violence at the time. Despite having a code of conduct prohibiting such content, Facebook failed in moderating the inflammatory posts that proliferated. How did its content moderation system fail? Was it a 'glitch'?

Two content moderation mechanisms were available: human review and algorithmic flagging. Neither was effective. Although anti-Muslim violence fuelled by viral misinformation was a recognized challenge, Facebook employed only four Burmese-speaking content moderators in 2015, despite having 7.3 million active users in Myanmar. These moderators, based in Manila and Dublin, were far removed from the local context. Meanwhile, Facebook's automated systems struggled with the Burmese language, whose font did not transition to the global Unicode standard until 2019. In one striking example, the system translated an inciting Burmese post – 'Kill all the kalars that you see in Myanmar; none of them should be left alive' – as the innocuous English sentence: 'I shouldn't have a rainbow in Myanmar.' [3] Be it the language capacity of human reviewers or the automated systems, Facebook's struggle with Burmese was 'eminently solvable', but addressing it was just 'not a priority'.[4]

Of the many lessons that may be drawn from the atrocities in Myanmar, an elusive issue pertinent to internet governance and global inequality is: how

[2] A/HRC/39/64, 12 September 2018.

[3] *Kalar* is a derogatory term used to describe people of Indian or Arabic heritage, who are often Muslims, in Myanmar. Stecklow, S. 'Why Facebook Is Losing the War on Hate Speech in Myanmar', *Reuters*, 15 August 2018, www.reuters.com/investigates/special-report/myanmar-facebook-hate/.

[4] According to a whistleblower who was formerly the Director of Public Policy at Facebook. Wynn-Williams, S. (2025) *Careless People: A Cautionary Tale of Power, Greed, and Lost Idealism*. New York: Flatiron Books. 349–350.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

platforms manage multilingualism. The private sector has acquired outsized power in internet governance, with social media platforms not only curating the digital experience of its users but making significant impact on the societies their users live in, sometimes through seemingly mundane administrative decisions, such as what language skills they look for in content moderators they hire, which languages they include in their user interface, and which languages they translate their platform rules into.

Language 'is tied to how all kinds of resources are produced, circulated, and consumed, including how they are identified as resources at all'.[5] Language communities around the world have vastly divergent digital experiences and are impacted to a different extent when online harms are spread offline. These differences do not only come from the availability of content that has been produced in each language, but also from how platforms perform their gatekeeping role. They intersect with some of the most important challenges concerning internet governance today, including access and participation, safety and security, human rights and global inequality.

This Element explores the impact of platform governance on global inequality, focusing on language as a proxy. Section 1 introduces the current state of global digital divides and the digital linguistic landscape, illustrating the relationship between language and inequalities in the information society. Section 2 probes how transnational social media platforms serve and support a global, linguistically diverse, user base, by examining how decisions they make create divergent digital experiences for language communities. As a result, speakers of some languages enjoy much less protection from harm than others, less likelihood of having their voice heard, and fewer opportunities to engage with platform processes and governance. Section 3 articulates realistic goals of how these linguistic inequalities can be mitigated and discusses relevant levers. It places the findings in broader perspectives, including language hierarchies in the capitalist digital economy, the complexities of digital linguistic inclusion, linguistic competition in the digital field, and approaches to linguistic injustices.

Terminological Note

Platform

Throughout this Element, *platform* will be used to refer to digital services such as Facebook and YouTube. According to digital media researcher Tarleton Gillespie, platforms are online sites and services that

[5] Heller, M. and McElhinny, B. (2017) *Language, Capitalism, Colonialism*. Toronto: University of Toronto Press. 3.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

a) *host, organize, and circulate users' shared content or social interactions for them,*
b) *without having produced or commissioned that content,*
c) *built on an infrastructure, beneath that circulation of information, for processing data for customer service, advertising, and profit, and*
d) *moderate content and activity of users, using some logistics of detection, review, and enforcement.*[6]

These services are alternatively known as *digital/social/new media* or *social networking services* (SNS). The term *platform* is widely used, including by companies that operate them. Platforms are distinguished from traditional media in that they do not produce content. The term *platform* positions the companies as a neutral carrier of information, downplaying their control over the conditions under which content is produced and circulated. This is in line with their effort to claim legal immunity as intermediaries.

Global North and South

The *North–South* divide is a dominant yet contested framework for understanding global inequalities, replacing the three-world system and the developed-developing dichotomy. The United Nations Conference on Trade and Development uses the *Global North* to refer to affluent economies in Europe and North America, Japan, South Korea, Australia, New Zealand, and Israel. The *Global South* covers poorer economies in Latin America, Africa, the Middle East, Asia, and Oceania, excluding the aforementioned countries. Despite its geospatial connotation, the *North–South* divide does not correspond neatly to the hemispheres: India is considered part of the *Global South* despite being in the Northern hemisphere; Australia and New Zealand lie south of the equator yet belong to the *Global North*. The distinction is better understood as rooted in power asymmetries structured through historical colonial relations and economic dependency under global capitalism.

Critics argue that the concept oversimplifies today's complex realities, overlooking internal inequalities and marginalization within countries (the 'North in the South' and the 'South in the North'),[7] as well as the vast differences in population size, economic capacity, and geopolitical influence among *Global South* countries, which together account for two-thirds of the world's population.[8]

[6] Gillespie, T. (2018) *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. New Haven: Yale University Press. 18, 21.
[7] Thérien, J.-P. (2024) 'The North–South Distinction: From Consensus to Contestation'. *Global Studies Quarterly* 4(4),1–11, https://doi.org/10.1093/isagsq/ksae080.
[8] *Global Majority* has emerged as a term describing people living in the Global South, as well as people living in the Global North who have been racialized as ethnic minorities.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

In the digital context, the *North–South* dichotomy also fails to capture the technological rivalry between the United States and China, the latter still being sometimes classified as a developing country from the *Global South* despite its sizable global economic and technological power. The terms *Global North* and *Global South* will be adopted in this Element with awareness of their limitations.

Language Community

Language community will be used loosely to describe a group of people using the same language. Sociolinguists recognize that the concept abstracts and idealizes by glossing over the internal variation of language practices and the fuzzy boundary between members and non-members. The term, along with its variant *speech community*, has sometimes been defined using subjective instead of objective criteria, for example, whether speakers feel that they belong together, which is not the definition adopted in this Element.

Reference to *minoritized language communities* in this Element denotes speakers of languages that occupy a marginalized position nationally and globally. It therefore includes not only national linguistic minorities, whose languages typically do not enjoy official status and may or may not have developed a literary tradition, but also speakers of majority languages in the Global South whose languages may enjoy institutional support at the national level. It therefore entails heterogeneous populations that constitute more than half of the world's population.

1 Navigating Global Inequalities in the Digital Age

> The limits of my language mean the limits of my world.[9]
>
> Ludwig Wittgenstein

According to the World Inequality Database, wealth and income inequality within countries has been increasing since 1980. Despite overall global economic growth, global income inequality between countries in 2020 was at a level comparable to the peak of Western imperialism in the early twentieth century.[10] At the individual level, the top 10% wealth holders own three-quarters of global wealth, while the poorest half hold only 2% and receive only half of the income share it held in 1820. Although the world has become richer, governments have become poorer; in rich countries, governments' share of wealth is close to zero, meaning that all wealth is in private hands. The extreme gap between public and private wealth means that governments have limited capacity to tackle inequality effectively.[11]

[9] Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus*. Ogden, C. K. (trans.) (2010). London: Kegan Paul. 5.6.
[10] World Inequality Report 2026, https://wir2026.wid.world/.
[11] World Inequality Report 2026, https://wir2026.wid.world/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Among other factors, technological advancement is a key driver of the winner-take-all dynamics shaping today's economic order.[12] The digital economy is highly concentrated, with a small number of transnational corporations enjoying global market domination. Platforms exercise control over public discourse through moderating content and harness personal data for profit. Their priorities and decisions have far-reaching implications for global digital inequalities, which in turn can exacerbate income disparities and deepen the marginalization of already disadvantaged groups. As capitalism takes an informational turn and digital technologies increasingly permeate daily life, urgent questions arise: How does digitalization intersect with existing forms of social stratification? Are the inequalities it produces and reproduces inevitable or even irreversible? If language is a proxy to digital inequalities, can it also offer a point of intervention?

1.1 Global Digital Divides

Early discussion on digital inequalities was centred on the term *digital divide*. Coined to describe the connectivity gap in rural and urban America, it first appeared in newspaper articles and subsequently in a report by the National Telecommunications and Information Administration in 1995.[13] Following the rapid diffusion of the internet in the United States, scholarly interest in the concept peaked in the early 2000s, but it remains relevant in many parts of the world today. Although more people come online every year, the digital divide is still a lived reality globally. According to the International Telecommunication Union, roughly 74% of the world's population use the internet, while a quarter of the world's population are still offline as of 2025.[14] The penetration rate is only at 34% in least developed countries. Those who are not using the internet mostly live in rural areas, with the urban–rural divide being particularly prominent in low- to middle-income countries. Connectivity is also divided by gender: men are more likely to be online than women, especially in low- to middle-income countries.

While material access remains a global challenge, as the digital landscape evolves, the complex realities of digital divide have extended beyond questions of accessibility and affordability. In the Web 2.0 environment, where participatory media thrive, internet users actively generate content and interact with one

[12] Kapczynski, A. (2020) 'The Law of Informational Capitalism'. *The Yale Law Journal* 129: 1460–1515.

[13] 'Falling Through the Net: A Survey of the "Have Nots" in Rural and Urban America', National Telecommunications and Information Administration, July 1995, https://ntia.gov/page/falling-through-net-survey-have-nots-rural-and-urban-america.

[14] 'Measuring Digital Development: Facts and Figures 2025', The International Telecommunication Union. https://www.itu.int/itu-d/reports/statistics/facts-figures-2025/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

another. They go online to not only receive information but also to have a voice, form relationships, join communities, and gain access to economic opportunities. Digital inequality does not only involve access and consumption but also participation and production. With digitalization penetrating every aspect of life, the consequences of such divides are more far reaching than ever: certain populations could enjoy 'lower prices for products, greater chances of securing a job or finding a party to vote for, better education opportunities, more and better health information and treatment, opportunities to acquire new friends, and even more chance of forming romantic relationships'.[15]

Many countries in the Global South have leapfrogged in their adoption of information and communications technology, enjoying wireless internet without ever-experiencing wired internet, and using mobile devices without having had landlines and personal computers. While mobile technologies are instrumental in improving digital access, those who go online only through slower network and basic mobile devices, as is typical in the developing world, may not be able to take advantage of connectivity in the same way as those who are connected through a personal computer and high-speed internet. For one thing, mobile devices are designed more for consumption than creation, more for leisure than professional use; these tasks also require different levels of digital literacy skills.[16] Tellingly, only 1% of changes to Wikipedia are made through a mobile phone.[17] The digital divide, therefore, is not a simple dichotomy but is multi-layered,[18] encompassing device divide and usage divide.

1.1.1 Poor Internet for Poor People?

With the stated goal of bridging the global digital divide, transnational corporations have, for over a decade, offered zero-rated services to users in low-income countries, providing limited internet access for free. Though framed as philanthropic, these initiatives emerged at a time when platforms, whose stock price depends on continual user growth, were 'running out of road'.[19] Having saturated developed markets, they turned to harder-to-reach regions, where digital infrastructure was limited and new user acquisition more challenging.

These 'free' services are offered with caveats. For example, Meta collaborates with mobile operators to provide a programme called Free Basics in Asia,

[15] Van Dijk, J. (2020) *The Digital Divide*. Newark: Polity Press. 12–13.
[16] Mossberger, K., Tolbert, C. J., and Stansbury, M. (2003) *Virtual Inequality: Beyond the Digital Divide*. Washington, DC: Georgetown University Press.
[17] Napoli, P. M. and Obar, J. A. (2017) 'Second Class Netizens: Race and the Emerging Mobile Internet Underclass'. In Lind, R. A. (ed.) *Race and Gender in Electronic Media: Content, Context, Culture*. New York: Routledge. 293–311. https://doi.org/10.4324/9781315636801.
[18] Keniston, K. and Kumar, D. (eds.) (2003) *The Four Digital Divides*. Delhi: Sage.
[19] Wynn-Williams, *Careless People*.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Africa, Middle East, and South America.[20] The service allows users free and unlimited access to a selection of low bandwidth services, including Wikipedia, Google search, health-related websites, and news outlets that partnered with Facebook. Facebook Zero, which allows free access to a stripped-down version of Facebook on mobile phones, is part of a carrier's free data plan. It does not support video, audio, or other types of data-intensive traffic. Instead, multimedia content appears as redactions. Non-paying users can see links to websites on a Google search results page, the titles and descriptions of YouTube videos, or headlines of news articles shared on Facebook, without being able to read or watch them. Having to make sense of the world through clickbait headlines, which tend to mislead more than inform, these users are particularly susceptible to fake news and misinformation. They are also positioned as consumers, of mostly Western content and services, rather than creators and entrepreneurs.

Through Free Basics, Facebook therefore does not only moderate content on their own platform but also curate users' entire experience of the internet. It collects data on users' online activities, not just on Facebook. With its high penetration rate in the Global South, Facebook has been adopted not only as a social networking tool but also as an integral part of both public and private life. Indeed, Facebook has become the entirety of the online experience for many zero-rated service users. In a 2015 survey, more than half of the respondents from Brazil, India, Indonesia, and Nigeria saw Facebook as the internet.[21]

India has since banned Free Basics and made zero-rated services illegal, after antitrust concerns were raised about Facebook's gatekeeping role in this stripped-down version of the internet, which is also known as *a walled garden*. Some have even called zero-rated service a form of cultural imperialism or digital colonialism.[22] One tech entrepreneur said Facebook's approach was akin to British colonizers coming in to 'help' with tax collection and then proceeding to taking a cut.[23] On the African continent, Meta evolved its strategies and collaborated with civil society groups. Free Basics has been rolled out

[20] The cost of these free services is born by the mobile operators rather than the platforms themselves. Wynn-Williams, *Careless People*.

[21] Mirani, L. (2015) 'Millions of Facebook Users Have No Idea They're Using the Internet'. *Quartz*, 9 February, https://qz.com/333313/milliions-of-facebook-users-have-no-idea-theyre-using-the-internet.

[22] Solon, O. (2017) '"It's Digital Colonialism": How Facebook's Free Internet Service Has Failed Its Users'. *The Guardian*, 27 July. www.theguardian.com/technology/2017/jul/27/facebook-free-basics-developing-markets. See also Kwet, M. (2019) 'Digital Colonialism: US Empire and the New Imperialism in the Global South'. *Race and Class* 60(4): 3–26. https://doi.org/10.1177/0306396818823172.

[23] Bhatia, R. (2016) 'The Inside Story of Facebook's Biggest Setback'. *The Guardian*, 12 May, www.theguardian.com/technology/2016/may/12/facebook-free-basics-india-zuckerberg.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

successfully in at least thirty African countries.[24] By 2019, Free Basics was used by more than 100 million people in sixty-five countries.[25]

Responding to criticism that the programme violates net neutrality, the principle that all data traffic should be treated equally by the internet service provider, Facebook is testing a new app called Discover, which lets users freely access a low bandwidth version of any website, with a daily data cap. A study of the app's pilot in the Philippines suggests that Discover exercises editorial control in how the internet is presented to its users and makes some websites more visible than others. For instance, nearly half of the websites presented on Discover are owned by American companies. Moreover, while all video and audio content was removed, images were displayed selectively – they appeared on Facebook and Instagram but not on other sites tested. The study finds that the app's logic of redaction 'reproduces the very structural inequality that access to the internet frequently claims to ameliorate'.[26]

Whether zero-rated services truly help bridge the digital divide is contested. For many, limited internet access is still better than none. Others argue that such free services entrench existing inequalities and make it impossible for small local players to compete. There is some empirical evidence that Free Basics may not effectively reach the digitally disfranchised. A survey by the Alliance for Affordable internet of 8000 Free Basics users across eight countries found that only 12% had never used the internet before.[27] Most users relied on the app to supplement a paid plan or public WiFi. Questions also remain about how well Free Basics serves minoritized language communities, given that it operates in some of the most linguistically diverse places in the world. A review of the app in six countries – Colombia, Ghana, Kenya, Mexico, Pakistan, and the Philippines – found that the interface was available in English and one additional language, typically the national or majority language. Moreover, most services provided in the app were available in English only. This means that minoritized language communities in these countries were unlikely to benefit. Regardless of where one stands on zero-rated services, they constitute a critical piece of context for understanding global divides and social stratification that

[24] Nothias, T. (2020) 'Access Granted: Facebook's Free Basics in Africa'. *Media, Culture & Society* 42(3) (April): 329–48, https://doi.org/10.1177/0163443719890530.

[25] Nothias, 'Access Granted: Facebook's Free Basics in Africa'.

[26] Pei, L., Olgado, B. S., and Crooks, R. (2021) 'Market, Testbed, Backroom: The Redacted Internet of Facebook's Discover', in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI '21: CHI Conference on Human Factors in Computing Systems, Yokohama Japan: ACM), 1–13, https://doi.org/10.1145/3411764.3445754.

[27] Alliance for Affordable Internet. *The Impacts of Emerging Mobile Data Services in Developing Countries*. June 2016.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

information capitalism helps to perpetuate, as well as the influential role platforms play in shaping these dynamics.

1.1.2 Digital Inequalities and Intersectionality

Beyond disparities between high- and low-income countries, digital divides also intersect with existing inequalities related to gender/sexuality, race, disability, socioeconomic status, and so on, even among those who are connected to the internet. Compared with earlier works on the digital divide, a rapidly growing literature has recognized that information technologies are not a neutral resource. This shift in understanding reflects concerns about biased data being fed into automated systems ('garbage in, garbage out') and the lack of diversity among those who design and govern these technologies.

Online speech harms are not distributed evenly across different populations, especially along racial and gender lines, due to the popular adoption of artificial intelligence (AI) and the biases that machine learning reproduces. Algorithmic bias affects the life chances of marginalized populations when AI is applied to areas like banking, policing, education, and healthcare. For example, search engines can produce racially biased results,[28] automation tools can profile, surveil, and punish the poor,[29] and platform designs can reinforce gender-based violence.[30] Sensor technology often performs better on light skin than darker skin, and facial recognition software has infamously tagged black people as gorillas. Automated content moderation system has classified calling a man 'a dick' as aggressive, yet failed to flag threats of rape against women as attacks.[31]

The type of digital inequalities uncovered by these works is not fully captured by the idea of digital divide, which carries the implicit assumption that such divide can be bridged by digital solutions.[32] Instead, we are dealing with 'a form of exclusion and subordination built into the ways in which priorities are established and solutions defined in the tech industry'.[33] While technologies may not be intentionally designed to be racist, sexist, classist, and ableist, they

[28] Noble, S. U. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

[29] Eubanks, V. (2017) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, 1st ed. New York: St. Martin's Press.

[30] Chemaly, S. (2019) 'Demographics, Design, and Free Speech: How Demographics Have Produced Social Media Optimized for Abuse and the Silencing of Marginalized Voices'. In Brison, S. J. and Gelber, K. (eds.) *Free Speech in the Digital Age*. New York: Oxford University Press. 150–69. https://doi.org/10.1093/oso/9780190883591.003.0010.

[31] Chemaly, 'Demographics, Design, and Free Speech'. 150–69, at 160–64.

[32] Warschauer, M. (2003) 'Demystifying the Digital Divide'. *Scientific American*, http://education.uci.edu/uploads/7/2/7/6/72769947/ddd.pdf.

[33] Benjamin, R. (2019) *Race after Technology: Abolitionist Tools for the New Jim Code*, 1st ed. Medford, MA: Polity. 19.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

participate in perpetuating and sometimes exacerbating inequalities. These inequalities are not glitches – rather, they are 'enduring and constitutive' features.[34]

1.2 Digital Multilingualism

Language is both a tool of communication and a marker of identity. While identity carries a non-market value that resists quantification, the communicative value of the world's languages is hierarchically structured. Sociologist Abram De Swaan captures this in a global language system[35] that places the vast majority of languages – around 95% – in a peripheral, marginal position at the bottom of the pyramid. These languages have limited demographic weight and enjoy minimal institutional support. Above them are approximately a hundred central languages that are used by governments or in commerce, including languages such as Finnish and Dutch. Higher still are about a dozen supercentral languages that facilitate transnational communication, such as Arabic, Chinese, French, Hindi, Malay, Portuguese, Russian, Spanish, and Swahili. At the very top sits English, the global lingua franca of our present world and the only hypercentral language. A similar hierarchical stratification exists among varieties of a single language. For example, among global English varieties, Standard American English occupies the hypercentral position, followed by supercentral, central, and peripheral varieties.[36] While linguistic vitality online may not fully mirror offline use, it is unsurprising that supercentral and hypercentral languages dominate digital communication.

1.2.1 How Many Languages Are Online?

Estimating linguistic diversity in the world is a messy enterprise. Census data are often imprecise and out of date. Multilingual abilities are not consistently accounted for. Most fundamentally, languages exist on a dialectal spectrum rather than as distinct entities, and so any boundary drawing, any counting of 'languages', involves a level of idealization, which is more often based on political rather than linguistic differences. Despite so, a degree of strategic essentialism allows us to paint in broad strokes the state of global linguistic diversity. We know that the world's languages are distributed very unevenly among its inhabitants. Just twenty-three languages are spoken by more than half

[34] Benjamin, *Race after Technology*, at 80.

[35] De Swaan, A. (2010) 'Language Systems'. In Coupland, N. (ed.) *The Handbook of Language and Globalization*. Malden, MA: Blackwell. 56–76. https://doi.org/10.1002/9781444324068.ch2.

[36] Mair, C. (2013) 'The World System of Englishes'. *English World-Wide* 34(3): 253–78. https://doi.org/10.1075/eww.34.3.01mai.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

of the world's population, and over 88% of people speak one of the 200 largest languages as a first language. Indigenous peoples make up around 6% of the world population and speak more than 4,000 languages. The highest concentration of linguistic diversity is located in Africa, South Asia, and the Pacific.

Linguistic diversity has been on the decline following the development of modern societies.[37] Since the 1990s, UNESCO has been drawing attention to this issue by reporting on endangered languages and maintaining a world atlas of languages. According to the 28th edition of *Ethnologue,* there are 7,159 living languages worldwide.[38] Of these, only 6.9% are considered 'safe', meaning they are institutionally supported. About half are classified as 'stable', with ongoing intergenerational transmission. The remaining 43% are endangered, many of them spoken by fewer than 1,000 people.

The decline refutes simple diagnosis, as it results from a complex interplay of historical, social, economic, and political factors that vary across local contexts. Language change is influenced by patterns of language contact, community size, and social network structure.[39] Throughout history, different groups – hunter-gatherers, farmers, and herders – have spread their languages through migration and interaction. Colonization coerced language shift, as did industrialization. In the modern era, globalization and the monolingual norms tied to the rise of nation states continue to exert pressure on minoritized languages. Characteristics of modern society, such as formal schooling, urbanization, and labour mobility, are strongly associated with language decline,[40] while the relative isolation and self-sufficiency of a community tend to support language maintenance.[41]

On the internet, the great majority of languages are not exactly in decline – they have hardly left any digital footprint. Since computers were invented in the United States, early hardware and software systems were designed to meet the needs of the English language.[42] For example, the QWERTY keyboard and its variants are based on the Latin script. To perform everyday digital functions

[37] See Pavlenko for a wonderful account of how multilingualism thrived historically in human societies. Pavlenko, A. (ed.) (2023) *Multilingualism and History.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009236287.

[38] Ethnologue maintains the ISO 639 standard for language codes. See www.ethnologue.com.

[39] Trudgill, P. (2020) *Millennia of Language Change: Sociolinguistic Studies in Deep Historical Linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108769754.

[40] Bromham, L., Dinnage, R., and Skirgård, H., et al. (2021) 'Global Predictors of Language Endangerment and the Future of Linguistic Diversity'. *Nature Ecology & Evolution* 6 (2): 163–73, https://doi.org/10.1038/s41559-021-01604-y.

[41] Fishman, J. A. (1967) 'Bilingualism with and without Diglossia; Diglossia with and without Bilingualism'. *Journal of Social Issues* 23(2): 29–38. https://doi.org/10.1111/j.1540-4560.1967.tb00573.x.

[42] Graddol, D. (2000) *The Future of English? A Guide to Forecasting the Popularity of the English Language in the 21st Century.* The British Council. 30.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

such as texting and emailing, a language must undergo a process of digitalization. At a minimum, this requires an input method and, ideally, a script that is encoded in the Unicode Standard.[43] While major lingua francas quickly established a strong online presence, the growth of minoritized languages was initially constrained, especially given the dominance of written communication online. A major obstacle is that more than a third of the world's languages have not developed a written form, and many writing systems have not been standardized across communities of speakers. Even among languages that have successfully digitized their scripts, most are still classified as low-resource: they lack sufficient annotated data to support the development of digital tools such as speech recognition, machine translation, and other natural language processing applications.

Measuring digital linguistic diversity reliably is a complex challenge.[44] Digital linguistic diversity should ideally reflect user activity across the full range of online behaviours: not only browsing websites but also downloading files, playing games, instant messaging, posting on forums, and more. However, due to the vastness of online activity and the private nature of much of this data, most studies focus on a limited subset of the internet, such as specific platforms or discussion forums.[45] As a result, researchers often rely on indirect indicators. Two common measures are *user profiles* (based on the number of internet users in each language group) and *web presence* (based on the number of websites available in different languages). These metrics capture different aspects of linguistic visibility and each has its limitations. Measuring by user profiles may overestimate minority language use by assessing potential rather than actual online activity, and by assuming that users primarily operate in their native languages. Conversely, measuring web presence may underestimate minority language use, as web crawlers often miss informal and private communication channels such as emails, messaging apps, and mailing lists, where vernaculars frequently flourish and where social ties are built and maintained. Moreover, many language detection algorithms rely on written data, which risks

[43] The Unicode covers major writing systems in the world. The Script Encoding Initiative at the UC Berkeley Department of Linguistics is devoted to supporting the encoding of scripts that are still in the pipeline. See https://scriptencodinginitiative.github.io/.

[44] Gerrand, P. (2007) 'Estimating Linguistic Diversity on the Internet: A Taxonomy to Avoid Pitfalls and Paradoxes'. *Journal of Computer-Mediated Communication* 12(4) (July): 1298–1321, https://doi.org/10.1111/j.1083-6101.2007.00374.x; Prado, D. (2012) 'Language Presence in the Real World and Cyberspace'. In Vannini, L. and Le Crosnier, H. (eds.) *Net.LANG: Towards the Multilingual Cyberspace*. Caen: C&F éditions. 35–52.

[45] Wodak, R. and Wright, S. (2006) 'The European Union in Cyberspace: Multilingual Democratic Participation in a Virtual Public Sphere?' *Journal of Language and Politics* 5(2): 251–75, https://doi.org/10.1075/jlp.5.2.07wod.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

underrepresenting the digital vitality of primarily oral languages, further skewing our understanding of linguistic diversity online.

The broad trend is that the internet is becoming increasingly multilingual, according to the Observatory of Linguistic and Cultural Diversity in the Internet, which is maintained by computer scientist Daniel Pimienta. Its annual reports approximate the relative proportion of web content in 343 languages spoken by more than one million first-language speakers, based on a range of indirect indicators (such as the number of speakers connected to the internet).[46] While this method has limitations, the Observatory estimates that the share of English-language content declined from 30% in 2017 to 20% in 2025.

Assessing the digital presence or vitality of smaller languages is more difficult. A foundational study of language vitality online was conducted in 2013 by mathematical linguist András Kornai.[47] Using a machine classification model, Kornai assessed the digital vitality of the world's languages and concluded that over 96% were digitally 'dead' or effectively non-existent online. Only about 170 languages were classified as 'ascending' or having already 'ascended' into the digital realm, while approximately 140 showed signs of transitioning. Kornai's model makes a crucial distinction between passive web presence and active, interactive use of a language. For instance, extinct languages such as Classical Chinese or Latin may have substantial online resources, yet they are considered digitally impoverished because they lack active user communities engaging with the language in dynamic ways.

A group of researchers developed an updated method to assess digital language vitality by measuring digital language support. This approach classifies languages on a five-point scale: *still*, *emerging*, *ascending*, *vital*, and *thriving*, based on the extent to which they are supported by a set of 143 digital tools.[48] These tools cover a range of functions, including content availability, encoding methods, localized user interfaces, text processing, speech processing, meaning processing, and virtual assistant integration. Ethnologue has adopted this framework to track digital language vitality.[49] According to their 2023 estimates, only thirty-two languages, just 0.4% of the world's languages, are considered thriving and enjoy strong digital support. Another 3,816 languages

[46] For results, see www.obdilci.org; for method, see Pimienta, D., Blanco, Á., and de Oliveira, G. M. (2023) 'The Method behind the Unprecedented Production of Indicators of the Presence of Languages in the Internet'. *Frontiers in Research Metrics and Analytics*. 8. https://doi.org/10.3389/frma.2023.1149347.

[47] Kornai, A. (2013) 'Digital Language Death'. *PLoS ONE* 8(10): e77056, https://doi.org/10.1371/journal.pone.0077056.

[48] Simons, G. F., Thomas, A. L., and White, C. K. (2022) 'Assessing Digital Language Support on a Global Scale'. *Proceedings of the 29th International Conference on Computational Linguistics*, 4299–4305. https://aclanthology.org/2022.coling-1.379/.

[49] www.ethnologue.com/insights/digital-language-divide/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

(53%) have partial support, while 3,277 languages (46%) remain unsupported. Although 3.5 billion people speak one of the 32 digitally thriving languages, more than half of the world's 8 billion people lack strong digital support for their first language.

There are likely far more languages being used online than current estimates of digital language vitality and support suggest. The rapid growth of multimedia content, from videoconferencing to multimodal social media, has created new opportunities for oral languages to thrive in digital spaces. However, much of this content may go undetected by algorithms designed to process text-based data. As a result, the full extent of linguistic diversity online is likely underestimated.

1.2.2 How Multilingual Is User-Generated Content?

Users on major social media platforms generate content in a wide array of languages. Facebook, for example, estimates that over 160 languages[50] are in use on its platform. A large-scale study of X (formerly Twitter), analysing a 10% random sample of 118 billion tweets posted between 2009 and 2020, identified 173 content languages.[51] Interestingly, these figures align closely with András Kornai's 2013 estimate that around 170 languages had either ascended or were in the process of ascending into the digital realm, suggesting a relatively stable upper bound for the number of actively used digital languages over the past decade.

These figures likely vastly underestimate the true extent of linguistic diversity on the platforms. Using the Crúbadán web crawler, designed to look for texts in 2,200 under-resourced languages, the Indigenous Tweets Project has identified tweets in 185 indigenous and minoritized languages, most of which are neither interface languages on X nor supported by its integration with Google Translate.[52] Many of these languages have not fully adopted Unicode and would be classified as 'digitally still' under Kornai's criteria. Despite being

[50] https://ai.meta.com/blog/-xlm-r-state-of-the-art-cross-lingual-understanding-through-self-supervision/.

[51] Alshaabi, T., Dewhurst, D. R., Minot, J. R., et al., (2021) 'The Growing Amplification of Social Media: Measuring Temporal and Social Contagion Dynamics for over 150 Languages on Twitter for 2009–2020', *EPJ Data Science* 10(1): 1–28 (31 March), https://doi.org/10.1140/epjds/s13688-021-00271-0.

[52] See Scannell, K. P. (2007) 'The Crúbadán Project: Corpus Building for under-Resourced Languages'. *Cahiers Du Cental* 5(1): 5–15. The crawler works by maintaining lists of statistically unique words in each language and periodically searches texts for these words. See also Scannell, K. P. (2022) 'Managing Data from Social Media: The Indigenous Tweets Project'. In Berez-Kroeker, A. L., McDonnell, B., Koller, E., and Collister, L. B. (eds.) *The Open Handbook of Linguistic Data Management*. The MIT Press. 481–88. https://doi.org/10.7551/mitpress/12200.003.0046.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

actively used, their presence remains largely invisible. This invisibility stems in part from the limitations of off-the-shelf language identification tools (such as Langid.py and whatthelang), whose pre-trained models can typically recognize only several dozen to just over a hundred languages. The study that identified 173 languages on X, for example, relied on a classifier with a maximum capacity of 173 languages; any text outside that range was simply labelled 'unknown'.[53] This 'language recognition inequality', as sociolinguist Ana Deumert terms it, reflects a lack of investment in the development of automatic language detection systems for small and under-resourced languages.[54] Initiatives like the Indigenous Tweets Project help counteract this inequality by increasing the visibility of minoritized languages and enabling speakers to more easily locate and follow their language communities – voices that might otherwise be drowned out by dominant languages on the platform.

Technological affordances for multimedia have enabled languages that have not yet digitally ascended to participate in the digital public sphere. Even if such usage remains invisible to automated language identifiers, speakers of primarily oral languages can engage on social media platforms through audio and video formats. However, there are important caveats: users typically need some literacy in a platform's interface language to navigate it, and they must be able to afford the bandwidth required for multimodal communication.[55] According to a 2019 report published by the Pew Research Center, more than half of popular YouTube channels posted content in languages other than English.[56] Although there are no reliable estimates of how many languages are currently being used on YouTube or TikTok, both sign languages and oral languages are known to be present.

Languages detectable by language identifiers are not used evenly on social media platforms. For instance, on X, just eight languages account for approximately 80% of all tweets, with English, Japanese, Spanish, Arabic, and Portuguese being the most frequently used languages.[57] Similarly, a study employing language detection software to analyse nearly a million Reddit comments identified forty-four languages used on the platform, but non-English posts

[53] Alshaabi et al., 'The Growing Amplification of Social Media'.

[54] Deumert, A. (2020) 'The Invisibility of Linguistic Diversity Online'. In Vigouroux, C. B. and Mufwene, S. S. (eds.) *Bridging Linguistics and Economics*. Cambridge: Cambridge University Press. 81–108. https://doi.org/10.1017/9781108783101.004.

[55] Bandwidth usage differs greatly across continents. As of 2024, the average monthly traffic in Africa is 3.1GB per subscription, which is less than a quarter of the world average (13.9GB). www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-internet-traffic/.

[56] Van Kessel, P., Toor, S., and Smith, A. (2019) 'A Week in the Life of Popular YouTube Channels', www.pewresearch.org/internet/2019/07/25/popular-youtube-channels-produced-a-vast-amount-of-content-much-of-it-in-languages-other-than-english/.

[57] Alshaabi et al., 'The Growing Amplification of Social Media'.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

accounted for only 3% of the total comments, underscoring the dominance of English within the platform's user-generated content.[58]

The uneven distribution of languages on a platform is not inherently problematic, as the size of language communities naturally varies and some platforms are designed to serve specific target audiences. For instance, South Korea boasts highly successful platforms such as KakaoStory or Band (owned by Naver), which primarily cater to domestic users and Korean diasporic communities. From emoticons deeply embedded in Korean popular culture to features tailored for close-knit groups and fan interactions with K-pop idols, these platforms demonstrate a strong alignment with local culture and user preferences. These unique characteristics set them apart from global social media giants.

However, most countries lack competitive local platforms.[59] Even in South Korea, where domestic options exist, Facebook and Instagram surpassed them in popularity, illustrating the global reach and influence of these Western-based platforms. The dominance of a few global platforms can incentivise language shift and reinforce the predominance of English in online environments. While this globalizing force may put additional pressure on linguistic diversity, these platforms also allow content to be posted in any language, providing opportunities for 'writing back' and for communities to build and sustain themselves.[60]

1.2.3 Digital Language Divide

Concerns about the limited linguistic diversity on the internet have prompted significant attention from researchers and international organizations, especially during the 2000s. Notable initiatives include the United Nations' World Summit on the Information Society, which featured an action line focused on 'cultural diversity and identity, linguistic diversity and local content'. In 2003, UNESCO passed a resolution urging efforts to reduce language barriers and promote human interaction online.

The term *digital language divide* refers to the disparity in the quantity and quality of information available online across different languages. This divide is fundamentally epistemic, shaping who can access knowledge and whose knowledge is visible. While early accounts tended to describe it as an aggregate outcome of user behaviour, reflecting uneven content production across languages, this view needs updating, as platforms now play an increasingly active

[58] https://towardsdatascience.com/the-most-popular-languages-on-reddit-analyzed-with-snowflake-and-a-java-udtf-4e58c8ba473c.

[59] Notable exceptions include China, Japan, Korea, and Russia.

[60] In the sense of resisting the dominance of English, see Deumert, 'The Invisibility of Linguistic Diversity Online', at 92.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

role in mediating this divide, through their design choices, language support infrastructure, content moderation practices, and content recommendation algorithms.

The digital language divide was demonstrated by the Language Observatory Project, who looked up 100 science and engineering terms on Wikipedia and found that while all had English entries, only 30 appeared in European languages, 18 in Asian languages, and just 7 in African languages.[61] Although the internet promises to universalize access to information, the availability of scientific knowledge, library resources, up-to-date information, and culturally relevant content varies dramatically depending on language.

Consider also the stark coverage inequalities on Google Maps. In a study that queried 3 million unique places, half were discoverable with English language searches, while only around 10% appeared in results for Indonesian, Arabic, or Chinese searches. The figures were even lower for Hindi (less than 5%) and Bengali (less than 1%), despite the fact that these are among the world's most widely spoken languages.[62] This disparity highlights how speakers of these languages cannot rely on digital tools to navigate the world to the same extent as English speakers, even though their languages are officially supported by Google.

The digital language divide helps explain the phenomenon of digital diglossia, in which multilingual users draw on different parts of their linguistic repertoires for distinct digital functions, often aligned with varying levels of prestige.[63] For example, they may use a dominant language online while reserving the use of a minority language to face-to-face interactions.[64] Or they may use a minority language for socializing and entertainment, but switch to a dominant language for learning and information-seeking. Speakers of less dominant languages rely on their foreign language skills to bridge epistemological gaps; for example, in the Caucasus, English proficiency strongly predicts internet adoption and use.[65] Consequently, digital language divide may promote language shift towards dominant languages.

[61] Yoshiki, M. and Kodama, S. (2012) 'Measuring Linguistic Diversity on the Web'. In Vannini, L. and Le Crosnier, H. (eds.) *Net.LANG: Towards the Multilingual Cyberspace*. Caen: C&F éditions. 118–39. https://doi.org/10.3917/e.cf.maaya.2012.02.0118.

[62] Graham, M. and Dittus, M. (2022) *Geographies of Digital Exclusion: Data and Inequality*. London: Pluto Press, https://doi.org/10.2307/j.ctv272452n.

[63] Diglossia is a concept attributed to Charles Ferguson. Ferguson, C. (1959) 'Diglossia'. *Word* 15: 325–40. https://doi.org/10.1080/00437956.1959.11659702.

[64] Mansfield, J. B. (2014) 'Polysynthetic Sociolinguistics: The Language and Culture of Murrinh Patha Youth'. The Australian National University. https://openresearch-repository.anu.edu.au/bitstream/1885/12687/1/Mansfield%20J%20B%20Thesis%202014.pdf.

[65] Pearce, K. E. and Rice, R. E. (2014) 'The Language Divide – The Persistence of English Proficiency as a Gateway to the Internet: The Cases of Armenia, Azerbaijan, and Georgia'. *International Journal of Communication* 8: 2834–59.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

1.3 Language and Inequalities in the Information Society

Language is a well-known proxy for social inequalities. Being key to accessing resources and opportunities, it intersects with economic inequality, cultural domination, and imparity of political participation.[66] It also serves as a basis for exclusion, with the language practices of marginalized groups often becoming racialized and stigmatized.

A plethora of terms have been invoked in recent decades to describe different kinds of language-based inequalities that people experience. *Linguistic oppression* refers to gross linguistic injustice that is usually associated with a policy of forced assimilation or enforced language loss, through the prohibition or punishment of the use of a language in settings such as the classroom, or forcibly removing children from their parents and depriving them of the opportunity to learn the linguistic and cultural practices of their community, as experienced by indigenous communities in different parts of the world. This leads to what is sometimes emphatically called *linguicide* or *linguistic genocide*.[67] The agent of oppression is normally the state. In contrast, *linguistic discrimination* or *linguicism*[68] refers to everyday linguistic injustices based on the language someone speaks or other speech characteristics they have, such as the accent they speak with in a second language. Such discrimination may happen at an interpersonal level, such as between landlords and tenants, or within institutions. The interdisciplinary field of language and law, sometimes also known as forensic linguistics, deals with linguistic disadvantage, linguistic barriers to justice, injustice arising from miscommunication, mistranslation, or misinterpretation of linguistic evidence, etc.

Advancements in digital technologies have brought many positive benefits to minoritized languages and their speakers around the world. In the era of mass media, these communities rarely had control over media production, distribution, or consumption.[69] By contrast, the participatory nature of Web 2.0 has opened new possibilities for language revitalization, allowing minoritized languages to gain visibility on a global scale without the high costs traditionally

[66] Piller, I. (2016) *Linguistic Diversity and Social Justice: An Introduction to Applied Sociolinguistics*. Oxford: Oxford University Press, https://doi.org/10.1093/acprof:oso/9780199937240.001.0001.

[67] Skutnabb-Kangas, T. (2000) *Linguistic Genocide in Education – or Worldwide Diversity and Human Rights?* Mahwah, NJ: Routledge. https://doi.org/10.4324/9781410605191.

[68] Skutnabb-Kangas, T. (2015) 'Linguicism'. In Chapelle, C. A. (ed.) *The Encyclopedia of Applied Linguistics*, 1st ed. Chichester, West Sussex: Wiley, 1–6, https://doi.org/10.1002/9781405198431.wbeal1460.

[69] Jones, E. H. G. (2013) 'Minority Language Media, Convergence Culture and the Indices of Linguistic Vitality'. In Jones, E. H. G. and Uribe-Jongbloed, E. (eds.) *Social Media and Minority Languages: Convergence and the Creative Industries*. Bristol: Multilingual Matters, 58–72, https://doi.org/10.21832/9781847699053-006.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

associated with content creation and dissemination.[70] Multimodal affordances, such as audio and video, cater to the cultural needs of communities that have not adopted a written tradition, enabling them to engage in digital communication in ways that reflect their linguistic practices. Mobile technologies have also helped overcome geographic isolation and structural disadvantage, supporting community cohesion and self-determination.[71] Social media now plays a vital role in language learning, maintenance, and revitalization, especially by connecting speakers across diasporas. These technological affordances have contributed to increased online linguistic diversity and are sometimes seen as a force for linguistic justice. New technologies may save old languages, or so the narrative goes.

However, linguistic inequalities still plague the digital sphere. The issue extends beyond mere access to technology to encompass how these technologies are designed and deployed. In fact, the concentration of private powers in Web 2.0 and their mediation and commodification of our public and private discourse might have exacerbated global inequalities. The next section will demonstrate how dominant technology platforms make language-related decisions that profoundly affect people's lives, reaching far beyond the conventional understanding of the digital language divide. To fully grasp the disparate impacts social media platforms have on diverse language users and communities, this study examines not only technological barriers and biases but also corporate policies and practices that produce unequal treatment. In fact, an excessive focus on technology and the belief that technology is always the solution have distracted us from more immediate remedies grounded in policy setting and equitable resource allocation.

2 How Language Shapes Platform Experience

> An adequate science of discourse must establish the laws which determine who (*de facto* and *de jure*) may speak, to whom, and how.[72]
>
> Pierre Bourdieu

The early promise of the internet as the land of freedom and an equalizing playing field for all has been replaced by pessimism about how it facilitates censorship, surveillance, and excessive commodification. Between these two extremes is the realization that information technologies are highly

[70] Jones, 'Minority Language Media, Convergence Culture and the Indices of Linguistic Vitality'; Deumert, 'The Invisibility of Linguistic Diversity Online', at 89.

[71] Dyson, L. E., Grant, S., and Hendriks, M. (eds.) (2015) *Indigenous People and Mobile Technologies*. New York: Routledge. https://doi.org/10.4324/9781315759364.

[72] Bourdieu, P. (1977) 'The Economics of Linguistic Exchanges'. *Social Science Information* 16(6): 645–68. 648. https://doi.org/10.1177/0539018477016006.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

configurable, which invites analyses as to how these technologies are deployed within particular business models to generate profit and produce power.[73]

This section explores how users' platform experiences vary depending on the language(s) they speak, shaped not only by technological design but also by corporate policies and operational processes. There is no claim here that platforms maintain coherent or explicit language policies. Rather, language-related decisions inevitably arise across multiple facets of platform operation: from user interface design and content moderation to content recommendation, research and development, and human resource allocation.

Save for my personal communications with some relevant organizations, data presented below are obtained from the public domain. These include information aggregated from platform interfaces, data voluntarily shared by platforms (such as user policies and transparency reports), whistleblower disclosures from former employees, media coverage, publications by NGOs and think tanks, academic research, and other publicly accessible materials.

The primary focus of analysis is Meta, the parent company of Facebook, Instagram, and Threads, with other platforms discussed where relevant. Despite an ageing user base, Facebook remains highly significant: as of 2025, it boasts over 3 billion active users and is the leading social network in 157 out of 167 countries.[74] Its Free Basics programme, which provides free access to a selection of data-light services, gives Facebook an outsized influence in developing countries, positioning it as a key player in either bridging or exacerbating global inequalities. To its credit, Meta has established an independent body called the Oversight Board to improve its public accountability in content moderation. The Board, which handles final appeals of content moderation decisions, comprises over twenty distinguished individuals from around the world, including lawyers, journalists, human rights activists, and law professors. Their opinions provide rare insights into the platform's internal operations, usually inaccessible to outsiders.[75]

2.1 Accessing Platforms: Interface Languages

Efforts to localize interface design vary significantly across platforms. At the time of writing, Facebook's website setting contains 114 language options, and X is available in 48. Reddit is available in eleven languages, all of which are

[73] Cohen, J. E. (2019) *Between Truth and Power: The Legal Constructions of Information Capitalism*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780190246693.001.0001.

[74] According to the World Map of Social Networks, as of 2023.

[75] Leung, J. (2022) 'Shortcuts and Shortfalls in Meta's Content Moderation Practices: A Glimpse from its Oversight Board's First Year of Operation'. *Comparative Law and Language* 1(2): 55–73. https://doi.org/10.15168/cll.v1i2.2365.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

European. Although the number of interface languages might be one indicator of a platform's linguistic inclusivity, the reality is more complex and nuanced. For one thing, the list of interface languages often includes closely related linguistic varieties, such as Spanish (Spain) and Spanish (Mexico). Moreover, although all interface languages are presented as fully equivalent options, their usability varies, with some language versions containing significant content gaps. For instance, if a user selects Inuktitut or Hausa as their interface language, Facebook (as of 2025) prompts them to choose a backup language from a list of fifty options for situations where the preferred language is not supported.

Across platforms, from search engines to social networking sites, a common strategy for localizing interfaces is to crowdsource translation services from volunteers, enabling companies to expand their markets worldwide.[76] Facebook localized its platform primarily through community efforts, which may be considered an example of the gift economy and digital free labour.[77] In 2008, it launched an application called 'Translate Facebook', allowing community translators to submit their work in languages chosen by Facebook, which was then reviewed and voted on by other volunteers. Although any user can volunteer, they must translate from and navigate the application in US English. The first two years of the translation application brought 108 interface languages to Facebook.[78] However, participation in community translation dwindled over time and the application was retired in 2022. This decline highlights the limitations of unwaged digital labour, as localization efforts depend heavily on volunteers' motivation, as well as their access to technology, language skills, and leisure time.

Platforms' investment in interface languages is clearly shaped by market considerations. Languages selected tend to be widely spoken or serve as the dominant language of a given country. The largest transnational platforms have gone beyond national and official languages to include some regional and even artificial ones, such as Frisian or Esperanto. However, decisions about which languages to include, or exclude, can be politically charged. In December 2022, Signal, an encrypted messaging app popular among political activists, disabled its Uyghur and Cantonese interfaces in favour of using 'the official language of a given region'.[79] The decision drew criticism from activists who pointed out that

[76] Osborn, D. (2010) *African Languages in a Digital Age: Challenges and Opportunities for Indigenous Language Computing*. Ottawa: International Development Research Centre.

[77] Terranova, T. (2000) 'Free Labor: Producing Culture for the Digital Economy'. *Social Text* 63, 18(2): 33–58. https://doi.org/10.1215/01642472-18-2_63-33.

[78] Lenihan, A. (2011) '"Join Our Community of Translators": Language Ideologies and/in Facebook'. In Thurlow, C. and Mroczek, K. (eds.) *Digital Discourse: Language in the New Media*. New York: Oxford University Press. 48–64, https://doi.org/10.1093/acprof:oso/9780199795437.001.0001.

[79] According to Meredith Whittaker, President of the Signal app. https://twitter.com/mer__edith/status/1606329476597649410.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Uyghur and Cantonese are the primary languages used in Xinjiang and Hong Kong – regions where China's human rights practices have faced international scrutiny. Signal ultimately reversed the change after facing public pressure.

Platforms exhibit a monolingual ideology in the language options they provide. Users are typically required to select a single interface language, while platforms privilege standardized orthographies and disregard variation. For example, although K'ichee speakers in Guatemala regularly code-mix with Spanish, Facebook reinforces rigid linguistic boundaries by releasing a version of its app in pure K'ichee, avoiding even nativized Spanish loanwords.[80]

When a user's preferred language is not supported as an interface option, they can still access and contribute content, as long as they have sufficient proficiency in another language to navigate the platform. Content can be created in any language that has a digital input method, so across all major platforms, the number of content languages far exceeds the number of interface languages. For example, while YouTube has localized its interface into 83 languages, it anticipates content in as many as 230 languages, as indicated by its video language settings. In practice, many multilingual users are so accustomed to navigating the internet in a dominant language that they may not even switch to a local language interface when it is available.

For example, a survey found that Zimbabwean users of Google often do not select African language options on the webpage, even when they are aware of their availability.[81] One explanation for this behaviour lies in prevailing language ideologies: English is perceived to have more functional power than local languages when it comes to accessing information.[82] By contrast, Zimbabweans do use African languages on social media, reflecting a broader pattern in which English is preferred for instrumental purposes, while local languages are reserved for socializing. Similar trends have been observed elsewhere. A 2011 study found that although nearly all Tunisian Facebook users surveyed had selected the French interface, more than half used Arabic to communicate with others on the platform.[83] Likewise, in Egypt, while users were evenly split between Arabic and English interface preferences, roughly

[80] Romero, S. (2016) 'Bill Gates Speaks K'ichee'! The Corporatization of Linguistic Revitalization in Guatemala'. *Language & Communication* 47: 154–66. https://doi.org/10.1016/j.langcom.2015.08.001.

[81] Mpofu, P. and Salawu, A. (2020) 'African Language Use in the Digital Public Sphere: Functionality of the Localised Google Webpage in Zimbabwe'. *South African Journal of African Languages* 40(1): 76–84, https://doi.org/10.1080/02572117.2020.1733833.

[82] Mpofu, and Salawu, 'African Language Use in the Digital Public Sphere'. 76–84.

[83] El Zaim, A. (2012) 'Cyberactivism and Regional Languages in the 2011 Arab Spring'. In Vannini, L. and Harvé Le Crosnier, H. (eds.) *Net.LANG: Towards the Multilingual Cyberspace*. Caen: C&F éditions. 325–36. https://doi.org/10.3917/e.cf.maaya.2012.02.0324.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

three-quarters of them used Arabic in their interactions.[84] In addition to language ideologies, users' preferences may also reflect the quality and completeness of localization work, which can affect the usability and appeal of the local-language interface.[85]

Sometimes, digital diglossia arises out of necessity, as minoritized language speakers have no choice but to acquire some level of foreign language proficiency in order to engage with digital technologies. For example, kigay, an Aboriginal community in northern Australia, developed functionally specific digital literacy – learning just enough English to navigate menu options and file names on mobile devices, but not enough to read or write extended English texts.[86] This has significant implications: users may accept terms and conditions in English without fully understanding what they are agreeing to, and may struggle to comprehend platform rules, policies, or enforcement mechanisms.

2.2 Availability and Quality of Machine Translation

Major social media platforms have developed machine translation capability to help users engage with content originally shared in other languages. Unlike a public service or a utility website, where users typically choose one preferred language for all website functions, social media companies are financially motivated to accommodate multilingual users across their full linguistic repertoires, as doing so increases user engagement and time spent on the platform. Both Facebook and X, for instance, allow users to select interface and content languages separately. Facebook enables automatic translation of posts from any of ninety-two languages into a user's chosen language. On X, users can click a 'translate post' button to view content in their preferred language; this function relies on the Google Translate API, which supports 249 languages.[87] X also lets users specify, from a list of sixty-seven languages, which languages they would like to see in recommended posts. YouTube offers caption translation in 125 languages, while TikTok currently provides automatic captions and translation in nine.[88]

Of the approximately 160 languages known to be used on Facebook, its machine translation capacity covers fewer than 60% of them. Machine translation models consistently perform well only for high-resource languages that

[84] El Zaim, 'Cyberactivism and Regional Languages in the 2011 Arab Spring'. 325–36.
[85] Deumert, A. (2014) 'Sites of Struggle and Possibility in Cyberspace. Wikipedia and Facebook in Africa'. In Androutsopoulos, J. (ed.) *Mediatization and Sociolinguistic Change*. Berlin: De Gruyter. 487–514. https://doi.org/10.1515/9783110346831.487.
[86] Mansfield, 'Polysynthetic Sociolinguistics'.
[87] As of April 2025. https://cloud.google.com/translate/docs/languages
[88] https://insense.pro/blog/tiktok-translation

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

have a vast bilingual corpus. The great majority of the world's languages, which lack tens of millions of parallel sentences, are considered resource-poor.[89] In 2022, Meta launched its 'No Language Left Behind' project, claiming to have developed a model capable of translating among 200 languages by using AI to construct bilingual datasets from monolingual data, thereby addressing the data scarcity for low-resource languages.[90] Google has also expressed the ambition to extend its machine translation coverage to the next thousand languages.[91] Although the development of tools like Google's Transformer and OpenAI's GPT (Generative Pre-trained Transformer), based on large language models that are trained using internet content, promises major breakthroughs in machine translation, the performance gap across languages remains substantial, especially for low-resource languages with non-Latin scripts.[92]

Machine translation reduces the cost of accessing information and enables instant cross-cultural communication. It helps break down communication barriers in social interactions and tourist encounters. However, its limited reliability in low-resource languages poses risks, particularly in high-stakes situations such as seeking healthcare advice or receiving emergency alerts in politically volatile regions.

Content moderation on platforms often relies partially on machine translation. The accuracy of these translations can significantly impact moderation outcomes, as demonstrated in cases reviewed by the Oversight Board. In two instances involving hate speech, the Board overturned Facebook's decisions to remove content when discrepancies were found between Facebook's English translation and that of the Board's translators.[93] Erroneous translations can lead to unnecessary content removal, account suspensions, or other serious consequences. In 2017, Israeli police arrested a Palestinian man for posting 'يصبحهم' ('yusbihuhum', meaning 'good morning') on Facebook. The platform's machine translation rendered the phrase as 'hurt them' in English or 'attack them' in Hebrew, leading to the arrest without an Arabic-speaking officer verifying the original post.[94] This incident highlights how the questionable

[89] Haddow, B., Bawden, R., and Barone, A. V. M., . et al. (2022) 'Survey of Low-Resource Machine Translation'. *Computational Linguistics* 48, (3): 673–732, https://doi.org/10.1162/coli_a_00446.

[90] NLLB Team, Costa-jussà, M. R., Cross, J., et al. (2022) 'No Language Left Behind: Scaling Human-Centered Machine Translation', 6 July, https://arxiv.org/abs/2207.04672.

[91] Bapna, A., Caswell, I. and Kreutzer, J., et al. (2022) 'Building Machine Translation Systems for the Next Thousand Languages', http://arxiv.org/abs/2205.03983.

[92] Ahuja, K., Diddee, H., Hada, R., et al. (2023) 'MEGA: Multilingual Evaluation of Generative AI', https://doi.org/10.48550/arXiv.2303.12528.

[93] 2020-002-FB-UA and 2021-007-FB-UA.

[94] Hern, A. (2017) 'Facebook Translates 'Good Morning' into 'Attack Them', Leading to Arrest'. *The Guardian*. 24 October. www.theguardian.com/technology/2017/oct/24/facebook-palestine-israel-translates-good-morning-attack-them-arrest.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

quality of machine translation is often not apparent to users, including law enforcement, because language options are presented as equivalent, and no warnings are provided for machine-translated texts, especially in low-resource languages.

Beyond the risk of inaccurate translation, machine translation can also perpetuate stereotypes and prejudices. Since output is generated based on the probability of word sequences from the training data, biased patterns present in the source texts can be reflected in the translations. For example, when input texts contain gender stereotypes, machine translation may reproduce these biases, such as translating a gender-neutral phrase into English with the pronoun 'he' for a doctor, reinforcing gender stereotypes.[95]

The expansion of machine translation could have long-term effects on language communities. Conceivably, being able to access foreign language content in one's first language can disincentivize language shift towards dominant languages. However, some critics worry that the replacement of foreign language learning with commercial, 'industrially prosthetic' multilingualism ultimately encourages individual monolingualism, as users might feel less compelled to learn foreign languages when they can depend on machine translation.[96]

2.3 Interacting with Platforms: Warnings and Notifications

Platforms often do not consistently communicate with users in their chosen interface language, which is especially problematic for warning labels and other important notifications. This issue hampers efforts to combat misinformation, as many social media sites apply labels to content that may be false, misleading, disturbing, sensitive, or dangerous.

A study comparing Facebook, X (formerly Twitter), YouTube, and TikTok found that for users browsing the platforms in nine of the most widely used languages – Spanish, Indonesian, Portuguese, Hindi, Chinese, Arabic, French, Russian, and Bengali, respectively – content labels are not consistently translated into users' preferred languages when they shared the same English content that has been flagged as containing misinformation.[97] TikTok performed best, translating the labels across all nine languages. Facebook translated some

[95] Farkas, A. and Németh, R. (2022) 'How to Measure Gender Bias in Machine Translation: Real-World Oriented Machine Translators, Multiple Reference Points'. *Social Sciences & Humanities Open* 5 (1): 100239, https://doi.org/10.1016/j.ssaho.2021.100239.

[96] Gramling, D. (2021) *The Invention of Multilingualism*, 1st ed. Cambridge: Cambridge University Press. 7. https://doi.org/10.1017/9781108780667.

[97] Bradshaw, S. and McCain, M. (2022) 'Lost In Translation: Language Gaps in Social Media Labels'. *Lawfare* (blog), 4 February, www.lawfaremedia.org/article/lost-translation-language-gaps-social-media-labels.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

labels, but only partially in Spanish, Russian, and Bengali. X offered some translated labels but displayed others in English. Most concerningly, YouTube's untranslated labels were not shown in English either – they were simply hidden, meaning that users received no warning at all about the nature of the content.

One might assume that if misinformation is shared in English, then providing content labels in English would suffice. However, this overlooks the fact that English-language misinformation is often shared by and circulated among non-English speakers. In fact, limited English proficiency might affect one's ability to assess the credibility of a post, heightening the need for warning labels in the user's preferred language.

For instance, although Spanish is officially supported by X, a study of the 2020 US election found that warning labels on misinformation were not automatically translated into Spanish, even when users had set it as their default interface language. The research team notes that this 'follows a broader trend observed throughout the election season, in which non-English language policy enforcement fell distinctively behind even when the narrative themselves were the same across languages'.[98] They further noted that fact-checking measures were significantly less effective for Spanish- and Chinese-speaking communities compared to English-speaking users.

A different study that examined over 100 pieces of coronavirus-related misinformation content posted on Facebook in six different languages has similarly found that speakers of some languages are at a greater risk of exposure to misinformation than others. These include English and five supercentral languages: Arabic, French, Italian, Portuguese, and Spanish. The study suggests that the platform's effectiveness in issuing warning to users about the quality of content does not only hinge on its machine translation capacity but also on its ability to detect violating content in different languages. Even for content flagged by third-party fact checkers, Facebook struggled to apply warning labels and reduce visibility. This issue was evident for English-language posts, where 29% of verified misinformation lacked warning labels. The failure was even more pronounced in other languages: 68% of such content in Italian and 70% in Spanish went unlabelled.[99] These lapses meant that false claims, such as using hairdryers or high doses of vitamin C for coronavirus prevention,

[98] Center for an Informed Public, Digital Forensic Research Lab, Graphika, and Stanford Internet Observatory, *The Long Fuse: Misinformation and the 2020 Election* (Stanford Digital Repository: The Election Integrity Partnership, 2021), 106, https://purl.stanford.edu/tr171zs0069.

[99] Avaaz, 'How Facebook Can Flatten the Curve of the Coronavirus Infodemic' (AVAAZ, 15 April 2020).

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

reached tens of millions of Facebook users without any indication that the content was misleading or dangerous.

Taken together, these studies show that even for widely used languages, content detection and user notifications do not function nearly as effectively as they do in English. Since machine learning systems perform best on resource-rich languages, the situation is likely far worse for less commonly used ones. The disproportionate risk of exposure to misinformation that non-English language speakers face can have direct impact on their health, safety, and wellbeing. Misinformation also has cumulative effects on the society by affecting elections, fuelling conspiracy theories, inciting inter-group conflicts, and undermining trust in institutions, experts, and the media.

Another type of notification that users receive informs them that their post has been removed for violating platform policy. The Oversight Board has explicitly recommended that Facebook notify users of the specific rules their content has breached. This guidance stemmed from concern over Facebook's vague and inconsistent application of its Hate Speech Community Standard across multiple cases.[100] Yet as of 2021, Facebook had implemented detailed violation notices only for English-language users, with vague assurances that the feature would be extended to other languages 'in the future'.[101] In the meantime, non-English speakers are often left in the dark about why their content was removed. Some may assume that their content was removed because the platform disagreed with their views.

2.4 Access to Platform Rules

Platforms usually have an external facing set of content rules they make available to its users, and a much more elaborate set of rules available only internally. Facebook's external facing rules, called Community Standards, are published in various language versions, but they are not created to be equal: the authoritative version is in US English. Its Transparency Center notes that 'the US English version of the Community Standards reflects the most up-to-date set of the policies and should be used as the primary document'.[102] Inconsistencies across language versions have been noted;[103] in one case, the English version of the Standards was already confusing, and its translation into Urdu became

[100] Such as 2020-003-FB-UA and 2021-002-FB-UA.

[101] 'Facebook Q1 2021 Quarterly Update on the Oversight Board', n.d., https://about.fb.com/wp-content/uploads/2021/07/Facebook-Q1-2021-Quarterly-Update-on-the-Oversight-Board.pdf. In 2023, Meta said such notifications were being tested in Arabic, Spanish, and Portuguese (2023-007-FB-UA, 2023-008-FB-UA; 2023-09-IG-UA). Implementation remained partial as of 2024.

[102] https://transparency.fb.com/en-gb/policies/community-standards/.

[103] Such as in 2022-005-FB-UA.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

incomprehensible.[104] More fundamentally, the Community Standards are not available in all the languages Facebook otherwise claims to support. Even where translations exist, updates could be slow. The Oversight Board has observed, for example, that updated versions in Urdu and UK English were still unavailable five months after the US English version had changed.[105]

As of 2025, Community Standards on Facebook were published in ninety-seven languages – about 85% of the languages the platform officially supports and 60% of the content languages known to the platform. This leaves a substantial number of users without easy access to the rules that govern their behaviour on the site. Users who cannot read the Community Standards in their preferred language receive no clear notice of what is permitted or prohibited. Their content may be removed or their accounts suspended without their understanding of what rule was broken. Nor can they easily determine whether they have grounds to report content that may be harmful or offensive to them. For example, Facebook has not translated its Community Standards into Pashto or Dari, the two official languages in Afghanistan, leaving the platform's five million Afghan users unable to access the rules in a language they speak. Considering the importance of these platforms for social networking, public discourse, and political expression, especially in countries with limited independent media, inadequate access to platform rules has significant implications for free speech and due process.

The Oversight Board has flagged the lack of translated Community Standards in several of its decisions. In 2021-003-FB-UA, for example, the Board expressed concern that Community Standards were not available in Punjabi, a language with 118 million speakers globally, including 30 million in India, and one of the fifteen most widely spoken languages in the world. Although the Board recommended that Community Standards be translated into Punjabi,[106] it used notably softer language when addressing other language groups in similar circumstances. The Board told Facebook that it 'should also aim to make its Community Standards accessible in all languages widely spoken by its users'. While it rightly noted that this would 'allow a full understanding of the rules' and help users 'engage with Facebook over content that may violate their

[104] In 2024-031-FB-MR, the Board found the use of the term 'outing-risk group' in the religious context odd in the Coordinating Harm and Promoting Crime Policy.

[105] 2022-005-FB-UA.

[106] Douek calls this translation gap a 'fairly obvious' problem, and the Board's recommendation to translate the Community Standards into Punjabi as something that 'hardly requires expertise in freedom of expression to recognize as important'. In her critique of the Board's impact, she suggests that this kind of problems 'speaks more to the Meta's negligence rather than the Board's vigilance'. Douek, E. (2024) 'The Meta Oversight Board and the Empty Promise of Legitimacy'. *Harvard Journal of Law & Technology* 37(2): 373–445.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

rights', it failed to define what qualifies as 'widely spoken'. The vagueness is concerning, given the Board's own observation that the absence of Community Standards in Punjabi means that its speakers' inability to access platform rules may have effects that 'raise human rights concerns'. Presumably the same human rights concerns apply to other minoritized language speakers. Moreover, these language gaps are systemic rather than mere oversight. The closing of these gaps therefore requires a structured and comprehensive approach, not piecemeal recommendations on a case-by-case basis.

2.5 Multilingual Capacity in Content Moderation

Content moderation on social media platforms is governed by a combination of local laws and the platform's internal rules, with the latter typically encompassing a broader scope. For instance, Meta restricts speech that involves violent or criminal behaviour, poses safety concerns, or is deemed objectionable or inauthentic. Moderation activities can occur either before content is published, known as ex ante moderation, or after it has been shared, known as ex post moderation. Given the volume of content that is shared on social media every day,[107] this moderation work relies heavily on automation. Users can also flag content they find objectionable, prompting platforms to review these reports. Human moderators are employed to identify and remove violating content and to review flagged material. Meta, for example, has 15,000 human content moderators who speak about fifty languages and work in more than twenty outsourced sites worldwide.[108] Their automated tools can detect hate speech in about thirty languages and terrorist-related messages in nineteen languages. Despite being an industry leader, Meta's capacity, both human and technological, is insufficient to handle moderation across all content languages. This leads to a situation where multilingual users may experience significantly different speech environments on the same platform, with content being more effectively moderated in certain languages than others.

Speakers of low-resource languages are at a greater risk of encountering more abusive speech or disinformation when platform rule enforcement in their language is inaccurate, delayed, or entirely absent.[109] This lower enforcement accuracy also increases the likelihood that speakers of these languages will be subjected to erroneous censorship, where legitimate content is wrongly

[107] According to one estimate, 4 petabytes of data are generated on Facebook alone each day in 2023. See https://edgedelta.com/company/blog/how-much-data-is-created-per-day.

[108] Barrett, P. M. (2020) *Who Moderates the Social Media Giants? A Call to End Outsourcing*. New York: NYU Stern Center for Business and Human Rights.

[109] For example, the Oversight Board notes significant delay in Meta's moderation of content in Haiti due to lack of investment. 2023-21-FB-MR.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

removed or suppressed. Currently, platforms do not publish data on the accuracy of content moderation as a factor of language. The Oversight Board has recommended that Meta include error rates broken down by country and language for each of its Community Standards in transparency reports.[110] However, implementation of this recommendation from 2021 is still "in progress" as of February 2026.[111]

The lack of effective content moderation in certain languages has led to dire consequences. In 2021, when Meta CEO Mark Zuckerberg appeared before the United States House of Representatives, Congressman Tony Cárdenas questioned him about Facebook's ability to handle misinformation in languages other than English.[112] In particular, he highlighted the problem of Spanish-language disinformation in the United States, which has led to anti-vaccine conspiracies, voter suppression, and hate crimes. During the 2020 election campaign, for example, the Trump campaign ran Spanish-language ads that falsely claimed Joe Biden was endorsed by Venezuelan President Nicolás Maduro. This underscores how gaps in moderation across different languages can have harmful impacts on democratic processes.

As these cases illustrate, harm spawned on platforms can easily lead to real world harm, affecting both platform users and non-users. To help mitigate risks, Meta developed a registry of at-risk locations that helps guide its resource allocation in content moderation. Having deemed Ethiopia as a Tier 1, or the highest, at-risk country since late 2020, Meta developed language classifiers, machine learning tools trained to automatically detect potential violations of the Community Standards, in Amharic and Oromo, two of the most widely used languages in Ethiopia. Despite such efforts, an Amnesty International report found Meta too slow to respond to harmful content. Between 2020 and 2022, armed conflict in the Tigray region resulted in the death of up to 600,000 civilians. Meta currently faces a $1.6 billion lawsuit for allegedly fuelling the Tigray War, which cited language as a key source of inadequate moderation, as the company could only review content in four out of the eighty-five languages spoken in Ethiopia.[113]

Whether it is done by human or machine, content moderation tends to suffer from pragmatic deficiency, which is often considered a trade-off against efficiency, responsiveness, and scalability.[114] Machine learning systems have a

[110] In 2021-003-FB-UA.

[111] https://transparency.meta.com/oversight/oversight-board-recommendations/

[112] See transcript at www.congress.gov/event/117th-congress/house-event/111407/text.

[113] 'Ethiopia: "A Death Sentence for My Father": Meta's Contribution to Human Rights Abuses in Northern Ethiopia' (Amnesty International, 30 October 2023).

[114] Leung, J. (2022) 'Shortcuts and Shortfalls in Meta's Content Moderation Practices: A Glimpse from Its Oversight Board's First Year of Operation'. *Comparative Law and Language* 1(2): 55–73, https://doi.org/10.15168/cll.v1i2.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

limited ability to interpret context and are often insensitive to linguistic variation and cultural particularities. Human moderators, meanwhile, operate under significant time pressures and typically have restricted access to broader speech context, leading them to rely on scalable rule-based systems.[115] These rules, while efficient, are prone to generating false positives – incorrectly flagging or removing content. For example, Instagram's DeepText identified 'Mexican' as a slur because of it frequently co-occurred with the word 'illegal' in the dataset. Automated content moderation systems are built upon datasets that contain human biases, which most likely disadvantage vulnerable groups, according to former United Nations Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression David Kaye.[116]

The lack of linguistic and cultural sensitivity is a built-in feature of Meta's current system of content moderation. Meta, for example, assign 'language-agnostic' reviewers to moderate content through machine translation for languages where the company lacks human or algorithmic expertise but are supported by its machine translation capacities. These reviewers may not be familiar with the region where the content was flagged. Language-specific reviews are not supported when Meta deems that the demand for content moderation in a language to be low, but no detailed criteria have been published. In a telling example, three human moderators decided that a video portraying violence against gay men in Nigeria did not violate Community Standards.[117] Although the video caption was in English, all the verbal exchanges in the video were conducted in Igbo, a major language spoken by 33 million people in Nigeria. It turned out that the human content reviewers might not have been familiar with Igbo. Meta explained that its automated systems identified the language as English because of the caption, and routed the case to English-speaking reviewers. Even if Igbo had been correctly identified, Meta's policy does not support large-scale content moderation in Igbo for Nigeria. Instead, content in Igbo is processed through machine translation by language-agnostic reviewers unfamiliar with Nigeria's cultural context. It takes contextual knowledge to appreciate the danger that the video poses to its victims: same-sex relationships are criminalized in Nigeria, and hostility towards homosexuals is widespread. When the case was appealed, Meta again misidentified the language, this time as

[115] According to one estimate, content moderators on average have under 150 seconds to make a decision. Koetsier, J. (2020) 'Report: Facebook Makes 300,000 Content Moderation Mistakes Every Day', *Forbes*, 9 June, www.forbes.com/sites/johnkoetsier/2020/06/09/300000-facebook-content-moderation-mistakes-daily-report-says/.

[116] 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression' (United Nations General Assembly, 29 August 2018), 9.

[117] 2024-041-FB-UA

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Kiswahili, further illustrating the system's deficiencies. The video eventually went viral, amassing over 3.6 million views over five months before the Oversight Board intervened and ordered its removal.

Context sensitivity also depends on how platforms categorize markets and allocate resources. In one case, a Cuban woman's protest speech was flagged as hate speech by both Meta's automated 'hostile speech' classifier and its human reviewers. While her words may have appeared dehumanizing when read literally, the Oversight Board found that they should have been permitted when understood as a call to action within the context of political protest.[118] Meta explained that Spanish content from Spain receives country-specific review; content from Venezuela, Honduras, and Nicaragua belong to a single regional queue. In contrast, content from Cuba is lumped into Spanish-language content generally because it is not considered a 'separate market'. As a result, reviewers may lack the cultural and contextual expertise specific to Cuba. A similar issue arises with Arabic: Meta has developed a content filter targeting bullying and harassment for 'general language Arabic', even though Arabic is spoken across a vast geographical area from Middle East to North Africa and regional Arabic varieties are not mutually intelligible. It is doubtful whether Meta's enforcement tools are sufficiently attuned to these dialectal differences.[119]

Content moderation processes have been shown to exhibit biases that disproportionately affect minoritized communities. A human rights impact assessment conducted in response to the May 2021 crisis in Israel and Palestine revealed significant disparities in how Meta's moderation systems handled content in different languages. Specifically, there was over-enforcement in Arabic content, meaning a higher rate of erroneous removal of Palestinian voices, and under-enforcement in Hebrew content, allowing more violating content in Hebrew to remain undetected. After adjusting for population size, these disparities indicate that Arabic speakers experienced more frequent misclassification and censorship, raising concerns about fairness and discrimination.[120] While Meta has developed an Arabic hostile speech classifier to address these issues, there is no equivalent tool for Hebrew. The assessment notes that such differential impacts on Arabic speakers have implications for the human rights principle of non-discrimination.

[118] 2023-014-IG-UA [119] 2023-006-FB-UA; 2023-039-FB-UA

[120] *Human Rights Due Diligence of Meta's Impacts in Israel and Palestine in May 2021*, Business for Social Responsibility, September 2022, www.bsr.org/reports/BSR_Meta_Human_Rights_Israel_Palestine_English.pdf.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

A platform's internal processes significantly influence the consistency and quality of content moderation across different languages. For example, Facebook's Internal Implementation Standards, that is, the internal guidelines that are used to enforce Community Standards, are provided in English, even for moderators who work in other languages.[121] This reliance on English-language guidelines could increase the likelihood of enforcement errors and raise human rights concerns for minoritized populations.[122] Meta has declined to translate the internal guidelines, asserting that its moderators are fluent in English.[123]

The uneven distribution of content moderation resources across languages reflects broader issues of market prioritization and resource allocation. Notably, the Oversight Board has raised concerns about the limited number of Urdu reviewers – fewer than fifty as of mid 2022 – despite the significant size of the Indian market, which has the highest number of Facebook users, accounting for approximately 12.2% of Facebook's monthly active users in 2024. This disparity suggests that resource investment is not necessarily proportional to the user base size or the regional risks posed by dangerous groups.[124] The Board's concern was highlighted in a case where a post flagged for review under the High Impact False Positive Override, which is designed to identify and correct mistaken content removals, was never examined due to language capacity bottlenecks.

Whistleblowers who worked at Meta have revealed significant disparities in the company's allocation of attention and resources across different markets. According to The Facebook Papers, a set of internal documents leaked to the media in 2021, 87% of Meta's global budget for classifying misinformation was dedicated to the United States, despite US users constituting only about 10% of the platform's daily active users worldwide. Conversely, only 13% was allocated to the rest of the world.[125] Another whistleblower revealed that the company would prioritize campaigns in the US or Western Europe and foreign adversaries such as Russian or Iran, while ignoring abuses in smaller nations.[126] For example, she lamented that instances like fake engagement by a former president of Honduras and human rights violations in Azerbaijan were largely ignored. Moreover, moderation errors tend to receive more attention and correction when flagged by high-profile accounts with many followers in the

[121] As flagged by the Oversight Board in multiple decisions, such as 2022-003-IG-UA.
[122] 2021-003-FB-UA [123] Oversight Board 2022 Annual Report. [124] 2022-005-FB-UA
[125] https://facebookpapers.com/.
[126] Wong, J. C. (2021) 'How Facebook Let Fake Engagement Distort Global Politics: A Whistleblower's Account'. *The Guardian*, 12 April, sec. Technology, www.theguardian.com/technology/2021/apr/12/facebook-fake-engagement-whistleblower-sophie-zhang.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

United States.[127] Errors involving non-English content are less likely to be noticed or addressed.[128]

2.6 Access to Appeal Processes

No platform has guaranteed users the right to appeal a content moderation decision, even though the opportunity to contest wrongful decisions can enhance procedural fairness. Meta accepts appeals for most types of violations, and appeals may be handled by human or by its automated systems. However, not all appeals will be attended to when the company's review capacity is exceeded. The Oversight Board has observed that non-English content is far less likely to be noticed, flagged, and given additional attention for escalated review (2022-003-IG-UA). Once users have exhausted Meta's three-tiered internal appeal system,[129] they can choose to further appeal to the Oversight Board. Unfortunately, the likelihood of their cases being heard is very low. In 2022, the Board received approximately 1.3 million cases but only published decisions on twelve of them.

Members of the Board are selected from different regions of the world, with efforts made to ensure that the cases they review have a balanced geographical distribution that reflects Meta's user base.[130] According to their bylaws 3.1.3, at least one of the five panel members assigned to each case must originate from the region primarily affected by the content.[131] While this approach is commendable, the appeals they receive do not proportionally mirror the demographics of Meta's global users. In 2023, 38% of appeals originated from the United States and Canada, and 26% from Europe; only 12% came from Asia Pacific and Oceania, another 12% from Latin America and the Caribbean, with just 5% from Central and South Asia, 5% from the Middle East and North Africa, and 2% from sub-Saharan Africa. Yet, as of 2025, only 9% of Meta's users are based in the United States and Canada. Fewer appeals do not necessarily suggest fewer moderation problems but may instead highlight procedural and outreach shortcomings. As discussed, many minoritized language groups cannot access platform interface and rules in their own language and might not be aware of

[127] Power dynamics may be observed across and within language groups. Even within a language group, powerful speakers often enjoy a free pass for violating content based on newsworthiness, and unwarranted censorship they experience are rectified quickly to contain bad press, while the average user's wrongfully removed content may never be reinstated. See Klonick, K. (2018) 'The New Governors: The People, Rules, and Processes Governing Online Speech'. *Harvard Law Review* 131: 1958–1670.

[128] 2022-003-IA-UA [129] Klonick, 'The New Governors'.

[130] Overarching Criteria for Case Selection. https://oversightboard.com/sr/overarching-criteria-for-case-selection.

[131] Their bylaws, last updated June 2025, https://www.oversightboard.com/wp-content/uploads/2025/07/Oversight-Board-Bylaws-June-2025.pdf.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

appeal processes. They may also lack resources such as time and literacy skills needed to navigate the appeals system.

According to the Oversight Board's 2022 bylaws (Section 4.3), Facebook and Instagram users could submit an appeal to the Oversight Board in any language, with their statements translated into English for review. However, in the most recent updated bylaws (2023, 2024 and 2025), this specific section appears to have been removed for reasons that have not been publicly disclosed. Despite this change, the Oversight Board has confirmed through personal communication that the policy remains the same: users can still appeal in any language.[132] The Board's spokesperson clarified that appeals submitted in languages other than English are processed via machine translation into English, alongside the original text. This raises the earlier discussed issue that not all languages are equally well resourced when it comes to machine translation.

The Board's website is currently available in twenty-eight languages, but not all content is translated in all these languages. For instance, announcements regarding new cases are not consistently available in every language, leading to gaps in accessible information for non-English speakers. Similarly, on Meta's Transparency Centre, the availability of content varies across different interface languages, with many non-English language versions missing certain materials. In cases where translations are absent, content defaults to English.

2.7 Participation in Platform Governance

Platforms are a service provider to its users, who have no opportunity to influence how the service is run when they sign up. However, since online platforms have an outsized impact on public life and free expression, scholars have come to expect that they operate with due process, accountability, and even democratic governance.[133] Platforms have, to a varying extent, responded to such expectations, not least as a means of gaining user trust. Large social media platforms and search engines have implemented transparency measures, such as sharing information about content moderation, content restrictions, and government requests for data. They have also occasionally offered users opportunities to help shape platform policies or to provide input to content moderation processes.[134] For example, Facebook has conducted participatory governance

[132] My communication with an Oversight Board spokesperson took place in July 2023.

[133] Klonick, K. (2018) 'Facebook Released Its Content Moderation Rules. Now What?' *The New York Times*, 26 April, sec. Opinion, www.nytimes.com/2018/04/26/opinion/facebook-content-moderation-rules.html; Douek, 'The Meta Oversight Board and the Empty Promise of Legitimacy'.

[134] In one extreme example, Elon Musk polled X users on whether he should step down as CEO, saying that he would abide by the result. Some would call such measures 'transparency theatre', with Tiktok's physical transparency center being a good example. Alex Heath, (2023) 'Inside

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

experiments, including user votes on site policies; however, these efforts often faced limitations, such as low voter turnout, which led the platform to proceed with policy changes opposed by most voters.[135] Platforms also engage experts, civil society organizations, and partner groups in fact-checking and policy development. Incorporating user input into governance structures can enhance perceptions of legitimacy and procedural justice – users are more likely to feel their voices matter when they have opportunities to influence decisions, receive transparent explanations, and see rules applied fairly.

Other than its automated systems and professionalized content moderation workforce, Meta's content moderation model also relies on user flagging and input from external consultants. During the development of the Oversight Board, Meta conducted extensive global consultations, including twenty-eight workshops and roundtable discussions attended by over 650 individuals from eighty-eight countries.[136] While questions can be raised about how well these participants represented Facebook's then-2.5 billion users, it is evident that Meta tried to enhance the legitimacy of its content moderation framework through these consultations.

Unlike traditional judicial processes that are open to the public and involve direct participation by the parties involved, the Oversight Board conducts its deliberations privately behind closed doors. However, similar to how courts sometimes consider the opinions of individuals and organizations not directly involved in a case, known as amicus curiae or 'friend of the court', the Oversight Board seeks to incorporate broader public input on the cases it reviews. When the Board selects a case, it announces this on its website and social media channels, provides a brief, anonymized description of the case, and invites the public to submit comments within a fourteen-day window. The graph in Figure 1 shows, for the cases that the Board dealt with in its initial years,[137] the regional origins of the public commenters compared to the regions primarily concerned by each case.

As illustrated in Figure 1, regardless of the geographic region a case pertains to, the majority of public comments submitted come from the United States,

TikTok's Transparency Center'. *The Verge*, 3 February, www.theverge.com/2023/2/2/23583491/tiktok-transparency-center-tour-photos-bytedance.

[135] Barely 1% of users voted; 30% was required for the vote to be binding. The voting happened in 2009 and 2012. See Dobby, C. (2012) 'Facebook Site Governance Vote Wraps up with Fewer than a Million Ballots Cast'. *Financial Post*, 10 December, https://financialpost.com/technology/facebook-vote-governance-policy-changes.

[136] 'Global Feedback and Input on the Facebook Oversight Board for Content Decisions', *Meta* (blog), 27 June 2019, https://about.fb.com/news/2019/06/global-feedback-on-oversight-board/.

[137] The case number dates back to the year of 2020, but the Board only started publishing its judgments in 2022.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

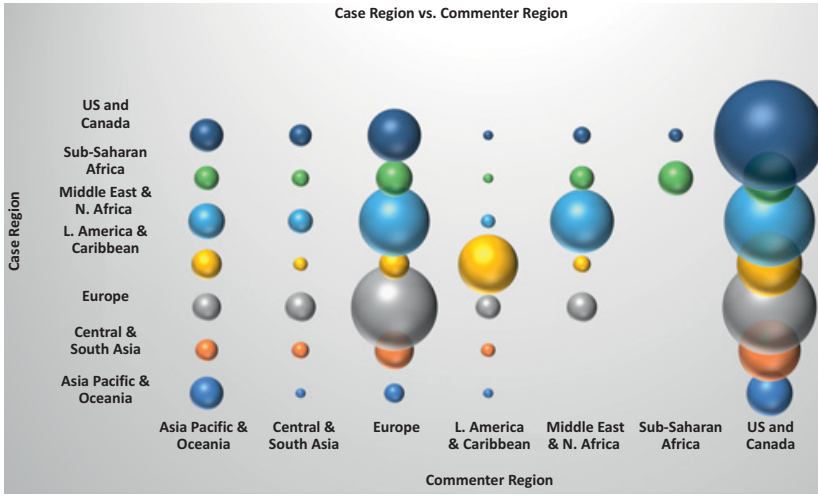


Figure 1 Where the commenters come from versus which region the case is primarily concerned with in the Board's published opinions dated between 2020 and 2024. The graph excludes two cases (2021-001-FB-FBR, on Donald Trump's posts on 6 January 2021, and 2024-004/005/006-FB-UA, on the Palestinian protest slogan From the River to the Sea) which are significant outliers in the data because of the number of comments they attracted. The graph follows how Oversight Board divides the world into regions.

Canada, and Europe. Many of these submissions originate from organizations engaged in related issues such as disinformation or hate speech. This pattern of responses is likely influenced by how comments are solicited and the accessibility of the process. Notably, most comments, regardless of the commenters' location, are submitted in English. Douek argues that this reliance on voluntary, often English-language submissions from wealthier, English-speaking regions means the public comment process may not be genuinely democratic or representative.[138] It tends to favour well-resourced or powerful interest groups for two main reasons: First, although the Board has translated its judgements into eighteen languages, its communication channels, including social media, are predominantly in English, and some cases only accept public input in English. Second, since public commentary is voluntary and unpaid, economic inequalities can limit participation from populations with fewer resources. In a roundtable meeting I attended with some of the Board members in September 2023, they described the public comments system as an amazing opportunity to

[138] Douek, 'The Meta Oversight Board and the Empty Promise of Legitimacy', at 36.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

ask Meta questions, but acknowledged that soliciting sufficiently diverse comments remains a challenge.

An additional challenge to public commentary on or scholarly research into the Board's cases is that the actual content is usually not disclosed due to privacy concerns. Instead, the Board provides a summary of the post in question, which means that observers and researchers do not have access to the actual language or specific details that may have led to content removal. Finally, there's the further question of the actual impact of the public commentary. Although the Board briefly summarizes the opinions received in each case, it does not clearly indicate how, or if, these public inputs impact the final rulings.

2.8 Amplification and Demotion of Content

Platforms do not merely transmit what users post; they actively shape what users see.[139] Apart from moderating content, platforms also actively engage in content recommendation. They curate and personalize feeds by selecting posts, advertisements, and out-of-network content that they believe will maximize user engagement. The dissemination, personalization, and amplification of content are driven by algorithms optimized for user engagement, based on datasets including browsing history, user demographics, sentiment analyses, and other behavioural factors.[140] The inner workings of these recommendation models are largely opaque to users. Resourceful individuals and companies can exploit these systems by paying for increased engagement or manipulating AI-driven algorithms, such as by using bots, to increase exposure. How content is presented to users has profound implications where people use social media as the main source of news and information. Notably, elections and political outcomes have been influenced, sometimes decisively, by how content is disseminated and amplified on platforms like Facebook.

Social media platforms' business model is based on advertising revenue and data extraction. As a result, content go viral not only because they resonate with many users, who like, share, or comment on posts, but also because the platform's algorithms actively promote these posts to garner more attention. Conversely, posts that do not receive attention are not always irrelevant to people's interests – they might not have the opportunity to receive audience exposure. A well-observed problem with this business model is that the algorithms create an echo chamber or filter bubble that reinforces users' viewpoints

[139] Gillespie, *Custodians of the Internet*, at 21.
[140] 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression'.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

even when they are not supported by facts, making the society more polarized and people less empathetic towards one another.[141]

Meta has not publicly disclosed the specifics of how its content recommendation algorithms function, and it has declined to provide the Oversight Board with detailed insights into these systems. This lack of transparency has raised concerns about how platform design may incentivize the promotion of sensationalist or provocative content (2021-016-FB-FBR). Generally, content predicted to be engaging to a broad audience is boosted, while posts expected to generate low engagement are deprioritized. Such algorithms tend to reward outrage and inflammatory material, often at the expense of minority voices or nuanced discussions.[142] The language of a post can influence how widely it is disseminated, although major platforms do not openly reveal the criteria they use to prioritize content, partly out of fear that users could manipulate or 'game' the system. But there is evidence of English posts boosted at the cost of other languages. In 2023, components of Twitter's recommendation algorithm were open-sourced, providing some insights into its content ranking. While these disclosures do not fully clarify how factors are calculated, they reveal a bias toward English-language content: tweets in English tend to be promoted more than those in other languages, and tweeting in a language different from one's interface language is often viewed unfavourably by the ranking system.[143] Essentially, the algorithm manifested a monolingual English ideology: it did not appear to promote other popular languages on the platform, such as Japanese or Spanish; the distinction is simply English and non-English. Tweets written in a language deemed to be 'unknown' to the user are heavily demoted.[144] Although this is likely to be a measure to reduce noise on a user's feed, it also reduces opportunities for exposure to cross-cultural perspectives through machine translation and prevents multilinguals from taking full advantage of their linguistic repertoires.

This analysis indicates that having equal opportunities to participate and post on social media platforms does not necessarily guarantee equal opportunities to be heard. Content creators who produce content in English are more likely to benefit from amplification and wider dissemination due to platform biases in content ranking algorithms. Such disparities in content visibility are especially

[141] Pariser, E. (2011) *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. London: Penguin.

[142] According to a whistleblower, Facebook charges less for ads that are more incendiary. Wynn-Williams, *Careless People*.

[143] See https://github.com/twitter/the-algorithm/blob/main/src/thrift/com/twitter/search/common/ranking/ranking.thrift.

[144] See https://github.com/twitter/the-algorithm/blob/main/src/thrift/com/twitter/search/common/ranking/ranking.thrift.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

significant in regions lacking strong independent media, where social media often functions as a vital platform for free expression. During crises and conflicts, when the risk of physical violence is heightened, the ability to share information becomes even more critical. For example, a human rights assessment report on the conflict between Palestine and Israel in May 2021 shows that the ability of Palestinians to share information about their experiences was more heavily curtailed than their Israeli counterparts, not only because of Meta's differential ability in moderating content in Arabic versus Hebrew but also because of harm-mitigation measures that reduced the visibility of all repeatedly reshared content.[145] These 'break-the-glass' measures, meant to be emergency responses to extraordinary circumstances, are seemingly content neutral but can disproportionately impact certain geopolitical contexts. The practice of algorithmically reducing the visibility of a user's posts without notification, sometimes called shadow banning, further limits individuals' ability to communicate and be seen, effectively silencing or marginalizing voices without their knowledge.

In summary, minoritized voices on social media platforms are often suppressed or erased due to linguistic and cultural biases inherent in content moderation and risk management practices. Additionally, when the social and communicative needs of these communities conflict with the platforms' corporate goal of maximizing user engagement, their content may be demoted. This combination of structural biases and algorithmic incentives can significantly hinder diverse voices from being heard and represented online.

2.9 Understanding the Impact

The capacity of global populations to express themselves, access information, and engage in the digital economy and public discourse is now largely mediated by private social media platforms. The language-related decisions made by these private entities, regarding policy design, content moderation, resource allocation, and algorithmic prioritization, carry significant political implications. These choices can result in material injustices, affecting not only active platform users but also the broader communities they are part of. Such linguistic injustices manifest as *epistemic inequalities* (limits on access to knowledge and information), *participatory inequalities* (barriers to meaningful engagement), *representational inequalities* (underrepresentation or misrepresentation of cultural identities), and *distributive inequalities* (unequal distribution of resources

[145] *Human Rights Due Diligence of Meta's Impacts in Israel and Palestine in May* 2021, Business for Social Responsibility.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

and opportunities). Collectively, these disparities undermine the inclusivity and fairness of digital spaces for diverse language communities.

Social media platforms facilitate access to vast amounts of user-generated content, enabling peer-to-peer connectivity among language communities worldwide. However, persistent digital language divides remain, even for languages officially supported by these platforms. While advancements in machine translation improve access to information across languages, they also pose risks of errors and inaccuracies, leaving speakers of low-resource languages more vulnerable to misinformation and miscommunication. The quantity and quality of content available in different languages vary significantly, partly due to uneven platform capacities to filter harmful, dangerous, or misleading material. What starts as an epistemic gap – limited access to reliable information – can quickly escalate into crises threatening the safety and well-being of both platform users and non-users. These divides are not solely the result of uneven information availability, a bottom-up phenomenon, but are also influenced by top-down factors such as disparities in investment for harm mitigation and moderation efforts. Ultimately, the divides are mediated by power relationships, as the next section will explore further.

In the era of interactive media, assessing linguistic inequalities must extend beyond mere consumption to include content production, as active participation is essential for minoritized language speakers to have their voices heard and to meaningfully engage in the digital economy. Structural barriers such as device and connectivity disparities, along with the limitations of zero-rated services, contribute to the production gap between the Global North and the Global South. It has been observed that the great majority of people who use crowd-sourcing forums, from mailing lists and blogs to social media, do not produce content actively: around 90% are passive consumers of information, 9% contribute occasionally, and 1% is actively generating content.[146] Usage divide is a behavioural phenomenon and does not necessarily point to unfairness, but uneven patterns of digital production may reveal underlying structural barriers. In the US, voices of the working class are marginalized; survey data found a class-based digital production gap among those who are online,[147] with education being the most important predictor of production. Since digital participation involves labour, often unpaid, those with limited time and

[146] Nielson, J. (2006) *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities*. 8 October. See www.nngroup.com/articles/participation-inequality/. The ratio of passive consumers goes up to as high as 99% on a low-engagement platform like Wikipedia, which has over a billion visitors monthly, but only 0.011% of them contribute actively. See https://en.wikipedia.org/wiki/Wikipedia:About

[147] Schradie, J. (2011) 'The Digital Production Gap: The Digital Divide and Web 2.0 Collide'. *Poetics* 39(2): 145–68. https://doi.org/10.1016/j.poetic.2011.02.003.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

resources are less able to contribute.[148] For minoritized language users, these structural barriers, such as difficulties in understanding platform rules, being notified of content removals, or facing algorithmic biases favouring majority languages, intersect with pre-existing inequalities. This participation gap translates into unequal access to economic opportunities, as content creators who reach larger audiences can monetize their work and benefit from the digital economy, further marginalizing those who are unable to do so.

Inequality in digital content production directly fuels representational inequality, which pertains to how different peoples and issues are depicted online. For example, in 2011, the Wikipedia entries for Jerusalem varied significantly: the Arabic version described it as the largest city in occupied Palestine, while the Hebrew version called it the capital of the State of Israel.[149] Such conflicting narratives exemplify how representations are contested, especially around sensitive topics like historical events, land claims, and identity. Distorted or incomplete digital representations often reflect global structural inequalities. For example, although Wikipedia exists in 360 languages,[150] some editions in smaller languages were created by non-speakers residing in the Global North based on dictionaries and machine translation. An American teenager had produced half of the articles in Scots, a minority language in Europe, by inserting Scots words into English sentences. A large number of articles written in isiXhosa, which is spoken as a first language by 8 million people in South Africa, were found to be created by contributors in the Global North who had very limited knowledge of the language.[151] Although Wikipedia entries are not permeant and can be corrected through communal efforts, such efforts require sufficient native speakers to be online and be prepared to volunteer their labour. Social media platforms contribute to representational inequality through algorithmic amplification of certain viewpoints over others, often to the detriment of minoritized communities. Input on platform governance, such as public commentaries that the Oversight Board receives, also tends to be dominated by perspectives from high-income countries. Representational inequality reinforces existing economic, political, and cultural domination, and challenges the optimism that digital media offers a more democratic and inclusive marketplace of ideas than traditional mass media.[152]

[148] Deumert, 'The Invisibility of Linguistic Diversity Online'.
[149] Graham and Dittus, *Geographies of Digital Exclusion*, at 4.
[150] As of February 2026. See https://en.wikipedia.org/wiki/List_of_Wikipedias.
[151] Deumert, 'The Invisibility of Linguistic Diversity Online'.
[152] Benkler, Y. (2006) *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Distributive inequality manifests in the unequal allocation of resources across different platform languages, affecting everything from the translation of rules and policies to content moderation efforts. This uneven distribution means that the transformative opportunities, and the risks, associated with connectivity are not shared equitably among users worldwide. Certain communities face language bottlenecks that limit their ability to access platform information or have their content properly moderated, while others benefit from more robust support. Although platforms naturally prioritize geopolitical contexts that are most relevant to their strategic interests, neglecting less prominent or marginalized communities can have serious human rights implications.

The following section examines how such inequalities might be mitigated, what they reveal about the structure of the global digital order, and why conceptions of linguistic justice must account for digital transformations, particularly with regard to competition in the global linguistic market and the role of private actors in shaping linguistic hierarchies.

3 Interventions and Interpretations

> Languages, like individuals and states, are constantly involved in the conflicts of precedence and domination which are a ceaselessly repeated feature of the hierarchies in which they exist.[153]
>
> Jean Laponce

Although a world in which all languages are treated equally is not attainable, many of the linguistic inequalities discussed in the last section can be contested. This section begins by setting realistic goals and identifying the levers, such as legal interventions, market forces, and community engagement, that can facilitate progress toward greater linguistic equity on digital platforms.

In the latter part of the section, I will situate digital linguistic inequalities within broader theoretical frameworks. I argue that these inequalities are not isolated issues but are embedded as a core feature of information capitalism. I will differentiate between inclusion and inequality, challenging the assumption that inclusion is inherently emancipatory, and highlighting the complexities related to agency, risk, and power that accompany efforts toward inclusion.

Using Pierre Bourdieu's theory of social stratification, I will analyse the global linguistic market created within the digital sphere, where competition underlies access, influence, and resource distribution in the political economy of digital platforms. This perspective questions whether traditional concepts like the digital divide or digital exclusion adequately capture the evolving nature of digital inequalities in today's interconnected and competitive environment.

[153] Laponce, J. A. (1987) *Languages and Their Territories*. Toronto: University of Toronto Press. 3.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Finally, I will assess the applicability of existing solutions proposed for addressing linguistic injustice in the digital realm. I argue that these approaches need to be expanded and adapted to account for the complexities of the digital field, emphasizing the need for systemic strategies to promote linguistic equity in the digital age.

3.1 The Goals

Linguistic inequalities both reflect and intersect with deep-seated global divides and digital platforms may have unintentionally amplified them. Addressing these issues is complex. There is no straightforward solution that can fully undo or reverse entrenched global power asymmetries, and social and political challenges extend beyond legal and technological remedies. Nonetheless, many of the negative impacts stemming from platform policies and processes can be mitigated through targeted actions. This section outlines achievable goals for the near future.

3.1.1 Setting Minimum Standards

Given the vast diversity of languages and the uneven sizes of language communities on digital platforms, it is unreasonable to expect all platforms to provide their services fully in every human language. A realistic approach to goal setting must recognize resource constraints and practical limitations.

Establishing industry-wide standards about multilingual content moderation is a viable first step. Such standards would not only set minimum expectations but also compel platforms to invest in language detection so they can identify the all range of languages that are being used on their sites. In 2014, violence broke out in Myanmar because of a viral post on Facebook. Its content team could not understand the post and would leave it on the site for hours while they tried to reach the only Burmese-speaking moderator Facebook had, who was based in Dublin and who went out to dinner without his work laptop.[154] As the Burmese-speaking moderator workforce grew to just two in 2015, issues persisted, as one of them appeared to be allowing anti-Muslim content on the site and removing posts by civil society and peace activists.[155] The example illustrates that both responsiveness and quality assurance in multilingual moderation requires proper resourcing. Although Facebook hired dozens of content moderators after the genocide in Myanmar, it would need to have around 800 of them in order to match the ratio of reviewers to users in Germany.[156] While uniform

[154] Wynn-Williams, *Careless People*. [155] Wynn-Williams, *Careless People*.
[156] Roose, K. and Mozur, P. (2018) 'Zuckerberg Was Called Out Over Myanmar Violence: Here's His Apology'. *The New York Times*. 9 April. www.nytimes.com/2018/04/09/business/facebook-myanmar-zuckerberg.html.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

ratios across all languages will not be practical given uneven geopolitical risks and resource constraints, establishing a minimum ratio of human content moderators relative to the user base for each language is highly desirable. It would reduce platforms' reliance on a whack-a-mole approach to content risk mitigation. Furthermore, hiring moderators from within the country and ensuring they understand local politics enhances moderation quality and cultural sensitivity. Transparency reporting should include data on the number of moderators per language, distinguishing between employees and contractors. When platforms cannot meet minimum moderation standards, they should proactively inform users in their preferred language that moderation capacity is limited.

Given the enormous volume of daily content, mandating human review for all posts is impractical and may need to be limited to posts that surpass certain engagement thresholds. While automated language technologies offer valuable support, their effectiveness is limited by factors such as data availability, especially for low-resource languages, and their current technological shortcomings. To address these challenges responsibly, transparency must be significantly enhanced. Platforms should publicly disclose the accuracy rates of both algorithmic and human moderation efforts, broken down by language. Maintaining a repository of moderation decisions for public scrutiny can also foster accountability and trust.

Similarly, transparency regarding machine translation quality across different languages is essential; platforms should provide clear information about potential translation inaccuracies, particularly for low-resource languages. When machine translation mediates platform communication, users should be warned about possible quality issues in their selected interface language. Implementing these transparency measures can help users better understand the limitations and risks of automated moderation and translation, promoting informed participation. Decisions about whether human review is necessary should be determined not only by the level of location risk but also by the extent to which the language involved is low-resourced.

Ensuring that platform rules are available in all officially supported languages and that content moderators have access to implementation standards in their preferred working language is another practical step. As noted by Meta's Oversight Board, the absence of such translations can raise significant human rights concerns, including barriers to understanding and exercising rights on the platform (Sections 2.4 and 2.5). Platform communication, including notifications and warnings, should be delivered in the user's chosen interface language. This approach helps lower participation barriers, ensures users fully understand platform policies, and supports due process by allowing users to access processes like appealing moderation decisions in their preferred language.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

However, a potential risk exists if minimum language support standards are set solely based on a platform's interface languages: platforms can decide to simply eliminate them. Facebook, for example, has a history of adding interface languages only to remove them a few years later.[157] This could be mitigated by developing an industry standard that requires platforms to adopt a language that crosses certain usage threshold as an interface language.

3.1.2 Decentralizing Platform Governance

For Facebook, most content moderation is done algorithmically, especially for English. For central and peripheral languages which are not served by content moderation algorithms, moderation often relies on reviewing content through machine translation. Many languages used on these platforms remain outside the scope of both machine and human moderation efforts altogether. Mitigating inaccuracy in algorithmic content moderation or machine translation is not straightforward, as AI operates in a black box. Gaps in content moderation can only be mitigated to a small extent through transparency reporting and user warnings, but these will not do much for speakers of languages that platforms do not officially support.

Professionalized human moderation is costly to scale, but it can operate at a larger scale if non-professionals are involved. Some platforms, such as Reddit and Facebook groups, appoint community managers to moderate content. A community moderation system may supply the linguistic and cultural competence that centralized content moderation often lacks. Indeed, large platforms appear to be converging towards community moderation, involving potentially all users rather than a selected number of community managers. In 2025, Meta announced that it is planning to introduce a community notes programme. Bluesky and X already have a similar system. For example, X uses a community-driven approach to identifying misleading content, allowing users to provide context that can help others understand or evaluate a post. Community notes are publicly displayed if they are sufficiently upvoted as 'helpful' by users 'from different points of view',[158] while notes sufficiently downrated as 'not helpful' do not appear. An author who disagrees with a community note attached to their post can request for an additional review, which is not done by X's moderators but again by users who signed up to be contributors.

[157] For example, Pulaar and IsiXhosa have been removed as an interface language on Facebook, as of 2025.

[158] Defined as contributors who have sometimes disagreed in their past ratings. See https://communitynotes.x.com/guide/en/about/introduction.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

One benefit of community moderation is that it protects marginalized communities from having to compete for attention and resources in the centralized content moderation system. While crowdsourced moderation is economical for the platforms and provides an opportunity for people with relevant cultural, linguistic, and contextual knowledge to contribute, the approach has limitations. Posts need to be circulated sufficiently widely for a note to gain enough votes before it is visible to others. People who are not experts may upvote seemingly believable statements that are not factually correct. Fundamentally, the challenge of arriving at bipartisan notes in a heavily polarized world where people 'can't agree on basic facts'[159] might mean that not many posts receive meaningful contextual notes. Finally, community moderation raises the question of unwaged labour. Considering its benefits and limitations, community moderation may be better adopted as complementary to rather than a full replacement of a top-down moderation system.

The use of experts fills a critical gap between centralized moderation system and crowdsourced community moderation. Experts may supply local knowledge (of culture, politics, etc.), verify facts, or provide expertise on particular issues such as child abuse and domestic violence. Unfortunately, the industry seems to be moving away from its reliance on experts, which is likely to dampen efforts in risk mitigation. For example, Twitter's (now X) Trust and Safety Council was dissolved in 2022, and in January 2025, Meta announced the termination of its third-party fact-checking program.[160]

Another way of decentralizing platform governance is for platforms to distribute its decision-making authority more horizontally, with transparent allocation of resources and regional content adjudication units who are attuned to the sociopolitical conditions of their user communities.[161] These units can adapt and apply platform rules in ways that better reflect local norms and sensitivities. Decisions from the local and regional units could ultimately be reviewed by a higher-level international team. Users and civil society could participate not only in content moderation but also in company decision-making about platform policies and processes. This could take the form of multi-stakeholder councils established in

[159] Alex Mahadevan, director of MediaWise, cited in Lorenz, T., Oremus, W., and Merrill, J. B. (2022) 'How Twitter's Contentious New Fact-Checking Project Really Works', *Washington Post*, 10 November, www.washingtonpost.com/technology/2022/11/09/twitter-birdwatch-factcheck-musk-misinfo/.

[160] Meta's decision to end its third-party fact-checking programme coincided with the second presidency of Donald Trump. In justifying the decision, CEO Mark Zuckerberg said independent experts were 'too politically biased' and he wanted to prioritize free speech. https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/.

[161] Jhaver, S., Frey, S., and Zhang, A. (2023) 'Decentralizing Platform Power: A Design Space of Multi-Level Governance in Online Social Platforms'. *Social Media + Society* 9(4): 1–12. https://doi.org/10.1177/20563051231207857.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

the countries or regions where platforms operate, involving users and civil societies, to enhance democratic governance.[162]

3.2 The Levers

The good news is that major transnational platforms possess the resources necessary to implement the recommendations outlined above. The bad news is that private actors do not have the same responsibilities to their platform users as governments have to their citizens. Companies are free to choose which values to prioritize, or to prioritize none at all.[163] The question is then how they may be incentivized or coerced to act in ways that promote responsible governance. This section further focuses on the levers for change, acknowledging that some are clearly more accessible than others and each has its own limitations.

3.2.1 Leveraging Market Forces

Apart from areas such as privacy or criminalized speech, where governments impose rules and expect compliance, platforms mostly regulate themselves. Industry self-regulation primarily serves as a way to pre-empt more intrusive government regulation. However, government regulation is not the only source of constraints shaping the business environment in which platforms operate. Other factors, such as market pressures, public expectations, reputation considerations, and international norms, also influence platform practices and governance.

Since the mandate of a corporation is to pursue profit, measures that promote inclusivity and bring social benefits are only adopted as far as they align with this mandate. Widening linguistic inclusion could enlarge a platform's user base and potentially increase profit, but platforms see diminishing returns as they expand their digital language support to smaller languages and less advanced economies. Advertising or data extraction income generated from very small language communities is unlikely to provide a return on investment on the cost of digital language support.[164] Similarly, content moderation is crucial for maintaining user trust and attracting advertisers; platforms risk losing credibility and advertising revenue if their environments are filled with abuse and toxicity. At the same time, certain markets are valued much more than others in content moderation efforts. For example, although only 9% of Facebook's

[162] Kaye, D. (2019) *Speech Police: The Global Struggle to Govern the Internet*. New York: Columbia Global Reports.

[163] Klonick, 'The New Governors', at 1617.

[164] Platforms can generate advertising revenue from a market even without providing service to users there. Facebook is banned in China but China is its second largest market of advertising revenue after the United States.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

users were based in the United States and Canada, North America accounted for 38% of Meta's total revenue in 2024, highlighting how economic value influences resource allocation in moderation efforts.[165]

Corporate social responsibility is a good example of companies responding to social and market pressures and showing sensitivity to consumer and investor expectations around ethical and sustainable business practices. Since platforms compete with one another in the same field, they actively manage their brand image in order to maintain and grow their user base. Efforts to build legitimacy, such as publishing transparency reports and establishing governance structures like the Oversight Board, are part of this strategy. As Gillespie notes, users possess collective power over platforms, not necessarily as individuals or groups but in aggregate.[166] There are historical examples where users have collectively abandoned one platform, like Friendster, in favour of another. The court of public opinion remains a powerful force: advertisers, for instance, are reluctant to have their ads appear alongside harmful or abusive content.[167] Notably, many advertisers withdrew from X (formerly Twitter) after the platform significantly reduced its content moderation capacity, fearing damage to brand safety and reputation.

Civil society has played an active role in collaborating with platforms to improve digital language support. A notable example is the translation of Facebook's interface into Inuktitut, led by the Pirurvik Centre, a translation and learning organization based in Iqaluit, Nunavut, in Canada.[168] Beyond direct collaboration, civil society can also influence industry standards; for instance, a coalition of organizations, advocates, and academics developed the Santa Clara Principles on Transparency and Accountability in Content Moderation. These principles set out minimum standards for content moderation, including the requirement that reports, notices, and appeal processes be available in the user's language and that users should not be disadvantaged based on their language, country, or region. First introduced in 2018 and updated in 2021 to incorporate perspectives from the Global South, the Santa Clara Principles have been endorsed voluntarily by major tech companies such

[165] *Annual Report*, Meta Platforms Inc., 30 January 2025. Asia Pacific contributed 27% and Europe 23%, while the rest of the world, including Africa, Latin America and the Middle East, contributed 11%.

[166] Gillespie, *Custodians of the Internet*, at 23.

[167] Roberts, J. J. (2020) 'TikTok's Appalling "Holocaust Challenge" Signals Facebook-Style Problems Ahead', *Fortune*, 8 September, https://fortune.com/2020/09/07/tiktok-holocaust-challenge-trends-facebook-problems-content-policing/.

[168] Nunavut Tunngavik Inc. *Meta announces Inuktitut as official language on Facebook desktop*. 8 July 2022. www.tunngavik.com/news/meta-announces-inuktitut-as-official-language-on-facebook-desktop-new-language-settings-to-support-nunavut-tunngavik-incorporateds-efforts-to-strengthen-and-promote-the-vitality-of-inuit-lan/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

as Apple, Meta, Google, Reddit, and Twitter (now X). While adoption is not mandated, these efforts contribute to shaping norms in the field of platform governance and content moderation.

Engaging directly with platforms remains a significant challenge. Civil society has been contributing to fact checking and to alerting platforms about content violations, though their voice is often ignored. For example, during the Rohingya genocide, civil society organizations in Myanmar had difficulty escalating to and getting a prompt response from Meta's content moderation team.[169] While civil society can now also engage with the Oversight Board by submitting comments on specific cases, the actual influence of these inputs is still uncertain.

Another market dynamic that could potentially benefit the public arises from competition within the tech industry. Dominant players like Apple and Google have significant control over smartphone operating systems and can exert pressure on platforms such as Meta. For instance, in 2021, Apple introduced enhanced privacy controls on its iPhones, enabling users to opt out of being tracked by apps like Facebook. This move impacts advertising practices and data collection strategies of platforms reliant on targeted ads. Additionally, companies like Microsoft, whose business model is not so dependent on ad revenue, have advocated for regulatory measures requiring platforms that circulate news, such as Facebook and Google, to share revenue with local news organizations that produce the content.[170] These industry battles can shape market behaviours and potentially lead to outcomes that favour consumer interests and the sustainability of diverse media ecosystems.

If we take platforms' word for it, enabling voice is at the core of their business. Meta declares 'voice' as its 'paramount value', while YouTube says their mission is 'to give everyone a voice and show them the world'.[171] Many express an interest in supporting minoritized language communities, though they seem to be exclusively focused on scalable technological solutions (see machine translation initiatives by Google and Meta discussed in Section 2.2). However, external actors can exert pressure on platforms to take these values more seriously by exposing shortcomings in their digital support for minoritized languages, especially when there are compelling examples of how this is affecting people's lives and leading to unjust outcomes. For example, following the Rohingya genocide, Meta assisted in transitioning the Burmese font to

[169] See their open letter to Mark Zuckerberg: https://embed.documentcloud.org/documents/4432469-Myanmar-Open-Letter-to-Mark-Zuckerberg/?embed=1.

[170] Bradford, A. (2023) *Digital Empires: The Global Battle to Regulate Technology*. New York: Oxford University Press.

[171] https://about.youtube/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Unicode, a necessary step for better language support. Although it should not take a genocide for a platform to collaborate with local governments to improve language encoding, the reality is that platform policy changes tend to be reactive rather than proactive. Pursuing visibility and influence in the attention economy often means playing by its rules, which require access to technology and digital literacy skills. Even anti-capitalist activists leverage commercial social media platforms instead of dedicated activist spaces to reach broader audiences.[172]

Although market dynamics evolve over time and technological advancements may reduce the costs of digital language support and enhance content moderation effectiveness, market forces generally tend to deepen rather than alleviate linguistic inequalities. As legal scholar Owen Fiss observes, the market 'does not ensure that all relevant views will be heard, but only those that are advocated by the rich'.[173] While platforms have taken on the role of custodians of the internet, relying on them to undertake redistributive efforts to support marginalized languages and communities is risky. Without deliberate intervention or regulation, these inequalities are likely to persist or even worsen.

3.2.2 National and Regional Regulation

To understand the regulatory landscape of the tech industry, we need to begin with Section 230 of the Communications Decency Act, a piece of law in the United States that has shaped the internet as we know it today. Section 230 offers immunity to online intermediaries from tort liability arising from content posted by third parties on their sites. The law aims to strike a balance by protecting users' free speech rights while incentivizing platforms to moderate offensive or harmful content. This law has created a safe harbour for American platforms, which enjoy the right but not the responsibility to moderate content, to grow and dominate the global digital order. The American approach is libertarian, favouring free speech, free market, and limited government intervention.[174] This approach has enabled innovation but also raised issues about accountability.

Governments around the world are increasingly stepping in to regulate and limit the vast private power held by digital platforms. The European Union (EU) has been a trailblazer in establishing and asserting digital rights. Even in the United States, there is growing momentum for more comprehensive digital regulation, signalling the waning era of self-regulation by platforms. Contrary

[172] Deumert, 'The Invisibility of Linguistic Diversity Online'.
[173] Fiss, O. M. (1986) 'Free Speech and Social Structure'. *Iowa Law Review* 71(5): 1405–1426. 1412–1413.
[174] Bradford, *Digital Empires*.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

to the broad immunity granted by US law, most countries in Europe and South America impose conditional liability, which means platforms are not held liable for third-party content as long as they lack actual knowledge of unlawful material and respond promptly to removal requests once notified. In China and many Middle Eastern countries, intermediaries are held strictly liable for content on their platforms, which compels them to proactively remove unlawful content.[175]

Current legislative initiatives are primarily concerned with data privacy and content regulation aimed at combating issues like terrorism, hate speech, misinformation, and cyberbullying. Heavy-handed and reactive government regulation threatens free speech by focusing on what must be taken down, often neglecting the importance of what should be preserved for open debate. Such regulation can inadvertently undermine dissent and open discourse, while also increasing private power by outsourcing public responsibilities to private bodies and entrenching platforms' role as stewards of public space.[176] Moreover, a one-size-fits-all regulatory approach can stifle market competition, as larger platforms possess greater resources to ensure legal compliance, potentially reinforcing their dominance. Digital enclosures, spaces where user behaviour is closely monitored, often operate under opaque processes, with decisions that are difficult to challenge legally.[177] For instance, while content removed for violating platform rules can sometimes be appealed, users like those on Meta cannot contest removal decisions initiated by governments.[178] Although platforms have played vital roles in documenting government abuses and amplifying political dissent, especially in authoritarian contexts, they generally conform to local laws, raising concerns about collateral censorship.[179] This phenomenon occurs when private actors, out of fear of legal repercussions, censor content beyond what the law mandates, thus amplifying restrictions on free expression and risking the suppression of legitimate speech.[180]

Platform regulation is often driven by motives beyond harm reduction and individual rights protection. Countries that export technology use platforms to power their economy and compete with other states.[181] Countries who are at the receiving end of foreign technologies may use regulation to limit the influence

[175] Gillespie, *Custodians of the Internet*. [176] Kaye, *Speech Police*.

[177] Andrejevic, M. (2007) 'Surveillance in the Digital Enclosure'. *The Communication Review* 10: 295–317. https://doi.org/10.1080/10714420701715365.

[178] See 2022-007-IG-MR. [179] 2022-002-FB-MR.

[180] Law enforcement requests can lead to lawful content being reviewed against platform rules and taken down; this kind of requests may be made informally, without necessary oversight and safeguards. States may also ask social media companies to provide user data.

[181] See relevant discussion in Jin, D. Y. (2013) 'The Construction of Platform Imperialism in the Globalization Era'. *TripleC* 11(1): 145–172, https://doi.org/10.31269/triplec.v11i1.458, and Bradford, *Digital Empires*.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

of foreign powers, aligning with nationalist goals or national security concerns.[182] On the other hand, many governments, authoritarian and democratic alike, may not want to regulate too strictly, because they rely on platforms for mass surveillance and political influence.[183] For example, Myanmar's military regime had weaponized Facebook in a massive misinformation campaign. Similarly, the revelations by whistleblower Edward Snowden exposed US government's extensive surveillance operations, which involved accessing data stored on tech platforms' servers.

The architecture of the internet, invented in the United States and exported worldwide, facilitates control, 'not just by other democratic governments, but by any government, however repressive'.[184] As Eubanks observes, marginalized groups are disproportionately monitored and tracked, and digital inequalities are often experienced as a social group rather than as individuals.[185] Such surveillance poses significant threats to dissidents and minority communities, undermining their expression and safety.[186] For example, on Douyin, the Chinese version of TikTok, streamers who spoke ethnic languages would receive a warning to switch to Mandarin or to have their livestreams cut off, and the company was pressured to develop algorithms that automatically detect Uyghur and automatically cut off its speakers.[187]

Rather than focusing heavily on content moderation, regulation may target upstream issues such as system designs, organizational processes, and business decisions, which are in place before any user content is generated. Features such as content recommendation algorithms, moderation workflows, and procedures for handling user complaints can significantly influence the safety and fairness of the speech environment. Regulation can mandate greater transparency in platform policies, decision-making processes, and algorithmic operations,

[182] Techno-nationalism encourages countries to develop or promote domestic technologies and try to become technologically self-sufficient. This is most evident in China, and in Russia to some extent. If more and more countries enact policies that reduce their interdependence in the digital realm, the internet might be carved up into fragmented online spaces according to political borders and becomes the splinternet.

[183] See Wynn-Williams, *Careless People*, for accounts of Facebook helping politicians win elections in exchange for their commitment to not regulate platforms.

[184] Lessig, L. (1998) *The Laws of Cyberspace*. https://cyber.harvard.edu/works/lessig/laws_cyberspace.pdf. 13.

[185] Eubanks, *Automating Inequality.*

[186] Political activists around the world have been arrested, convicted, and imprisoned for their social media posts or likes. Chala, 'Ethiopian Protester Sentenced to Six Years Behind Bars for Facebook Posts'; Palatino, 'In Thailand, Facebook "Likes" Can Land You in Jail'; Macha, 'In Tanzania, Expressing Political Opinions on Social Media Is Becoming Increasingly Dangerous'; Amnesty International, '5 Countries Where Your Online Comments Could Land You in Jail'.

[187] Lu, S. (2023). 'I Helped Build ByteDance's Censorship Machine'. *Protocol*. 18 February 2021.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

encouraging platforms to open their systems to public scrutiny and ensuring that procedural fairness is upheld.

Few jurisdictions have used legislation to specifically tackle questions of linguistic inclusion and inequality on digital media platforms. States do routinely regulate language use by private companies, but their interest tends to be limited to languages that enjoy de jure or de facto official status.[188] These regulations typically require that product packaging, warning labels, and customer service be provided in one or more official languages. For example, in Canada, product warnings and food labelling must be displayed in both English and French, and Quebec's language law requires businesses make their website available in French. Some jurisdictions have adopted legislation requiring social media platforms to provide digital language support in selected languages. The Digital Services Act (DSA), enacted by the EU, requires big tech to publish their terms and conditions in the official languages of all EU member states where they operate (Article 14(6)). They must also include in their terms and conditions

> information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review, as well as the rules of procedure of their internal complaint handling system. (Article 14(1))

The DSA includes provisions that mandate platforms communicate with users in an accessible and user-friendly manner regarding content moderation decisions, user complaints, and related processes. While the DSA does not oblige platforms to provide full digital support in all EU official languages, it communicates such an expectation in key areas. Similarly, India's 2021 Information Technology Rules stipulate that intermediaries must inform users its rules and regulations, privacy policy, and user agreement in their preferred language, which may be English or any of the country's twenty-two official languages.

Apart from enhancing accessibility and transparency, media law can serve to protect local languages and cultures. Many countries impose quota on traditional mass media to support local content, such as cinema screen quotas, or requirements for local language programming on radio and television. For example, in Canada, the Canadian Radio and Telecommunications Commission mandates that at least 35% of content broadcast by commercial

[188] Governments often prioritize the protection and promotion of their official or national languages to foster national unity and cohesion. Out of the thousands of languages spoken worldwide, only around a hundred have some official status, leaving the vast majority with limited or no institutional backing.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

radio stations be Canadian, with television programming comprising at least 55% Canadian content. French-language radio stations are also required to ensure that at least 65% of their music selection is in French. Such local content requirement offers protection to the Canadian creative industries and promotes content diversity in the media. However, the impact of such laws has been weakened by the emergence of digital media where American cultural exports are globally dominant. The Online Streaming Act, enacted in 2023, requires commercial online broadcasting services to contribute to Canadian cultural productions and to ensure that Canadian content is easily discoverable. Although digital platforms do not create original content, they curate it through moderation and recommendation and could play a role in ensuring the availability of quality content in local languages and consolidating their position in the global marketplace. It is conceivable that states require a minimum percentage of recommended content on platforms to be in a certain language.

Not all measures that could benefit minoritized language communities are explicitly focused on language. Antitrust, which involves regulation that prevents anti-competitive practices by companies, could limit the concentration of market power among a few transnational platforms and open space for alternative services. Global tax equity is another example. The digitalized economy has allowed platforms to derive income from users aboard without paying tax, making investments, or creating jobs locally. To address this, countries such as France and India already impose a digital services tax that targets firms with significant economic presence in their country. Additionally, the Organisation for Economic Co-operation and Development (OECD) is working on a multilateral agreement that aims to reform the international tax system to give governments the right to tax large multinational corporations that make sales in their countries, even if these companies do not have a physical presence there.[189] Getting transnational corporations to pay their fair share of taxes can help mitigate the erosion of tax base caused by digital globalization, thereby strengthening governments' capacity for public spending and redistributive efforts.

Not all governments have equal leverage when dealing with large tech companies. When problems arise, smaller or less influential countries struggle to attract the attention of big tech, negotiate effectively, or enforce their laws due to limited regulatory capacity. When a country imposes stringent regulations, platforms may simply choose to exit that market and focus on more profitable

[189] See www.oecd.org/en/topics/base-erosion-and-profit-shifting-beps.html. In the global trade war that the US instigated in 2025, President Donald Trump threatened to retaliate against countries that impose foreign digital services taxes on US companies.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

regions. To bolster their bargaining power, countries can collaborate through regional alliances such as the African Union or ASEAN.

3.2.3 Human Rights and International Law

Government overreach into individual freedoms has prompted the creation of the Global Network Initiative (GNI), an industry initiative that sets standards for how internet companies should respond to illegitimate government requests that violate international human rights. Human rights principles not only constrain state actions but can also serve as a guide for government regulation of corporate behaviour. Any substantive discussion of transnational platform governance inevitably involves both national and international legal orders.

Could international law be used to mitigate digital linguistic inequalities and other human rights impacts experienced by platform users? Given the considerable power platforms have over the rights of their users, it is pertinent to ask what human rights obligations they have. As things stand, international law remains fundamentally state-centric. Historically, non-state actors like corporations have been largely excluded from direct human rights obligations. Human rights law has primarily focused on protecting individuals from abuses by states, and it places the duty to uphold rights squarely on governments. Nevertheless, there is growing recognition that corporations should also bear human rights responsibilities. On rare occasions, some companies, primarily in the extractive sector, have provided reparations such as financial compensation and guarantees of non-repetition to victims of human rights violations. The heavy reliance on states as the primary protectors of human rights poses significant challenges in a globalized and digitally mediated world, where harms often cross borders and where corporate practices frequently bypass or outpace national legal oversight.

Human rights obligations have been extended to non-state actors, but only through non-binding human rights instruments, which are sometimes called soft law because they lack legal enforceability. The United Nations Global Compact (UNGC), for example, articulates fundamental responsibilities that corporations have in sustainability and social responsibility. One of its core principles is that corporations should avoid complicity in human rights abuses. Similarly, the Guiding Principles on Businesses and Human Rights (UNGP) provides a framework for applying the rights enshrined in the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR) to corporate conduct. While the ICCPR and ICESCR are binding treaties for the states that have ratified them, the UNGP itself lacks the force of law.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Beyond the United Nations (UN), other intergovernmental organizations (IGOs) have also issued human rights guidance for corporations. The OECD, which is an IGO representing high-income countries, has developed guidelines for multinational enterprises on responsible business conduct,[190] which include human rights as a key area of business responsibility. The High Commissioner on National Minorities of the Organization for Security and Co-operation in Europe (OSCE) has also weighed in, issuing a series of recommendations that aim at preventing conflicts involving national minorities. Its 2019 recommendation focuses on the digital environment. The Tallinn Guidelines on National Minorities and the Media in the Digital Age[191] encourages state intervention to ensure that online media spaces are inclusive, safe, and diverse in viewpoints. These guidelines are directed at participating states rather than corporations.

As we have seen, Meta has set up its own quasi-judicial body to oversee speech moderation on their platforms. The Oversight Board voluntarily adopts the UNGP in its adjudication as a relevant standard, alongside the company's Community Standards and five values (voice, authenticity, safety, privacy, and dignity).[192] Specifically, the Board assesses whether restrictions to freedom of speech on the platform meet the requirements of legality,[193] legitimate aim, and necessity and proportionality as laid down in Article 19, para. 3, of ICCPR. Following the due diligence framework recommended by the UNGP, the Board has called on Meta to assess its moderation practices, for example, by conducting a human rights due diligence review on how it moderates content in Arabic and Hebrew.[194] Whether such initiatives arose from self-interest, such as preempting regulation, or corporate social responsibility, the voluntary implementation of UNGP demonstrates the standard-setting effect of soft law. Considering the miniscule volume of content moderation cases that the Board handles, some may criticize the voluntary implementation as performative as its scope of application remains very limited. Platforms can fundamentally improve their content moderation practices by adopting human rights law as a standard for all stages of their content moderation processes and enhance transparency and democratic participation in how they make rules and decisions. These standards would be much more transparent and defensible against national laws than rules that platforms come up with themselves.[195]

[190] mneguidelines.oecd.org/mneguidelines/. [191] www.osce.org/hcnm/tallinn-guidelines.

[192] Meta does not provide a clear definition of these values, but see the following web page for their elaboration: https://about.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards/.

[193] The principle of legality requires that rules used to limit expression be clear and accessible.

[194] See https://about.fb.com/news/2022/09/human-rights-impact-meta-israel-palestine/.

[195] Kaye, *Speech Police*.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

One area of the UNGP that has seen limited voluntary implementation is access to remedy, which is important because, even with the best intentions, not all human rights violations can be prevented. Despite acknowledging that its platform had been used to incite violence in Myanmar, Meta declined a request from Rohingya groups in Bangladesh to contribute USD $1 million towards an education project in refugee camps. For a company that earned USD $164.5 billion in 2024, the refusal reflects poorly on its commitment to accountability and raises questions about the seriousness with which it approaches reparative justice.[196]

The UNGP also requires companies to carry out human rights due diligence across all aspects of their operations, which has not yet become common practice. While the Oversight Board has acknowledged the human rights implications for minoritized language communities in some of its rulings, it has stopped short of recommending a comprehensive human rights assessment of how platforms' decisions on digital language support and broader language management affect linguistically diverse populations.

There is growing recognition that private actors, whose platforms serve as key venues for the exercise of civil and political rights, and whose decisions carry significant human rights consequences, lack sufficient accountability under existing human rights frameworks. Given the influence these companies exert over the public sphere, there is a compelling case for assigning them additional responsibilities, potentially treating them as a distinct class of public interest companies.[197] The international human rights regime could also evolve to establish mechanisms that allow individuals to seek remedy directly from such actors.

3.2.4 Civic Technology and Contextualized Development

Putting normative and regulatory pressure on platforms could potentially constrain private actors and mitigate inequalities, but not fundamentally change the political economy of our digital public fora. What real alternatives exist that can free us from both corporate profiteering and government overreach? One possibility lies in civic technologies: digital infrastructures designed to serve public interest, rather than commercial or state agendas.[198] These include open-source

[196] *Annual Report*, Meta Platforms Inc., 30 January 2025.

[197] Jørgensen, R. F. (2018) 'Human Rights and Private Actors in the Online Domain'. In Land, M. K. and Aronson, J. D. (eds.) *New Technologies for Human Rights Law and Practice*, 1st ed. Cambridge: Cambridge University Press, 243–69, https://doi.org/10.1017/9781316838952.011.

[198] Zittrain, J. (2009). *How to Get What We All Want*. Cato Institute. www.cato-unbound.org/2009/05/06/jonathan-zittrain/how-get-what-we-all-want/.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

platforms, cooperative networks, and peer-governed digital spaces that prioritize transparency, inclusivity, and democratic control. A prominent example known globally is Wikipedia, which is built and maintained largely by volunteers rather than profit-seeking entities.

But civic technologies face formidable challenges: they require strong institutional safeguards against bad actors, sustainable governance models, and a critical mass of users and contributors to remain viable. Even Wikipedia has not overcome the problem of global digital divide: most of its contributions come from a small minority of contributors, who unevenly distributed around the world. Its format and policies, such as the requirement that contributions need to be attributed to written sources, may not align with the epistemic traditions or communication norms of some communities.[199] Culturally sensitive governance designs are more likely to emerge when digital tools are localized and rooted in the languages and practices of specific communities. Yet, many speakers of digitally low-resourced languages also face scarcity in other domains: limited financial means, low literacy rates, and lack of technical capacity. These constraints make it difficult for them to develop or maintain localized digital infrastructures. This is where collaboration among governments, researchers, and civil society becomes vital. Such a strategy demands a long-term vision. In Africa, the grassroots organization Masakhane strengthens continent-wide research effort on natural language processing in African languages by building communities and sharing resources such as datasets and research findings.[200] The vast reach and entrenched dominance of global platforms discourage competition, making alternative models harder to sustain, but civic spaces do not always have to be built from scratch. Decentralized platforms such as Mastodon (a nonprofit) and Bluesky (a for-profit benefit corporation) provide infrastructure that allows communities to host their own servers, define local governance norms, and manage interactions according to their own values. These spaces may not rival dominant platforms in scale, but they offer important experiments in reimagining digital public space.

Language technologists have shown the value of participatory research in developing digital technologies for minoritized language communities. For instance, a study involving over 400 participants representing more than thirty low-resource African languages helped generate benchmarks for understudied language pairs and provided methods for evaluating machine translation outputs.[201] In another case, researchers collaborated with a Banjara farming community in India, whose language lacks a written script, to collect spoken

[199] Graham and Dittus, *Geographies of Digital Exclusion*. [200] See www.masakhane.io/.

[201] Nekoto, W., Marivate, V., Matsila, T., et al. (2020). 'Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages'. In *Findings of the Association for*

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

language data and develop an AI model capable of responding to spoken queries related to local agricultural practices.[202]

Aside from regulating global platforms operating within their countries, governments can create policies that will improve the conditions of their local languages and cultures in the digital environment. They can facilitate the digitalization of local languages, subsidize domestic content productions, and support the development of technologies that are attuned to local linguistic and cultural contexts. For example, the Department of Science and Technology in South Africa has funded a national network of human language technology researchers to strengthen collaboration in supporting the country's official languages in their digital transition.[203] In Spain, the government coordinated a top-down advocacy campaign to preserve the letter ñ by banning non-Spanish keyboards.[204]

Responsible development in digital technologies does not only require context sensitive solutions but also planning and resourcing across multiple spheres. Research from the Global South shows that rapid technological transfers from the North can aggravate social inequalities if this does not happen in tandem with other forms of development. For example, the arrival of social media has failed to create a vibrant public sphere in low-income economies in Africa, such as Zimbabwe, where persistent democratic and literacy divides mean that the technology has disproportionately benefited the elite.[205]

3.3 Linguistic Inequalities under Information Capitalism

It is no revelation that English dominates the digital realm. But this is not only because English is economically privileged and attracts more users in online exchanges. The powerful private actors who shape today's digital public sphere actively prioritize the interests of English speakers through platform design and governance. As a result, non-English speakers encounter a speech environment that is less safe, struggle more to be heard, and have limited access to platform

Computational Linguistics: EMNLP 2020, 2144–2160. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.findings-emnlp.195.

202 Reitmaier, T., Raju, D. K., Klejch, O. et al. (2024) 'Cultivating Spoken Language Technologies for Unwritten Languages'. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)* 614: 1–17. https://doi.org/10.1145/3613904.3642026.

203 Mlambo, R. and Matfunjwa, M. (2025) 'Human Language Technology Tools for Indigenous South African Languages and Their Potential Use', *Literator* 46(1): 1–11. https://doi.org/10.4102/lit.v46i1.2049.

204 Or, I. G. (2022) 'Typographical Advocacy in the Age of Digital Encoding'. *Language Policy* 21: 527–543. https://doi.org/10.1007/s10993-022-09619-2.

205 Moyo, L. (2020) 'The End of the Public Sphere: Social Media, Civic Virtue, and the Democratic Divide'. In Ragnedda, M. and Gladkova, A. (eds.) *Digital Inequalities in the Global South*. Cham: Palgrave Macmillan. 269–285. https://doi.org/10.1007/978-3-030-32706-4_13.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

processes and decision-making. At one end of the spectrum, content from some language communities is disproportionately censored; at the other, content in certain languages is scarcely moderated at all. Such linguistic inequalities are a facet of global structural inequalities in the political economy created by platforms.

In a book confronting biases in technology, data journalism professor Meredith Broussard reminds readers that problems in the tech sector are often dismissed as glitches – temporary and unexpected errors in need of a quick fix. She points out that many of these biases are not accidental but are embedded and structural.[206] They cannot be solved simply by tweaking code. In the same vein, it is not the advancement of information and communication technology itself, but the business decisions made by private actors that deepen inequalities across language communities. This distinction matters: the former naturalizes inequality as an inevitable by-product of progress and shifts the burden onto those left behind to catch up; the latter opens the door to institutional, political, and regulatory solutions.

There are two common rhetorical strategies that platforms use to publicly respond to content moderation failures. The first is the normalization of deviance: 'We are sorry, but mistakes happen.'[207] This narrative obscures the root causes of failure, deflects responsibility, and gradually desensitizes the public to rule violations, recasting them as routine rather than unacceptable. The second is the rhetoric of perpetual improvement: 'We will do better – our technology is improving.' This response is rooted in technochauvinism, the belief that technology is always the solution, regardless of the nature or complexity of the problem.[208]

Content moderation is inherently difficult because it involves navigating conflicting values. For instance, there may be a compelling public interest in sharing graphic content that documents human rights abuses, yet the same content can also pose serious risks of harm. Moderation is further complicated by users' ingenuity in evading controls, such as through the use of Algospeak – coded language designed to bypass automated moderation systems (such as using 'unalive' to refer to suicide or death, or the watermelon emoji for Palestinian resistance). However, these challenges should not be conflated with the persistent underperformance of moderation in certain languages, which often stems from chronic underinvestment. While perfection is not

[206] Broussard, M. (2023) *More than a Glitch: Confronting Race, Gender, and Ability Bias in*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/14234.001.0001.

[207] Halpern, S. (2 May 2021) 'Facebook and the Normalization of Deviance'. *The New Yorker*. www.newyorker.com/news/daily-comment/facebook-and-the-normalization-of-deviance.

[208] Broussard, M. (2019) *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/11022.001.0001.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

expected, platforms should be held to a baseline standard of content moderation across all of their content languages. They should also provide transparent, language-specific reporting that demonstrates how these standards are being upheld, as discussed earlier in this section.

The normalization of deviance has enabled a disregard for marginalized regions and populations who are not 'big on people's minds',[209] ignoring and invisibilizing their suffering. The way these transnational media corporations concentrate their resources on markets that their stakeholders care about is similar to how attention flows around the world. The public pay more attention to news happening in places that share geographic, linguistic, historical, or cultural similarities with themselves.[210] News value is also related to power relationships in a Euro-American centered world order. For example, the United States is heavily reported on by national media around the world due to its perceived news value, but the media in the United States do not pay nearly as much attention to the rest of the world.

The structure of the digital economy is reflected in its part in the global labour market. Platforms rely heavily on low-wage labour from different parts of the world to support their content moderation system and their AI technologies.[211] For example, Facebook employs content moderators working in call centres in the Philippines, Ireland, Mexico, Turkey, India, and Eastern Europe. Similarly, data labelling for training AI models is largely outsourced to annotators in the Global South, who do not only work long hours and receive low pay, but are also often exposed to graphic materials that affect their mental health in the long run.[212] Instead of being asked to contribute their linguistic and cultural knowledge, they are typically required to label data based on Western perspectives and in dominant languages.[213]

The political economy created by platforms contributes to global economic inequalities, which see cheap labour from the Global South powering the world economy,[214] while the benefits of production are reaped disproportionately by

[209] Halpern, 'Facebook and the Normalization of Deviance'.

[210] Kwak, H., An, J. and Salminen, J., et al. (2018) 'What We Read, What We Search: Media Attention and Public Attention among 193 Countries'. *Proceedings of the 2018 International World Wide Web Conference*, 893–902. https://doi.org/10.1145/3178876.3186137.

[211] These crowdworkers are recruited through agencies such as Amazon's Mechanical Turk, Appen, and Scale AI.

[212] Hao, K. and Hernández, A. P. (2022) 'How the AI Industry Profits from Catastrophe'. *MIT Technology Review.* 20 April. www.technologyreview.com/2022/04/20/1050392/ai-industry-appen-scale-data-labels/.

[213] For a comprehensive ethnography on commercial content moderators, including their work conditions and their geographical distribution, see Roberts, S. T. (2019) *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press.

[214] Hickel, J., Lemos, M. H., and Barbour, F. (2024) 'Unequal Exchange of Labour in the World Economy'. *Nature Communications* 15: 6298. https://doi.org/10.1038/s41467-024-49687-y.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

the Global North. Dal Yong Jin has called the asymmetrical relationship of interdependence between transnational technology companies and the majority of people and countries *platform imperialism*.[215] Imperialism has been reinvented through the digital ecosystem, which is primarily controlled by the US corporations, but with increasing competition from China. China, now a major technology powerhouse, has extended its influence by supplying technological infrastructure to developing countries through initiatives like the Digital Silk Road project. Although the rise of China has punctured the North–South divide, most countries in the Global South remain subservient to a digital world order characterized by US–China rivalry.

This new form of imperialism is no longer centred on nation states but transnational corporations, and capital accumulation is no longer based on territory as property but on intangible goods such as data and intellectual property. It is by and large a continuation of the North–South divide created by Western imperialism. Computer security specialist Bruce Schneier has gone as far as describing the relationship between users and platforms as 'more feudal than commercial': users provide free labour by generating data, which platforms package and sell for profit.[216] In this conception, users are not customers to platforms; they are the products that get sold to the real customers. Unlike older forms of imperialism, which relied on the coercive force of states, digital capitalism is arguably just as imperialistic, but it operates through soft power, versed in the discourse of consent, participation, and development.

3.4 Dilemmas of Linguistic Inclusion

It is often assumed that the presence of a minority language in a majority-language setting signals recognition and inclusion, and that translation practices advance linguistic justice by enabling cross-cultural communication. However, drawing on linguistic landscape studies of public signage, sociolinguist Philipp Angermeyer argues that inclusion can also be discriminatory, if such inclusion ties deviant behaviour to speakers of certain languages. In a striking example from a regional train in Germany, safety instructions are provided in German, English, French, and Italian, while the passenger bill of rights appears only in German. By contrast, a warning about fines for fare evasion is posted in German, English, French, and Turkish. Turkish speakers are thus implicitly marked as potential fare evaders, yet their rights and safety are not equally

[215] Jin, 'The Construction of Platform Imperialism in the Globalization Era'.
[216] Schneier, B. (2015) *Data and Goliath: The Hidden Battles to Capture Your Data and Control Your World*. New York: W. W. Norton. 58.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

addressed. Angermeyer terms this phenomenon punitive multilingualism.[217] A systematic feature of punitive multilingualism is that public order signs written in subordinated languages are often ungrammatical or unidiomatic, as the texts are regularly produced through machine translation without input from actual speakers.

The concept of *punitive multilingualism* can also be extended to the more complex realm of digital technologies. On social media platforms, making interfaces, rules, and processes available in multiple languages is generally seen as promoting access and improving procedural fairness. But digital inclusion is not an unqualified good. Often, this is because the existing structures into which communities are being included are themselves flawed or harmful. For instance, platforms that algorithmically amplify content based on engagement metrics tend to favour sensationalism over nuance – conditions that are often unfavourable to minoritized communities. In such cases, inclusion may reproduce or even exacerbate inequalities. Inclusion is therefore not always the solution, at least not until the broader conditions of participation are improved. As a former Facebook employee reflected in hindsight, Myanmar might have been 'far better off' if Facebook had never entered the country.[218]

Linguistic inclusion does not always improve conditions for the communities being included. Although inadequate content moderation threatens user safety, excessive moderation can stifle free speech. While it is necessary to moderate content in different languages, developing algorithms that detect speech transgression in some languages but not others may have the effect of limiting speech unevenly across language communities and amplifying some voices and suppressing others in a public debate. When a language community is made visible to platform management through irregularities, such as political or military conflicts, platforms might mitigate risks through shadow banning the whole community. This may reduce the spread of disinformation, but it can also silence legitimate voices, including those of citizen journalists. Content moderation is a double-edged sword: just as language can be both oppressive and emancipatory, both under-moderation and over-moderation can lead to harm.

Another reason why inclusion is not always emancipatory is the risk of miscommunication when the quality of translation is poor. The misuse of machine translation in legal and medical settings can cause serious harm; for

[217] Angermeyer, P. S. (2023) 'Translation as Discrimination: Sociolinguistics and Inequality in Multilingual Institutional Contexts'. *Language in Society*. 52(5): 837–59. https://doi.org/10.1017/S0047404522000422.

[218] Wynn-Williams, *Careless People*, at 359.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

example, an asylum seeker's credibility was undermined because their application had been generated using machine translation.[219] Applying such technologies to languages where accuracy is low can lead to unacceptable risks, especially when users are not warned about potential inaccuracies. Even when machine translation becomes more reliable, industrialized effort to develop technological 'solutions' to the 'problem' of multilingualism may ultimately narrow language practice and linguistic freedom in the long term.[220] Well-intentioned innovations can have unintended consequences: for example, people may feel less compelled to learn other languages if they believe AI will mediate all cross-linguistic communication. Meanwhile, to ensure effective multilingual transfer, technologies in cross-linguistic information retrieval and machine translation produce translingually controlled meanings at the expense of other variations in meanings that could create fuzzy ontologies for computer scientists. Meanings that are paralinguistic or untranslatable, or ones that are dependent on silence or implicature, are dispreferred.[221] In short, the desire to use technology to bridge immediate communication challenges may have longer term repercussions.

Moreover, overt inclusion and covert exclusion may happen to the same language community at the same time. For example, a platform may offer an interface in a given language but invest little in content moderation for that language, or it may moderate content without providing accessible appeal processes to its speakers. When users sign up to a platform, the range of language options in the user interface creates the illusion of access and choice. Presenting the language options as if they were equal performatively negates the language hierarchy without actually disrupting it. Visible forms of digital language support can obscure more critical but less visible forms of exclusion. Crucially, inflated claims about what big tech platforms can do for minority languages risk undermining local, community-led efforts to develop contextually appropriate solutions.[222]

Digital inclusion of a language is often implemented without involving its language communities in the process. A study of Guarani, an indigenous language co-official with Spanish in Paraguay, found that since 2006, speakers had developed a range of grassroots solutions to digitally encode the language. In 2013, Microsoft released an operating system with Guarani language

[219] Vieira, L. N., O'Hagan, M., and O'Sullivan, C. et al. (2021) 'Understanding the Societal Impacts of Machine Translation: A Critical Review of the Literature on Medical and Legal Use Cases'. *Information, Communication & Society* 24(11): 1515–1532. https://doi.org/10.1080/1369118X.2020.1776370.

[220] Gramling, D. (2021) *The Invention of Multilingualism*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108780667.

[221] Gramling, D. (2018) 'Supralingualism and the Translability Industry'. *Applied Linguistics* 41(1): 129–147. https://doi.org/10.1093/applin/amz023.

[222] Hadgu, A. T., Azunre, P., and Gebru, T. (2023) Combating Harmful Hype in Natural Language Processing. ICLR 2023. https://pml4dc.github.io/iclr2023/pdf/PML4DC_ICLR2023_39.pdf.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

support, followed by similar offerings from Facebook, Google, and Apple. While these developments may be hailed as milestones, they were introduced unliterally as an act of generosity without the advocacy, involvement, or feedback from speakers of the language.[223] Rather than being treated as collaborators, speakers were positioned as passive consumers of innovation, to whom language technologies could be marketed. Without engagement with language communities, digital inclusion could come across as corporate intrusion. For example, the Mapuche nation in Chile had threatened to sue Microsoft for releasing a version of its Windows operating system in their native language Mapudungun without their permission, and considered it an act of 'intellectual piracy' and a violation of their right to self-determination.[224]

One important reason why the absence of a language online should not be readily equated with its exclusion is the consideration of speaker agency. A study of minority languages in the Nordic countries found that while these languages have a limited presence on websites and in digital archives, interviewed speakers show little desire for their languages to have more extensive presence on the internet.[225] Digital tools such as online archives have certainly supported language documentation and revitalization in some contexts, for example, by providing access to learning materials, but not all communities share the same priorities. Some Indigenous cultures, such as the Māori in New Zealand, have refused to participate in digital libraries, because their knowledge traditions do not support the free circulation of information.[226] For many communities, digital 'ascendance' is not necessarily seen as a goal; instead, language revitalization is often rooted in relationships with people, place, and land.[227] As discussed earlier, the same digital tools used to support a language can also be deployed for surveillance, posing serious risks to marginalized groups facing displacement, criminalization, or state violence.[228] Ultimately,

[223] Or, 'Typographical Advocacy in the Age of Digital Encoding'.

[224] Long, G. (2007) 'Chilean Mapuches in Language Row with Microsoft'. *Reuters*. 9 January. www.reuters.com/article/business/chilean-mapuches-in-language-row-with-microsoft-idUSN22384122/. See also Mapuche letter to Bill Gates, www.mapuche.nl/english/05aug.htm. The author would like to thank an anonymous reviewer for relating this story.

[225] Dymet, M. (2019) 'Digital Language Divide in the European High North: The Level of Online Presence of Minority Languages from Northern Finland, Norway and Sweden'. *The Yearbook of Polar Law Online* 10(1): 245–74, https://doi.org/10.1163/22116427_010010012.

[226] Paolillo, J., Pimienta, D., and Prado, D. (2005) 'Measuring Linguistic Diversity on the Internet'. *United Nations Educational, Scientific and Cultural Organization*. https://unesdoc.unesco.org/ark:/48223/pf0000142186.

[227] Bird, S. (2020) 'Decolonising Speech and Language Technology'. *Proceedings of the 28th International Conference on Computational Linguistics*, 3504–3519. https://doi.org/10.18653/v1/2020.coling-main.313.

[228] Zuagg, I. (2024) 'Language Justice in the Digital Sphere'. In Liu, L. H. and Rao, A. with Silverman, C. A. (eds.) *Global language Justice*. New York: Columbia University Press. 244–273. https://doi.org/10.7312/liu-21038-011.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

speakers of a language may choose to preserve, abandon, or develop their languages, even though they may be acting under influence that is outside of their control.[229] Outside experts sometimes carry assumptions that clash with the perspectives of language communities, such as by viewing their language as data rather than as a connection to their land and culture, or believing that endangered languages should be documented and maintained in the purest form possible. Conflicting language ideologies may also develop within language communities. Facebook and Microsoft engaged young K'ichee' speakers to translate its platforms into the Mayan language, without seeking approval from the local governing body or consulting elders, who are traditionally regarded as authoritative custodians of the ancestral language. The young people view their language proficiency as a personal skill that can be commodified in the market, while older generations see the language as a sacred ethnic artefact that belongs to the community as a whole.[230] Linguistic inclusion by transnational corporations could therefore become a site of intergenerational ideological struggle.

Finally, it is important to remember that whether linguistic inclusion is experienced as emancipatory or punitive often depends on perception – and may be quite disconnected from the intentions of those who design the technology. Software localization, for example, is typically pursued as a market expansion strategy to increase profit, yet it may be interpreted by end users as a gesture toward bridging the digital divide. This gap between design intent and user interpretation reflects the concept of *interpretive flexibility*, the idea that technological artefacts are culturally constructed and can be understood differently by those who create them and those who use them.[231]

3.5 Competition in the Global Linguistic Market

Beyond its impact on individual users and non-users, globalized digital media affect the value of languages at a macroscopic level, with significant implications for the sociolinguistic experiences and life chances of their speakers and future generations of speakers. Pierre Bourdieu's sociology of social stratification, alongside its key concepts of field, capital, and habitus, offers a useful analytical framework. A field, or market, is a space in which actors, whose relations are defined by their access to resources, or capital (which may be economic, social, cultural, political, etc.), compete with one another based on

[229] Kulick, D. (1997) *Language Shift and Cultural Reproduction*. Cambridge: Cambridge University Press.

[230] Romero, 'Bill Gates Speaks K'ichee'! The Corporatization of Linguistic Revitalization in Guatemala'.

[231] Pinch, T. J. and Bijker, W. E. (1984) 'The Social Construction of Facts and Artefacts'. *Social Studies of Science* 14(3): 399–441. https://doi.org/10.1177/030631284014003004.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

rules and logic specific to the field.[232] Their behaviour is guided or constrained by predispositions, or habitus, such as their capacities and preferences. Linguistic practices are thus a product of the interaction between a linguistic habitus and a linguistic market;[233] for example, one may adopt a particular style or language variety in their linguistic repertoire based on its anticipated value in a market. By viewing language as a form of cultural capital and a component of social capital, Bourdieu reframes it as a means of action and power, rather than merely a tool for communication.

In Bourdieusian terms, one can therefore speak of information capital or digital capital that circulates in the digital field and conceptualize the digital divide as the uneven distribution of such capital. The digital field is dominated by the American big five (Google, Apple, Meta, Amazon, and Microsoft) and their Chinese counterparts (Alibaba, Tencent, ByteDance, etc.), who command enormous digital capital. These corporations are not only among the largest holders of personal data globally but also provide the foundational infrastructure on which countless other digital companies and content creators depend. Their dominance reinforces structural inequalities. For example, Google's AdWords has been described as a global, real-time, multilingual auction system for advertising placement – one that effectively constitutes a global linguistic market. In such a market, advertisers are incentivized to prioritize more profitable languages, reinforcing linguistic hierarchies and contributing to the marginalization of less economically valuable ones.[234]

Likewise, social media can be understood as a subfield within the broader digital field, where content creators seek to accumulate both economic and symbolic capital, while platforms extract and convert user activity into digital capital through data collection and analysis, in a digital economy financed by advertisers.[235] In this subfield, content creators compete for visibility, engagement, and monetization, just as platforms compete for user attention and market dominance. This competitive structure is shaped by the platforms themselves, which design the 'rules of the game'. As discussed earlier, these rules are often calibrated in ways that disadvantage users from minoritized language

[232] Ignatow, G. and Robinson, L. (2017) 'Pierre Bourdieu: Theorizing the Digital'. *Information, Communication & Society* 20(7): 950–966. https://doi.org/10.1080/1369118X.2017.1301519.

[233] Bourdieu, P. (1991) *Language and Symbolic Power*. Cambridge: Polity Press.

[234] Graham, R. (2017) 'Google and Advertising: Digital Capitalism in the Context of Post-Fordism, the Reification of Language, and the Rise of Fake News'. *Palgrave Communications* 3(45): 1–19. https://doi.org/10.1057/s41599-017-0021-4.

[235] Verwiebe, R. and Hagemann, S. (2024) 'Bourdieu Revisited: New Forms of Digital Capital – Emergence, Reproduction, Inequality of Distribution'. *Information, Communication & Society*, https://doi.org/10.1080/1369118X.2024.2358170.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

communities, thereby further devaluing peripheral languages and reinforcing existing linguistic hierarchies.

Linguistic products, such as literary works, that once circulated within the enclosed boundaries of local markets were to some extent protected from external competition. Similar to how state formation creates the conditions for a unified linguistic market, usually dominated by an official language within the state, digital media platforms have forged a global linguistic market, in which content from different corners of the world are put in much more direct competition with one another than before. Although these platforms have lowered the cost of publishing in minoritized languages, they embed incentives to prioritize language practices that maximize engagement, putting pressure on minoritized language practices.

Revisiting De Swaan's framework of global language hierarchies (see Section 1.2.1), we are reminded that the communication potential of a language is a strong motivation for learning a foreign language; as a result, language learning occurs mostly upward in the hierarchy. Few peripheral language speakers can afford to remain monolingual, but monolingualism is common among supercentral and hypercentral language speakers. Speakers of peripheral languages tend to have more reasons to be interested in what speakers of more central languages have to say, rather than the other way round. Translation therefore typically flows downward, such as from hypercentral to central;[236] when a peripheral language needs to be translated into another, it is sometimes bridged through a more central language. The use of a bridging language is known as pivot translation. For example, the European Union translates across its twenty-four official languages through using English, French, and German as pivot languages.

Users often use vernacular languages on the platforms to connect and socialize with their local or diasporic communities. However, peripheral language speakers may switch to a more central language when accessing platforms not only to be part of a potentially less hostile speech environment but also to access a wider range of content and to improve their chance of success in capital-generating activities. This practice, known as digital diglossia, is most commonly observed among internet users whose first languages occupy relatively low positions in the global language hierarchy (see relevant discussion in Section 1.2.3 and Section 2.1). These users are typically members of the educated elite in their societies, having acquired bilingual proficiency in both their local language and a more dominant, central language. Their command of

[236] Heilborn, J. (2010) *Structure and Dynamics of the World System of Translation*. *International Symposium on Translation and Cultural Mediation*. Paris: UNESCO.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

a central language enables them to navigate platforms more effectively. Yet this linguistic flexibility also presents a dilemma in content production. Using a central language may allow creators to reach wider audiences, but it also places them in competition with a much larger pool of content producers. Conversely, producing content in a peripheral language may limit their reach to a smaller, more niche audience, though with the advantage of less competition for visibility and attention.

Does digital diglossia lead to language shift? Although Ferguson argues that diglossia could remain stable,[237] Penelope Eckert contends that, in most circumstances, it continually evolves.[238] Using the displacement of Gascon by French as an example, she observes that as domains associated with the low variety become obsolete, the language can be stigmatized as that of those who had been 'left behind', while the domains of the high variety are exclusively tied to social and economic survival and success. In the case of digital diglossia, as digitalization permeates our daily lives, the domains that require the adoption of digitally thriving languages continue to expand. Social media platforms also foster language contact, which is the primary driver for language shift. Digital diglossia is therefore likely to incentivize language shift. Rather than remaining stable, individual bilingualism often represents a transitional phase in the process of language shift, especially when the diglossic functions are poorly compartmentalized and when bilingualism is widespread rather than being confined to a small, privileged class.[239]

Language is a hyper-collective good – 'the more people use it, the more valuable it becomes'.[240] Speakers of majority languages benefit from a location rent, profiting whenever an outsider learns their language and pushes its communication value up without effort on their part.[241] Language is a valuable commodity in many former colonial powers and their settler colonies, which export educational and cultural products worldwide. Education-related exports alone were valued at $39.7 billion for the United Kingdom in 2022, $50 billion for the United States in 2024, and $33 billion for Australia in 2024. This has led some to conclude that the global dominance of English is neocolonial or imperial.[242]

[237] Ferguson, 'Diglossia'.
[238] Eckert, P. (1980) 'Diglossia: Separate and Unequal'. *Linguistics* 18: 1053–1064.
[239] Fishman, J. A. (1967) 'Bilingualism with and without Diglossia; Diglossia with and without Bilingualism'. *The Journal of Social Issues* 23(2): 29–38. https://doi.org/10.1111/j.1540-4560.1967.tb00573.x.
[240] Grin, F. (2006) 'Economic Considerations in Language Policy'. In Ricento, T. (ed.) *An Introduction to Language Policy: Theory and Method*. New York: John Wiley & Sons. 81.
[241] De Swaan, 'Language Systems', at 64.
[242] Phillipson, R. (2013) 'Linguistic Imperialism'. In Chapelle, C. A. (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell. 3470–76. https://doi.org/10.1002/9781405198431.wbeal0718.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

The flip side of the coin is that the fewer people use a language, the less value it is perceived to have. As Bourdieu points out, the dominated are often complicit in their own subjugation, as they are susceptible to market forces. Speakers of peripheral language varieties collaborate in the devaluing of their languages, when they switch to a dominant language to improve their life chances. If language shift does happen with many speakers, which causes the demise of a language, it may be considered a tragedy of the commons, where individuals acting in their own self-interest lead to the depletion of a shared resource.

Bourdieu emphasizes the relationality and conflicts among actors in a field. The fact that linguistic products on social media platforms now compete in a global linguistic market, where infrastructure and material resources are accorded to a small number of dominant languages much more than others, positions minoritized language speakers, who constitute a global majority, as a digital underclass. This is highly significant to our understanding of digital linguistic inequalities, as well as digital inequalities more generally, and of the limitations of prevailing conceptual frameworks. In theory, digital divides can be bridged; the digitally excluded can become included; digital language support can be enhanced. However, neither digital divide nor digital exclusion pay sufficient attention to relationality in the digital field: the digital underclass are unlikely to emerge as winners in a global competition for attention. If they must compete globally for content moderation resources, they are unlikely to enjoy adequate protection from harm.

Under information capitalism, knowledge and information have increasingly become commodified – transformed from public goods into marketable assets. This commodification goes beyond economic transactions, influencing how core social goods such as due process, privacy, and human rights responsibilities are managed. Consequently, these foundational principles risk being subordinated by market logics such as efficiency, profitability, and platform metrics, rather than by legal obligations or democratic accountability.

3.6 Rethinking Linguistic (In)justice

Although many linguists who write about language-based injustices are concerned with the survival of minoritized languages, it is important to note that the consequences of these injustices are not only linguistic in nature. Discrimination based on language can hurt one's ability to secure housing and employment. Linguistic barriers to justice lead to lack of access to conflict resolution and judicial remedies. As the current study shows, gaps in digital language support and inequalities in content moderation and recommendation

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

can affect not only opportunities to participate in the digital economy but also the safety and well-being of platform users and non-users.

Sociolinguistic approaches to rectifying linguistic injustices is predominantly framed in the discourse of language rights.[243] Advocacy in this area often sees entrenchment of rights into domestic law as the ultimate goal; some assert that fundamental language rights should become linguistic human rights (LHR).[244] Currently, language-related rights embedded in international and regional human rights instruments are mostly tolerance-oriented and focus on individual rights grounded in principles of non-discrimination and fair trial; a narrower set of group-based minority rights also exists.[245] LHRs present a more radical proposal to address linguistic inequality by pushing for the extension of rights typically enjoyed by majority language speakers, such as the right to receive basic education in one's mother tongue and the right to use it in official contexts, to minoritized language speakers. These are promotion-oriented rights that require resourcing from the state. Despite its ambitions, advocacy for LHRs has gained limited traction. States are wary that if they supported minoritized language groups, they might also be sponsoring sub-state nationalism. When national law grants official status to minoritized languages, such law tends to emphasize political unity and have symbolic significance for language communities, but rarely translating into substantive rights with emancipatory power.[246] Fundamentally, the rights model positions the state as the primary actor, whether as aggressor or protector, while often overestimating its political will and economic capacity to fulfil the envisioned rights. This model has limited utility in the digital realm, where private actors have a dominating presence, as rights are primarily claimed against the state.

As an alternative to the rights approach, linguist Christopher Stroud has proposed an idea called linguistic citizenship, which encourages speaker agency and the development of political subjectivities beyond the bounds of existing power structures.[247] Although the emphasis on the agency of individuals is

[243] May, S. (2018) 3. Commentary – Unanswered Questions: Addressing the Inequalities of Majoritarian Language Policies' In Lim, L., Stroud, C., and Wee L. (eds.) *The Multilingual Citizen: Towards a Politics of Language for Agency and Change*. Bristol: Multilingual Matters. 65–72. https://doi.org/10.21832/9781783099665-006.

[244] Skutnabb-Kangas, T. and Phillipson, R. (2010) *Linguistic Human Rights: Overcoming Linguistic Discrimination*. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110866391.

[245] Leung, J. (2018) 'Language Rights'. In Visconti, J. (ed.) *Handbook of Communication in the Legal Sphere*. Berlin: DeGruyter. https://doi.org/10.1515/9781614514664-006.

[246] Leung, J. (2019) *Shallow Equality and Symbolic Jurisprudence in Multilingual Legal Orders*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780190210335.001.0001.

[247] Stroud, C. (2018) 'Linguistic Citizenship'. In Lim, L., Stroud, C., and Wee, L. (eds.) *The Multilingual Citizen: Towards a Politics of Language for Agency and Change*. Bristol: Multilingual Matters. 17–39. https://doi.org/10.21832/9781783099665-006.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

laudable, the strength of linguistic citizenship is also its weakness: by shifting the arena of contestation to spaces where community stakeholders can exercise full control of their linguistic practices, the framework does not present itself as a candidate that can be used to confront the structures that directly produce and reproduce social inequalities.[248] A more practical proposal was put forward by Docrat and Kaschula, which does not replace but supplement the rights approach.[249] Observing the limited success of implementing constitutional language rights in South Africa, Docrat and Kaschula propose applying the meaningful engagement framework, developed in the sphere of socioeconomics rights, to the implementation of language policy. This approach emphasizes effective consultation and mediation with stakeholders and requires that solutions be shaped by the views of those directly affected through democratic processes. Crucially, such engagement must take place in their own languages to ensure genuine participation.

Political philosophers have contemplated the normative question of how public institutions should handle societal multilingualism, particularly how state governments balance the demands of nation-building with language rights and accommodations for minority languages. While establishing a common language for public discourse may foster deliberative democracy, selecting a single language can also be perceived as exclusionary and unjust within a multilingual society. The dilemma is ultimately concerned with the relationship between governments and the governed. Social contract theory presumes that individuals consent to surrender some autonomy and liberty in exchange for security and a cooperative system of resource distribution. However, since social contract theories do not easily apply to private entities, this body of literature offers limited guidance on how private actors should navigate platform multilingualism. What is noteworthy for the purpose of our current discussion is that this literature seems to be moving away from the rights discourse. A volume edited by Will Kymlicka and Alan Patten suggests approaching linguistic justice not from the perspective of outcomes, such as language preservation or nation building, but procedures.[250] This approach aligns with recommendations from this Element: procedural safeguards that

[248] Wee, L. (2018) 'Essentialism and Language Rights'. In Lim, L., Stroud, C. and Wee, L. (eds.) *The Multilingual Citizen: Towards a Politics of Language for Agency and Change*. Bristol: Multilingual Matters. 65–72. https://doi.org/10.21832/9781783099665-006.

[249] Docrat, Z. and Kaschula, R. (2015) '"Meaningful Engagement": Towards a Language Rights Paradigm for Effective Language Policy Implementation'. *South African Journal of African Languages* 35(1): 1–9. https://doi.org/10.1080/02572117.2015.1056455.

[250] Kymlicka, W. and Patten, A. (eds.) (2007) *Language Rights and Political Theory*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780199262908.001.0001.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

will guarantee a minimum standard of treatment across languages supported by a platform.

Scholars working in intersections between language and economics have attempted to tackle the challenge of balancing between the economic advantages of having a limited number of standardized languages and the undesirable consequences of disfranchising minoritized language communities in developing language policies. Using communication in the EU as a case study, Ginsburg and Weber propose that the selection of core languages for communication should take into account linguistic proximity between groups and between languages.[251]

In sum, the linguistic (in)justice literature has largely been developed as a response to the linguistic homogeneity imposed by modern states following 19th century European nationalism. It has given limited consideration to new forms of linguistic injustices in the digital arena, as outlined in this Element. Nationalist governments often seek to unify a country through promoting a majority language, and may see minoritized language communities as impediments to nation-building. By contrast, corporations typically hold no animosity towards minoritized languages – they simply do not care. Their engagement with language policy is minimal and driven solely by commercial or legal necessity. As such, the linguistic inequalities documented in this Element are not the result of deliberate exclusion, but rather collateral damage inflicted as platforms trample their way to profit.

Some insights from the existing literature could potentially be transposed to the digital context, such as the emphasis on meaningful engagement, the analytical shift towards procedural justice, and prioritization methods in language policy. Overall, however, our conceptual toolkit will need to be expanded to tackle the rise of private powers in mediating public discourse, which is exercised largely through soft power. This task calls for multidisciplinary approaches grounded in the political economy of language, with a willingness to engage with government, law, and markets. Discussions in this Element invite reimagining of what linguistic injustice entails and what meaningful remedies may look like.

3.7 Conclusion

Platforms embody contradictory forces: they at once connect and isolate; they enable as well as suppress; they equalise yet also reinforce hierarchies; they democratize access while concentrating power; they promise freedom but

[251] Such as Ginsburgh, V. and Weber, S. (2011) *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton: Princeton University Press.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

impose control. Their technologies can be both friend and foe to marginalized language communities. With billions of users worldwide, platforms today do not only produce technologies, they also govern who may speak, to whom, and how.

This Element identifies platforms as a key site of contemporary struggles over language choice and linguistic revitalization, global linguistic competition, and language ideologies. It underscores the specific risks and harms arising from how platforms govern language-related matters, and calls for urgent intervention. Platforms like to tout technology such as AI as the catch-all solution to challenges ranging from content moderation to digital language support. However, technological breakthroughs require time, resources, and luck, and greater reliance on AI will likely diminish, rather than enhance, transparency and public accountability. Platform policies and processes can be reconfigured today, if the will is there.

On a global platform like Facebook, the majority of users participate in a language environment where content moderation efforts are heavily under-resourced compared to that available to English speakers. Platforms may not warn them when they come across disinformation or disturbing content; they may not be told what rules they need to follow; their content may be removed or violating content may remain because the content moderation system fails to understand their language; they may not be told why their post is removed or how they may challenge the decision; and their voice may not receive much of an audience when dominant voices are amplified.[252] These linguistic inequalities are much more complex than what is typically captured by discussions of the digital language divide or digital language support, and have received scarce attention in the literatures of digital inequalities and linguistic injustices.

Despite positioning themselves as conduits, platforms exercise control over content and make governance decisions that shape the conditions for speech. Even though this Element foregrounds linguistic inequalities, it does not posit that linguistic equality is possible. Following Piller,[253] it sees linguistic equality not as absolute equality among languages but as the overcoming of linguistic inequalities, by distributing opportunities and resources as fairly as possible. For example, while one cannot expect every language spoken in a country to enjoy official language status, linguistic accommodation through quality interpreting can be provided to mitigate barrier to justice for those who do not speak

[252] Even English speakers may experience some of these struggles, such as when platforms fail to label problematic content consistently or offer detailed explanations of their content removal decisions. However, the fact remains that the challenges are much bigger for minoritized language speakers.

[253] Piller, *Linguistic Diversity and Social Justice*.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

the language of the court. The Element uses linguistic inequalities to talk about the disproportionate risk, neglect, and inability to access platform processes that some language communities but not others experience on digital platforms. It advocates for a baseline of multilingual management not drawn by the market (see Section 3.1.1), as well as the decentralization of platform governance (see Section 3.1.2), using a multi-pronged approach (as detailed in Section 3.2). These proposals aim to give marginalized language communities a genuine chance to flourish online and to reduce their vulnerability to harm. They should not have to compete for visibility in a global neoliberal economy simply to receive basic protections, due process, and opportunities to be heard.

Linguistic inequalities on platforms are a feature, not a bug, of digital capitalism. Although many of these inequalities can be mitigated to some extent, assuming sufficient goodwill, it is difficult to find comfort in the reality that most of our online interactions, whether with friends and family or in relation to government communication and public debate, now take place on private platforms driven solely by profit; that we practice our civil and political rights through private actors who bear no accountable human rights responsibilities; that digital infrastructure and governance frameworks developed in the Global North are transplanted to the Global South to perpetuate economic dominance over the global majority. This Element has shown that language offers a powerful lens through which we can visibilize these deeper systemic problems, which are pertinent to our fundamental rights and freedoms. Language can provide a point of intervention, given its close ties with knowledge, participation, representation, and global redistributive justice. While certain impacts of linguistic inequalities can be mitigated through improving platform policies and processes, tackling these larger problems requires innovations in international law, the imposition of effective legal constraints on private power, a model of digital governance that gives voice to different stakeholders, and alternative visions of a digital public space.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Acknowledgements

Research for this Element was made possible by several grants and fellowships. I am grateful for the generous funding provided by the Hong Kong Research Grants Council (General Research Fund Project number 17613217) and the School of English at the University of Hong Kong. The Humanities and Social Sciences Prestigious Fellowship (Hong Kong) and the Luce East Asia fellowship at the National Humanities Center (USA) supported the relief from my regular duties to conduct research for this project.

I would like to thank all the people who have provided feedback on earlier versions of this work, including colleagues and friends at the University of Hong Kong and Wilfrid Laurier University, and attendees of seminars/conference presentations at McMaster University, Uppsala University, Central China Normal University, the Canadian Symposium on Language and Law, Canadian Law and Society Association, the Annual Meeting of the American Political Science Association, and Global Summit in Constitutionalism. Special thanks to Don Kulick and Ana Deumert for their insightful suggestions at various stages of my writing. The helpful comments from two anonymous reviewers of the manuscript are also appreciated. The project has benefited from the research assistance of Emily Tsz Ching Choi, Kelly Yung Chun Li, Carolyn Ching Yin Liu, Jeremy Vyn, and Eva Sze Chin Yu.

As always, I am indebted to Joseph Yang who served as an invaluable sounding board when I worked through my arguments and provided technical support whenever I needed it.

I dedicate this Element to my parents, whose evolving relationship with technology has not only brought confusion and amusement, but also inspired reflection.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Cambridge Elements

Forensic Linguistics

Tim Grant

Aston University

Tim Grant is Professor of Forensic Linguistics, Director of the Aston Institute for Forensic Linguistics, and past president of the International Association of Forensic Linguists. His recent publications have focussed on online sexual abuse conversations including *Language and Online Identities: The Undercover Policing of Internet Sexual Crime* (with Nicci MacLeod, Cambridge, 2020).

Tim is one of the world's most experienced forensic linguistic practitioners and his case work has involved the analysis of abusive and threatening communications in many different contexts including investigations into sexual assault, stalking, murder, and terrorism. He also makes regular media contributions including presenting police appeals such as for the BBC Crimewatch programme.

Tammy Gales

Hofstra University

Tammy Gales is Professor of Linguistics and the Director of Research at the Institute for Forensic Linguistics, Threat Assessment, and Strategic Analysis at Hofstra University, New York. She has served on the Executive Committee for the International Association of Forensic Linguists (IAFL), is on the editorial board for the peer-reviewed journals Applied Corpus Linguistics and Language and Law / Linguagem e Direito, and is a member of the advisory board for the BYU Law and Corpus Linguistics group. Her research interests cross the boundaries of forensic linguistics and language and the law, with a primary focus on threatening communications. She has trained law enforcement agents from agencies across Canada and the U.S. and has applied her work to both criminal and civil cases.

About the Series

Elements in Forensic Linguistics provides high-quality accessible writing, bringing cutting-edge forensic linguistics to students and researchers as well as to practitioners in law enforcement and law. Elements in the series range from descriptive linguistics work, documenting a full range of legal and forensic texts and contexts; empirical findings and methodological developments to enhance research, investigative advice, and evidence for courts; and explorations into the theoretical and ethical foundations of research and practice in forensic linguistics.

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press

Cambridge Elements

Forensic Linguistics

Elements in the Series

Online Child Sexual Grooming Discourse
Nuria Lorenzo-Dus, Craig Evans and Ruth Mullineux-Morgan

Spoken Threats from Production to Perception
James Tompkinson

Authorship Analysis in Chinese Social Media Texts
Shaomin Zhang

The Language of Romance Crimes: Interactions of Love, Money, and Threat
Elisabeth Carter

Legal-Lay Discourse and Procedural Justice in Family and County Courts
Tatiana Grieshofer

Forensic Linguistics in China: Origins, Progress, and Prospects
Yuan Chuanyou, Xu Youping and Lu Nan

Decoding Terrorism: An Interdisciplinary Approach to a Lone-Actor Case
Julia Kupper, Marie Bojsen-Møller, Tanya Karoli Christensen, Dakota Wing, Marcus Papadopulos and Sharon Smith

Forensic Linguistics in Southern Africa: Origins, Progress, and Prospects
Russell H. Kaschula, Monwabisi K. Ralarala, Eliseu Mabasso, Zakeera Docrat, Wellman Kondowe and Paul Svongoro

Trust, Discourse, and Corporate Corruption: The Case of Enron
Matteo Fuoli, Adam Nix, Alicia Wickert and Annina Van Riper

Doxxing Discourse
Carmen Lee

Digital Multilingualism and Platform Governance
Janny H. C. Leung

A full series listing is available at: www.cambridge.org/EIFL

https://doi.org/10.1017/9781009263177 Published online by Cambridge University Press