Wasafiri

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rwas20

Translation as Culture in the Age of the Machine

Naomi Wells

To cite this article: Naomi Wells (2022) Translation as Culture in the Age of the Machine, Wasafiri, 37:3, 77-80, DOI: 10.1080/02690055.2022.2067244

To link to this article: https://doi.org/10.1080/02690055.2022.2067244

TRANSLATION AS CULTURE IN THE AGE OF THE MACHINE

NAOMI WELLS

Until recently, the idea of machines instantaneously translating across multiple languages was the preserve of fantasies about the distant future. Now, however, rapid developments in artificial intelligence (AI) have meant that machine translation has become an increasingly mundane feature of our online lives. Many of us do not give a second thought to the automatic translation of words that we compose and share on social media, a process that often takes place without our explicit knowledge, let alone our consent. But despite its seeming banality, machine translation raises critical ethical and cultural questions, the stakes of which are rarely considered within a tech industry that deems translation to be little more than a neutral, instrumental solution to the challenge of communicating across languages.

Across the tech world, multilingualism is most often viewed as a technical 'problem' to be solved. Writing in 2017 on Facebook's use of AI to 'deliver better language translations', Mark Zuckerberg noted that 'Throughout human history, language has been a barrier to communication. It's amazing we get to live in a time when technology can change that' (qtd in Edmond). Zuckerberg's reduction of linguistic difference to only the obstacles it can pose demonstrates a profound ignorance of the cultural significance of language, and relatedly a major underestimation of the task of the translator, whether human or machine. As Michael Cronin notes, the instantaneous nature of machine translation further feeds into the impression that it is merely an act of language transfer 'akin to the automated sub-routines of digital processing' (47).

The recent damning allegations about Facebook's failures to protect its users illustrate the destructive consequences of Zuckerberg's simplistic, techno-utopian approach to languages. As revealed in the documents released by whistleblower Frances Haugen, one of the reasons the platform contributes to fuelling ethnic conflict in places like Ethiopia and Myanmar is its lack of human moderators for content in languages other than English and more globally dominant languages (Hao, 'The Facebook Whistleblower'). This underinvestment would appear to stem from an overestimation of the effectiveness of purely technical solutions to language-related challenges. In her deposition, Haugen stated that such an overreliance on AI for content moderation in different languages was 'doomed to fail'. Especially in contexts of heightened ethnic conflict, it seems obvious that human moderators with an understanding of not just the language but also the sensitive cultural and social dynamics at play would be vital. The case of Facebook highlights some of the harms that can result when languages are viewed as 'neutral' tools of communication detached from the environments in which they are spoken.

The context of a text is therefore crucial, yet as machine translation becomes embedded in social media platforms, this is increasingly subject to rapid change. Online written materials are more and more able to travel across languages in ways unforeseen by their original authors, sometimes taking on new and unexpected meanings for their readers. The introduction of texts into new settings has always been a consequence of translation, and one often celebrated for bringing literature and other cultural artefacts to new audiences. However, a key distinction in the way platforms like Facebook automatically translate content is the loss of human agency in the decision both to be translated and to undertake an act of translation. Anyone who has attempted translation knows that it is typically painstaking and meditative, requiring thoughtful consideration of both the original author's intent and the situations and backgrounds of new readerships. A human engaged in the translation of a politically or culturally sensitive text, for example, would be likely to exert an element of caution, or perhaps even decide the potential to cause harm prevented them from translating at all.

Routledge
Taylor & Francis Group

WASAFIRI / VOLUME 37 / NUMBER 3 / ISSUE 111 / SEPTEMBER 2022 / PAGE 77–80 / © 2022 WASAFIRI
HTTPS://DOI.ORG/10.1080/02690055.2022.2067244

In contrast, regulated only by what we know is a highly flawed process of moderation, Facebook and other platforms are instantaneously translating and publishing vast quantities of content in multiple languages with no regard for how translated material may be taken up or received across different social and cultural contexts.

This connects to a concept often discussed in online research known as 'context collapse', whereby digital content quickly travels across and consequently flattens or 'collapses' audiences and their environments (Marwick and boyd). This means that the original author is less able to predict or control how their message may be received and by whom, which can potentially contribute to online conflicts and misunderstandings. Although not yet addressed in research, instantaneous translation on social media platforms undoubtedly has the potential to heighten the effects of context collapse, by making it possible for texts to travel even further and more rapidly into new languages and settings. Given the risks and potential harms caused by different audiences responding to or taking up social media messages in ways unintended by the original author, the lack of human agency in the acts of translating and disseminating these messages to potentially vast global audiences is a major concern. The ongoing limitations of translation technologies, which mean many machine-translated texts are immediately recognisable as 'flawed' translations, may lessen this risk by making the reader more aware of the potential for error and misinterpretation. However, as these technologies become more sophisticated and thus less visible, the possibility that readers will see such translations as pure reflections of the original author's views and intentions increases the danger of misunderstanding, conflict, and the rapid spread of misinformation.

Nevertheless, despite the continued importance of human mediators in the task of translation, it is simplistic to suggest we should simply do away with translation technologies. Even leaving aside that it is arguably too late to turn back the tide, we must acknowledge that machine translation can still be a valuable tool. The pandemic has highlighted some of the uses of these technologies, such as facilitating the rapid circulation of vital public health messages across language communities.[1] For example, while continuing to rely on its large community of human volunteer translators, the Translators without Borders NGO (non-governmental organisation) has been collaborating with tech companies and universities to develop translated datasets for key Covid-19 terms and phrases. However, this initiative has also highlighted that inequality between languages poses a further challenge, as the state of development of technology to support machine translation varies wildly.

Such disparities compounded the language gaps in Facebook's approach to content moderation, with the platform's heavy reliance on AI apparently failing to account for the lack of access to effective technologies for the majority of the world's languages. In fact, the widely reported resignation of the computer scientist Timnit Gebru from Google was due to the attention she drew to this precise issue in a co-authored paper on the large language models (LMs) that underpin popular online translation tools like Google Translate (Hao, 'Language AI'). The paper highlights how the scale of these models and the vast resources and technological infrastructure required for their development and use serve to heighten inequalities. The authors emphasised that investments in these technologies are primarily focused on already dominant languages like English, even while the environmental impacts of their use are likely to fall on marginalised speakers of other languages:

> Is it fair or just to ask, for example, that the residents of the Maldives … or the 800,000 people in Sudan affected by drastic floods pay the environmental price of training and deploying ever larger English LMs, when similar large-scale models aren't being produced for Dhivehi or Sudanese Arabic? (Bender et al 612–13)

The article estimates that over 90 per cent of the world's languages, which together are currently spoken by more than a billion people, have little or no language technology support. In computer science, these are often called 'low-resource languages', referring to the fact that the models on which machine translation and related technologies are built require vast quantities of existing 'training text', or corpora, which for many of the world's languages are less readily available. Language technologies consequently often serve to further amplify languages that are already dominant, with the organisation Whose Knowledge? estimating that only about 500 of the world's 7,000+ languages are represented online (Vrana et al 7). Relatedly, Gebru

and her co-authors highlight how the reliance on large-scale training data drawn primarily from existing texts on the Internet risks encoding and reinforcing already hegemonic viewpoints and associated forms of racial and gender bias.

The most obvious solution to addressing such inequalities and better representing a wider range of languages and viewpoints online would seem to be to replicate the same approaches developed for English and thereby expand language technology support to more speakers. This, however, carries the risk of forcing distinct linguistic systems to conform to the norms and rules of English in ways that can perpetuate anglophone biases and worldviews. Attempting to create large written corpora for predominantly oral languages, for instance, calls to mind the historical figure of the missionary linguist, whose descriptions of languages from the colonial peripheries served to legitimise regimes of colonial difference (Errington). Equally, while more effective content moderation on social media platforms may appear a worthy goal, the flipside of such technologies is their close connection to forms of surveillance and control. With linguistic inequalities being intimately linked to histories of colonial violence and erasure, marginalised language communities have legitimate reasons to mistrust and fear the loss of ownership of their cultural and linguistic knowledge and practices.

Beyond the predominantly US-based creators of such technologies, there are, however, important initiatives to develop more community-led, ethical, and responsible language and translation technologies. Masakhane, for example, is a grassroots organisation aimed at supporting the development of such technologies 'in African languages, for Africans, by Africans'. Foregrounding African-centricity, their work is aimed at challenging a 'technological space that does not understand our names, our cultures, our places, our history'. Yet, as Gebru's co-authored paper highlights, the vast investments of Silicon Valley in resource-intensive language technologies risk crowding out the potential for these more ethical community-led approaches to carve out their own space in the technological landscape.[2]

Supposing, however, that it were possible to overcome such challenges and to develop more sophisticated and equitable translation technologies for all languages, allowing anyone to access and view the Internet in the language of their choice. While this might certainly seem more appealing than the monolingual dominance of English over the early Internet, Helen Kelly-Holmes posits that this goal of 'mass linguistic customization' might in practice lead to the creation of linguistic 'filter bubbles' that, like other forms of filter bubble, reinforce intellectual isolation (34). Kelly-Holmes envisions a world in which we are able to travel the global, multilingual web cocooned in our own monolingual bubble, which, through the further development of forms of wearable technology such as earpieces and microphones, could also be expanded to our offscreen interactions.

Kelly-Holmes's warning of a future in which automated translation eradicates the need to engage with any language other than one's own echoes similar discussions that have taken place in relation to comparative and world literature. Gayatri Spivak has been particularly critical of the illusion of quickly understanding the literatures of the world via underexamined processes of translation (*An Aesthetic Education*). While a renowned translator herself, Spivak emphasises that she also continues to 'resist translation' ('Translation as Culture' 23) and, in what might seem to be a pre-emptive retort to Zuckerberg's claim, those who view it as 'a mere convenience because we cannot learn every language in the world' (18). Spivak's concerns that an overreliance on translation may in practice serve to reinforce globalising monocultures echo warnings of a future in which machine translation makes learning languages an increasingly niche or even obsolete activity. If we understand language learning as so much more than an instrumental skill, and rather as a conscious orientation to the words and modes of thought of others (Harvey), we begin to understand how much is at stake with the development of these technologies.

A return to Spivak illustrates the importance of incorporating the forms of critical theory that inform research on language, culture, and translation into discussions of technological change. While machine translation has long been an established area of research in translation studies, Minako O'Hagan notes that the 'cultural turn' in translation studies in the 1990s led to a greater separation between those interested in human and machine translation respectively. However, the pervasiveness of digital technology today calls on us to reconsider the logic of attempting to isolate the cultural from the technological. Even if we might be tempted to disengage from the

overwhelming pace of technological change, might doing so merely cede greater influence to tech bros like Zuckerberg in shaping the linguistic and cultural transformations that digital advances will inevitably bring?

While computer scientists and related experts may be best placed to understand the technological processes that underpin digital tools and platforms, their work must be informed by much wider critical analysis of the cultural, societal, and linguistic effects of these tools.[3] The discipline of cultural studies itself arose in response to a period of rapid social and cultural change in post-war Britain that was at least partly connected to technological change, as well as from an ethical imperative to intervene 'in a world in which it would make some difference' (Hall 286). As the ever more critical role of AI in translation ushers new and profound transformations, we'd do well to reapply ourselves to this urgent task.

NOTES

[1] Sharon O'Brien does, however, also highlight the need to give serious consideration to the ethical questions raised by the use of such technologies in disaster management, with the risk of causing harm by conveying inaccurate information evidently heightened in an emergency setting.

[2] There are, however, attempts by some large tech companies to adopt more community-led approaches, such as the work led by Kalika Bali at Microsoft India. Bali is committed to challenging universalist technical solutions by working closely with language communities to co-design linguistically and culturally sensitive technologies (cited in Spence 16).

[3] This is not to diminish the vital critical work of computer scientists like Timnit Gebru; on the contrary, the resistance that researchers like Gebru face in challenging hegemonic approaches in computer science illustrates the importance of dramatically expanding the body of critical research in this area from both within and beyond the tech world.

WORKS CITED

Bender, Emily M, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?' *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2021, pp. 610–23, doi: 10.1145/3442188.3445922.

Cronin, Michael. *Translation in the Digital Age*. Routledge, 2013.

Edmond, Jennifer. 'A Recipe for Intimacy?' *KPLEX*, 10 May 2017, kplex-project.com/2017/05/10/a-recipe-for-intimacy/. Accessed 15 Jan. 2022.

Errington, Joseph. 'Colonial Linguistics.' *Annual Review of Anthropology*, vol. 30, no. 1, Oct. 2001, pp. 19–39, doi: 10.1146/annurev.anthro.30.1.19.

Hall, Stuart. 'Cultural Studies and Its Theoretical Legacies.' *Stuart Hall: Critical Dialogues in Cultural Studies*, edited by David Morley and Kuan-Hsing Chen. Routledge, 1996, pp. 261–94.

Hao, Karen. 'The Facebook Whistleblower Says Its Algorithms Are Dangerous. Here's Why.' *MIT Technology Review*, 5 Oct. 2021, www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms/. Accessed 15 Jan. 2022.

_____ . 'The Race to Understand the Exhilarating, Dangerous World of Language AI.' *MIT Technology Review*, 20 May 2021, www.technologyreview.com/2021/05/20/1025135/ai-large-language-models-bigscience-project/. Accessed 15 Jan. 2022.

Harvey, Lou. '"I Am Italian in the World": A Mobile Student's Story of Language Learning and *Ideological Becoming*.' *Language and Intercultural Communication*, vol. 16, no. 3, July 2016, pp. 368–83, doi: 10.1080/14708477.2016.1168049.

Kelly-Holmes, Helen. 'Multilingualism and Technology: A Review of Developments in Digital Communication from Monolingualism to Idiolingualism.' *Annual Review of Applied Linguistics*, vol. 39, Mar. 2019, pp. 24–39, doi: 10.1017/S0267190519000102.

Marwick, Alice E, and danah boyd. 'I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience.' *New Media & Society*, vol. 13, no. 1, Feb. 2011, pp. 114–33, doi: 10.1177/1461444810365313.

Masakhane. www.masakhane.io/. Accessed 19 Dec. 2021.

O'Brien, Sharon. 'Translation Technology and Disaster Management.' *The Routledge Handbook of Translation and Technology*, edited by Minako O'Hagan. Routledge, 2019, pp. 304–18.

O'Hagan, Minako. 'Introduction: Translation and Technology: Disruptive Entanglement of Human and Machine.' *The Routledge Handbook of Translation and Technology*, edited by Minako O'Hagan. Routledge, 2019, pp. 1–18.

Spence, Paul. *Disrupting Digital Monolingualism: A Report on Multilingualism in Digital Theory and Practice*. Version 1.0, Disrupting Digital Monolingualism (DDM), Language Acts & Worldmaking Project, 9 Dec. 2021, doi: 10.5281/zenodo.5743283.

Spivak, Gayatri Chakravorty. *An Aesthetic Education in the Era of Globalization*. Harvard UP, 2013.

_____ . 'Translation as Culture.' *Parallax*, vol. 6, no. 1, Jan. 2000, pp. 13–24, doi: 10.1080/135346400249252.

Translators without Borders. 'TWB Partners with Tech Leaders to Develop COVID-19 Language Technology for 37 Languages,' 6 July 2020, translatorswithoutborders.org/TICO-19-announcement. Accessed 15 Jan. 2022.

Vrana, Adele, Anasuya Sengupta, Claudia Pozo, and Siko Bouterse. *Decolonizing the Internet's Languages — Summary Report*. Whose Knowledge?, 2020, whoseknowledge.org/wp-content/uploads/2020/02/DTIL-Report.pdf. Accessed 15 Jan. 2022.