Ethical Considerations for Low-resourced Machine Translation

Levon Haroutunian
Department of Linguistics
University of Washington
`levonh@uw.edu`

Abstract

This paper considers some ethical implications of machine translation for low-resourced languages. I use Armenian as a case study and investigate specific needs for and concerns arising from the creation and deployment of improved machine translation between English and Armenian. To do this, I conduct stakeholder interviews and construct Value Scenarios (Nathan et al., 2007) from the themes that emerge. These scenarios illustrate some of the potential harms that low-resourced language communities may face due to the deployment of improved machine translation systems. Based on these scenarios, I recommend 1) collaborating with stakeholders in order to create more useful and reliable machine translation tools, and 2) determining which other forms of language technology should be developed alongside efforts to improve machine translation in order to mitigate harms rendered to vulnerable language communities. Both of these goals require treating low-resourced machine translation as a language-specific, rather than language-agnostic, task.

1 Introduction

The challenge of building machine translation systems for low-resourced languages is often seen merely a problem of data scarcity, but such a framing obscures the systemic differences between low-resourced languages and high-resourced languages, as well as between their corresponding speaker communities. Before building machine translation systems for low-resourced languages, it is important to consider the real impact that these systems can have on their intended users. This paper investigates the ethical implications for improvements to machine translation for low-resourced languages using a Value Scenarios (Nathan et al., 2007) framework to identify potential harms and construct recommendations to mitigate them.

The remainder of this paper uses the Armenian language as a case study and considers the specific circumstances of speakers in the Armenian Diaspora. I conducted interviews with four stakeholders (Armenian speakers who use machine translation systems) and analyzed their responses to identify common desires and concerns for machine translation. I then used these analyses to construct Value Scenarios, which illustrate potential unintended consequences of improving the quality of machine translation between English and Armenian to the level of current machine translation between English and other high-resourced languages.

In examining machine translation for Armenian, I hope to provide examples of the kinds of harms that may be caused to speakers of low-resourced languages in the development of machine translation tools. This paper is not meant as an exhaustive exploration of all possible harms; instead, it provides a starting point for considering how the specific circumstances of a language community can inform the creation of ethically-produced machine translation tools.

My findings show that machine translation for low-resourced languages should not be undertaken as an aggregated language-agnostic task, but should instead be approached in a language-specific way that contextualizes speakers' needs and other facets of existing language technology. Such an approach would allow NLP researchers to move away from a data-first paradigm that privileges high-resourced languages and varieties towards one that can take into account the particular circumstances of vulnerable groups in order to ensure that we build machine translation tools that are genuinely useful and reliable.

The rest of this paper is organized as follows. Section 2 provides background information on machine translation for low-resourced languages and describes Value Scenarios. Following that, Section 3 gives an overview of considerations for Armenian. Section 4 describes the methodology I followed to conduct stakeholder interviews, the results of

which are analyzed in Section 5. Based on interviewees' responses, two Value Scenarios are presented and analyzed in Section 6, and recommendations based on these scenarios are given in Section 7. Section 8 describes ethical considerations for this project, and Section 9 provides a conclusion.

2 Background

2.1 (Neural) Machine Translation for Low-Resourced Languages

In the past few years, neural machine translation (NMT) (Bahdanau et al., 2016) has become the dominant model for machine translation applications. NMT systems generally show a wide gap in translation quality for low-resourced languages and high-resourced languages (Caswell and Liang, 2020). This difference in performance has significant consequences for speakers of low-resourced languages. As Joshi et al. (2020) argue, the divergence in the quality of NLP applications for high-resourced languages and low-resourced languages can exacerbate conditions that lead to language decline.

There have been many efforts to improve machine translation for low-resourced languages (e.g. Gu et al., 2018; Lample et al., 2018; Ko et al., 2021; Zoph et al., 2016; Fadaee et al., 2017), many of which take a transfer learning approach that applies models trained on multilingual data to languages for which there is less data available. Many current NMT models are trained on large, uncurated datasets, such as Wikipedia dumps (Gu et al., 2018). As a result, NMT systems generally work better for standard language varieties than non-standard ones (Kumar et al., 2021).

When approaching the task of machine translation for low-resourced languages, it is important to consider the relationship between the NLP community and communities that speak low-resourced languages. As Nekoto et al. (2020) detail, understanding low-resourcedness as merely a lack of data is reductive, as this framing fails to capture the corresponding lack of linguistic and geographic diversity of NLP scholars. Nekoto et al. (2020) also describe how much of the work in machine translation is Anglo-centric, as it prioritizes the quality of translations to and from English.

Both of these observations point to a disconnect between the NLP community and speakers of low-resourced languages. As a result, language technologists building applications for low-resourced languages may not have a clear picture of the wants and needs of these languages' speaker communities, and may not understand the benefits and harms rendered to these communities by the language technology they build. The number of languages considered to be low-resourced is vast, and communities that speak low-resourced languages have diverse needs; often, these needs are very different from the needs of speakers of high-resourced languages (e.g. Joshi et al., 2019).

2.2 Value Scenarios

Incorporating stakeholders' perspectives is crucial for creating useful improvements in language technology for low-resourced languages (e.g. Nekoto et al., 2020; Joshi et al., 2019).

To do that, this work will draw on the principles of Value-Sensitive Design (Friedman, 1996), particularly Value Scenarios (Nathan et al., 2007), to identify ethical challenges in the improvements of machine translation for low-resourced languages. A Value Scenario imagines the systemic impacts of a proposed technology in order to anticipate and mitigate negative consequences before that technology is deployed. Considering how many efforts are underway to improve the quality of machine translation for low-resourced languages, it is pertinent to examine a range of impacts that these improvements may have for speakers of low-resourced languages. The utility of Value Scenarios in this use case is to illustrate the needs and concerns of a particular group of stakeholders in a way that generates specific avenues for harm mitigation.

The Value Scenarios framework (Nathan et al., 2007) identifies five key elements to explore: stakeholders (people who are impacted by the technology either directly or indirectly), pervasiveness (the effects of the technology when it has widespread use), time (the effect of the technology in short- and long-term scales), systemic effects (how the technology interacts with various areas of life), and value implications (potential positive and negative influences that impact use of the technology). My analysis in this paper will draw on each of these five components.

It should be noted that the purpose of Value Scenarios is not to generate predictions for the future, nor is it possible to use Value Scenarios to imagine every possible consequence of a proposed technology (Nathan et al., 2007). The task of preventing harms to marginalized groups is complex and

ever-changing, and requires the integration of many types of expertise and a variety of tools. This paper considers the use of Value Scenarios as one such tool.

3 Considerations for Armenian

In this section, I present two important challenges specific to the improvement of machine translation for Armenian: language variation and orthography. While the details of these challenges pertain to the particular situation of the Armenian language and of Armenian speakers, it is likely that improvements to other low-resourced languages would require similar considerations.

3.1 Variation

Modern Armenian has two main varieties: Modern Eastern Armenian and Modern Western Armenian. While these two varieties are mostly mutually intelligible, there are large distinctions between them in phonology, morphology, syntax, and lexical items.

The following characterization of Armenian-speaking populations is consistent with Eberhard et al. (2021). Eastern Armenian speakers are predominantly those in Armenia, Artsakh, Russia, and Iran. Western Armenian speakers are predominantly those in the United States, Lebanon, Georgia, and Argentina. There are about 3.8 million Eastern Armenian speakers and about 1.4 million Western Armenian speakers worldwide.

There is a large imbalance in the amount of data available in each of the two main varieties; for example, the Eastern Armenian Wikipedia (called simply Armenian Wikipedia) has about 291 thousand articles in early 2022, while the Western Armenian Wikipedia has only about 10 thousand.[1]

3.2 Orthography

The following description is based on Hagopian (2007). While all varieties of Armenian are written with the same alphabet, there are two sets of spelling conventions: Classical Orthography and Reformed Orthography. There are substantial differences between the two systems, though it is generally possible for someone who typically uses one orthography to read text written in the other. Speakers of all varieties of Western Armenian and of the *Barskahye* variety of Eastern Armenian use Classical Orthography, while all other speakers of Eastern Armenian use Reformed Orthography. The vast

Eastern, Reformed	իմ քույրը խոսում է
Eastern, Classical	իմ քոյրը խոսում է
Eastern, IPA	im ˈkʰui.rə ˈχo.sum e
Western, Classical	իմ քոյրս կը խոսի
Western, IPA	im ˈkʰui.rəs gə χo.ˈsi

Table 1: The phrase "my sister speaks" in both Eastern and Western Armenian. The transcriptions in IPA show roughly standard pronunciations for both main varieties.

majority of text in Armenian on the Internet is in Reformed Orthography. See Table 1 for an example of differences between varieties and orthographies.

Additionally, many Armenian speakers often write using ad-hoc Romanization rather than Armenian script, which is an additional challenge for machine translation.

4 Methodology

This section describes the process by which I conducted stakeholder interviews. This study was approved by the Institutional Review Board (IRB) at the University of Washington.

4.1 Participants

Four volunteer participants were recruited on the basis of their status as Armenian and English speakers who have previously used machine translation technology. To avoid the unwanted identification of these participants, the following description and the discussion in Section 5 include only details that are necessary for contextualizing the perspectives described in this paper.

Each participant speaks a different variety of Armenian: Standard Eastern Armenian, the *Karabaghtsi* (Artsakhi) variety of Eastern Armenian, the *Barskahye* (Iranian) variety of Eastern Armenian, and Standard Western Armenian. These varieties cover a large swath of Armenian speakers, although of course they do not constitute the totality of variation in Armenian.

While the stakeholders I interviewed are diverse in the varieties they speak, they are otherwise a somewhat homogeneous group. All of the interviewees are below the age of 35, and all have resided in the United States for the majority of their lives. Additionally, while all four interviewees speak Armenian at home and in some social settings, they also all speak English as their primary language in other settings. Therefore, the perspectives described in this paper reflect a partic-

---

[1]meta.wikimedia.org/wiki/List_of_Wikipedias

ularly diasporic and English-dominant experience of Armenian identity, which has a clear influence on the desires and concerns described in Section 5. It should also be noted that all of the interviewees were people in my own network, and do not form a representative sample of Armenian speakers as a whole or even of Armenian speakers in the United States.

4.2 Interviews

Before the interviews were conducted, participants were provided with a description of this study's purpose and the topics that would be discussed during the interview, along with a description of how their data would be stored and used. Each participant was interviewed separately over a video call that lasted between one and two hours in length. With the participants' consent, the interviews were recorded to aid later analysis. To protect the interviewees' privacy, all recording files are secured in accordance with the guidelines established by the Human Subjects Division at the University of Washington.

These interviews were conducted in a semi-structured format: some questions were pre-determined, and others were based on participants' responses in the moment. I constructed a set of basic questions for each topic I planned to discuss with interviewees, and these questions served as starting points to informal conversations. This format was chosen in order to illuminate comparisons between different interviewees' experiences with machine translation while allowing the course of each interview to be shaped by particulars of the interviewee's perspective. The length of the interviews was determined by the length of interviewees' responses.

Each interview covered the same set of topics, including the participant's 1) use of Armenian, 2) experience of being an Armenian person in online spaces, 3) use of machine translation technology, 4) desired improvements for Armenian-English machine translation, and 5) expected benefits and harms for Armenian-English machine translation. Below is a sample of the questions that I determined prior to the interviews; a complete list can be found in Appendix A.

- What is your experience using machine translation tools? How usable are they for you?

- What is your experience as an Armenian speaker online?

- When you translate from English to Armenian, do machine translation tools give you something that sounds like the way you would speak?

- When you translate from Armenian to English, do you run into any problems that relate to the way you speak Armenian?

- If machine translation for Armenian (to and from English) improved, how do you think it would affect you? How do you imagine other people (both Armenians and non-Armenians) would use it?

4.3 Limitations

To contextualize the results in Section 5 and the Value Scenarios in Section 6, it is important to acknowledge the limitations of this project.

First, as stated previously, the participant group forms a non-representative sample of Armenian speakers. There are only four interviewees, and they have similar backgrounds: they all live in the United States, they all speak English as a primary language, and they are all relatively young. Likewise, all of the speakers I interviewed indicated the same types of uses for machine translation and largely similar concerns. It is very likely that different results would have emerged if I had been able to interview a more diverse group of Armenian speakers, particularly if I was able to incorporate the perspectives of older speakers and those who speak a language other than English as a primary language. This is not to say that the needs and concerns identified below are any less relevant – merely that there are likely many other needs and concerns that I was not able to identify. The perspectives in this paper should not be taken as representative of all Armenian speakers.

Second, as stated in Section 2, the Value Scenarios approach cannot uncover every possible consequence of a proposed technology, since many harms are emergent. The harms described in this paper do not constitute the totality of potential negative impacts for low-resourced machine translation.

5 Results

Below is an overview of significant themes that emerged from my stakeholder interviews.

5.1 State of Current Machine Translation Tools for Armenian

In general, respondents said that they mostly used translation tools to look up words or short phrases. The most common usages respondents reported was to help them remember words that they already knew or to find words specific to Standard Eastern Armenian. One respondent said that she occasionally used machine translation to look up specific terms she knew in English but not Armenian in order to facilitate communication with family members who do not speak English well.

All interviewees were familiar with Google Translate[2], which has several features that they found useful. One such feature is transliterated output, which makes interpretation easier for interviewees who are unable to read Armenian or less practiced. Audio output was similarly useful. Respondents reported that they were usually able to find the English translation of an Armenian word they were not able to spell correctly, which was helpful.

On the other hand, every respondent reported a lack of trust in Google Translate's accuracy, with multiple respondents reporting that they usually verified the output with another Armenian speaker before incorporating it into their own speech. Additionally, all respondents noted that the output from Google Translate had a markedly Standard Eastern Armenian style, including the exclusive use of Reformed Orthography. As a result, only the respondent who speaks Standard Eastern Armenian reported that she was able to consistently get output from Google Translate that matches the way she speaks.

Most of the participants were also familiar with Nayiri Armenian Dictionary[3], which is an online resource that supports Western Armenian (in Classical Orthography) and Eastern Armenian (in Reformed Orthography). Nayiri, which is maintained by a small team of Armenian software engineers and linguists, incorporates a database of digitized Armenian dictionaries into its search. Respondents who used Nayiri reported that they trusted its output far more than they trusted that of Google Translate, but that Nayiri was more challenging to use: the site is less user-friendly, there is less forgiveness for misspellings, and Nayiri only supports single-word look-ups rather than phrase or sentence trans-

lations.

The respondent who reported the least amount of resources for her variety was the *Barskahye* speaker, who reported that she was unable to find any translation tool that outputs results in Eastern Armenian using Classical Orthography.

Respondents reported that neither tool was useful for translating full sentences or paragraphs in either direction. When respondents have tried to translate longer utterances on Google Translate, output was generally jumbled or nonsensical.

5.2 Desires for and Anticipated Benefits of Improved Machine Translation

The speakers I interviewed were enthusiastic about the prospect of improved machine translation tools, and each of them was able to identify both personal and communal benefits. Beyond a general improvement in translation quality, interviewees most strongly desired 1) expanded support for varieties other than Standard Eastern Armenian, and 2) output in Reformed Orthography, Classical Orthography, and in Roman characters.

There was a wide variety of potential uses that the interviewees identified for improved machine translation tools:

- *Language learning.* All of the stakeholders I interviewed said that they would hope to utilize improved machine translation tools to expand their own knowledge of the Armenian language, specifically to improve their vocabulary (in their own and other varieties) and to strengthen their literacy.

- *Transmitting urgent information.* Two stakeholders identified machine translation as a tool to help Armenians in the diaspora more rapidly understand urgent news coming out of Armenia and Artsakh. This is a particularly pressing need in the wake of the 2020 war between Artsakh and Azerbaijan.

- *Connecting to other Armenians.* Related to the above points, interviewees stated that they would use improved machine translation tools to better communicate with other Armenians, both those that speak their variety and those that speak other varieties. In particular, one interviewee spoke of the potential to use such tools to build bridges between diaspora communities and Armenia and Artsakh.

---

[2]translate.google.com
[3]nayiri.com

- *Connecting outsiders to Armenia.* One interviewee suggested that improved machine translation would bolster tourism prospects for Armenia, while another suggested that it would allow outsiders easier access to information and history that has thus far only been available in Armenian.

5.3 Concerns about harassment and disinformation

All respondents described seeing frequent online harassment against Armenians, generally from Turkish and Azerbaijani ultra-nationalists. According to respondents, there has been a substantial increase in harassment since the beginning of the 2020 war between Artsakh and Azerbaijan.

Two respondents reported receiving harassment on social media themselves, and all respondents reported seeing other Armenians be harassed. This harassment generally comes in the form of spam, specifically the use of particular emojis (e.g. Azerbaijani and Turkish flags, skulls, coffins, pigs, wolves, and knives) and inflammatory or disturbing hashtags. Other forms of harassment include 1) comments advocating violence against Armenians, denying the Armenian Genocide, celebrating the Armenian Genocide, and claiming Azerbaijani ownership of Armenian cultural monuments; 2) hateful or disturbing memes; and 3) videos of Azerbaijani soldiers desecrating Armenian churches and cemeteries, flying Azerbaijani flags over Armenian buildings and monuments, destroying Armenian homes and property, and in the worst cases, torturing and murdering Armenian soldiers and civilians.

All respondents stated that their relationship with social media changed in the wake of the war, with anti-Armenian harassment being one factor that influenced this change. When I asked respondents what negative impacts they could imagine from the deployment of an effective machine translation tool for Armenian, three of the four respondents independently brought up the potential for production of hateful content. These respondents expressed concerns that malicious actors could use improved machine translation to further their harassment of Armenians, either by using it to better understand posts written in Armenian and attacking creators of those posts, or to translate hateful messages into Armenian (which would potentially be more disturbing than hateful messages written in English).

Additionally, interviewees were concerned about the possibility of machine translation tools being used for disinformation campaigns and propaganda from Azerbaijani military forces.

5.4 Concerns about standardization

When presented with the possibility of machine translation tools being improved only for Standard Eastern Armenian and not for other varieties, three of the four interviewees expressed concern that this move would negatively impact speakers of Western Armenian and non-standard varieties of Eastern Armenian. Specifically, interviewees were concerned that the hegemony of Standard Eastern Armenian online, amplified by machine translation tools that exclusively produce output in Standard Eastern Armenian, would contribute to the common belief that Eastern Armenian written in Reformed Orthography is the most "correct" or "pure" form of Armenian.

6 Value Scenarios

In this section, I present two value scenarios that I have constructed based on the above findings. These value scenarios are intended to illustrate *potential* unwanted consequences of improving machine translation for low-resourced languages. They are not meant to be predictions of real events; rather, they are deliberately dark imaginings of the impacts that new technology could have (Nathan et al., 2007). The purpose of creating these value scenarios is to uncover considerations that may need to be made before developing improvements to machine translation for low-resourced languages.

While the two scenarios below are presented as separate outcomes, it should be noted they could occur simultaneously. The distinction between them is merely for the purpose of more easily illustrating different possible consequences.

6.1 Value Scenario 1

Thanks to advances in unsupervised neural machine translation, there have been large improvements in translation between English and languages with relatively large monolingual corpora; Standard Eastern Armenian is one such language. Due to these developments, machine translation in Standard Eastern Armenian on platforms like Google Translate is much more reliable than it used to be.

For people looking to learn Standard Eastern Armenian either to connect with their family or to visit Armenia on vacation, these applications are

very useful. However, speakers of minoritized varieties of Armenian receive none of these benefits. On top of that, machine translation for Armenian is now regarded as a solved task, so there is little motivation for expanding machine translation capabilities for other varieties.

Many more websites and platforms are able to support Armenian text and Armenian users, but it assumed that all of these users are willing and able to communicate in Standard Eastern Armenian written in Reformed Orthography. This contributes to the perception that Standard Eastern Armenian is the only legitimate form of Armenian, leading other speakers to feel alienated from their communities. Speakers of Western Armenian and non-standard varieties of Eastern Armenian alter their speech to fit in, or they avoid speaking Armenian at all when other languages are available. Artsakhi refugees of the 2020 war are ridiculed for their speech in their new homes in Armenia; many of them face additional burdens at school or work because their speech is seen as unintelligent.

Over time, other varieties' speaker populations decline, and the linguistic diversity of Armenian speakers around the world is replaced with homogeneity. Along with these varieties, numerous artifacts of minoritized Armenian cultures become less accessible and, in some cases, are lost. This is particularly painful for Western Armenian communities, for whom language was one of the most significant cultural resources that persisted in the wake of the Armenian Genocide.

Analysis In this scenario, improvements to machine translation only for the most high-resourced variety of Armenian exacerbate existing biases against speakers of lower-resourced varieties. The implicit standardization of one variety leads to further marginalization of the others, which has social and cultural consequences, including the erasure of distinct minoritized cultures.

6.2 Value Scenario 2

After substantial time and effort, improvements to machine translation tools are rolled out for a number of low-resourced languages, including Armenian. These improvements increase the accuracy of translation between English and Armenian to a level that is currently only seen among the most high-resourced language pairs. These improvements give Armenians in the diaspora better tools for developing their language skills, which allows some users to communicate more freely with their families and friends and connect with communities in Armenia and Artsakh.

On the other hand, Armenians are facing an extreme increase in online harassment. Turkish and Azerbaijani ultra-nationalists, seizing upon capabilities of newly released machine translation systems, gleefully descend into Armenians' DMs, retweets, and comments with translated messages expressing their hatred of Armenian people. Unlike the harassment that Armenians had been receiving previously, this time the comments are lengthier, more descriptive, and more disturbing – and they're in Armenian. While these comments are not translated perfectly, their meaning and intent is clear enough; the fact that they appear in the users' own language only adds more pain to the experience.

Because major social media platforms have yet to implement content moderation policies for content written in Armenian, the platforms are unable or unwilling to address this influx of harassment. Armenian users are able to delete messages containing harassment and block the senders' accounts, but this does not prevent trolls from making new accounts and sending more messages. For many Armenians on social media, this becomes an exhausting part of their daily routine. With all this effort expended, they still have to see the disturbing messages.

To escape harassment, many Armenian users, particularly those with large followings, leave social media for good. They are unable to use platforms like Twitter, Instagram, or Facebook to connect with friends and family or to engage with their communities. It becomes more challenging for Armenians to find job opportunities that are advertised on social media or to establish professional online profiles. Armenian artists and small business owners have to weigh the prospect of harassment if they maintain public profiles against losses in income if they don't.

The number of Armenian voices online gradually diminishes; in their absence, disinformation, anti-Armenian propaganda, and genocide denial flourishes.

Analysis It is crucial to account for the ways that machine translation interacts with existing technology, particularly on social media. Many Armenians already have to contend with harassment on social media, which affects their ability to engage with these platforms (as detailed in Section 5.3). If a new

machine translation tool is deployed without taking these circumstances into account, there could be dire consequences.

Improved machine translation can allow for a sudden proliferation of text in a low-resourced language like Armenian online, potentially from bad actors. To prevent unwanted harms, it is necessary for social media platforms to take proactive steps to support these language communities. In the above scenario, that means creating more robust content moderation policies and the infrastructure needed to enforce these policies. Depending on community-specific vulnerabilities, there are likely other possible harms that would need to be mitigated using other strategies.

7 Recommendations

The potential benefits of improved machine translation for low-resourced languages are enormous. The stakeholders I interviewed all named specific uses they would have for better translation tools, ranging from improving their literacy skills to strengthening their connections to their families and communities. The potential harms are enormous as well, as the above scenarios illustrate. Different speaker communities will have other uses for and concerns about machine translation (Paullada, 2020). Ensuring that improved machine translation tools maximize the benefits and mitigate the harms requires the NLP community to take explicit steps to collaborate with and support low-resourced language communities.

First, it is necessary to examine the particular wants and needs that language communities have during the planning stage of a project. This paper demonstrates the efficacy of a Value-Sensitive Design approach in surfacing a particular community's needs and anticipating potential harms before technology is built. The interviews described in Section 5 and the resulting Value Scenarios illuminate concerns that otherwise might only be apparent to Armenian speakers. Similar efforts can be undertaken with speakers of low-resourced languages to uncover other community-specific considerations. Value-Sensitive Design provides a number of other practical techniques for collaborating with stakeholders (Friedman et al., 2017), which may be useful in future efforts.

Second, we must consider what other facets of language technology should be developed alongside improvements to machine translation. The deployment of robust machine translation allows for the generation of large volumes of text in a low-resourced language, which can have negative impacts for language communities. These impacts are likely impossible to prevent without actions taken by entities outside of NLP; for instance, preventing the outcome described in Value Scenario 2 requires social media platforms to implement stronger content moderation policies in low-resourced languages. NLP researchers can, however, work to expand the capabilities of other facets of language technology (in this case, hate speech detection for Armenian) that can mitigate potential harms caused by improved machine translation.

Fulfilling these goals requires disaggregating the task of machine translation; rather than creating translation tools for many languages at once, each language should be considered separately. Doing this would undoubtedly be more resource-intensive than a language-agnostic approach, but it is a necessary step towards prioritizing the needs of low-resourced language speakers. The scenarios in Section 6 illustrate just a couple of the ways that speakers of low-resourced languages may have very different circumstances than speakers of high-resourced languages, both linguistically and geopolitically, that need to be taken into account when machine translation applications are deployed. In both scenarios, harms fall unduly on groups that are already marginalized: in the first scenario, minoritized Armenian speakers bear the brunt of these harms, while in the second, Armenians in general are impacted negatively. Treating machine translation as an abstract language-agnostic task, divorced from the specific needs of distinct groups of users, obscures harms like these. Worse, it risks exacerbating inequitable conditions.

Taking a language-specific, stakeholder-focused approach does more than prevent potential harms; it also builds better, more reliable technology. When researchers assemble datasets for languages they are not familiar with, they are often unable to verify the validity of a data source and may not be able to find an existing high-quality data source (Nekoto et al., 2020). This is illustrated by the difference in reliability that interviewees reported for Google Translate and Nayiri: while Google Translate has more useful features, Nayiri is more trustworthy because it is built by a team with deep knowledge of the language and communities it serves, using carefully curated resources that may be inaccessible to

outsiders.

The current paradigm of building NMT systems that rely on vast quantities of unlabelled data, whose size prevents careful curation (Bender et al., 2021; Paullada et al., 2021), makes it difficult to build systems that can account for language variation and serve users that speak minoritized varieties. As a result, machine translation systems cannot produce reliable and useful output for speakers whose varieties do not have substantial bodies of data. Building better and more equitable systems requires moving away from data-first approaches and investing in holistic methods that take into consideration the state of existing language technology and external circumstances of the communities in question, as well as developing higher-quality data sources (Hanna and Park, 2020; Paullada et al., 2021). This process does not need to begin from scratch; as with the example of Nayiri, low-resourced language communities may already have ongoing intra-community projects that would be fruitful sites for investment from and collaboration with NLP practitioners.

8 Ethical Considerations

As described in Section 4.3, the methodology used in this paper has a number of limitations that affect how these results may be generalized. Most prominently, the stakeholder group that I interviewed was small and represented only a small subsection of the perspectives of Armenian speakers.

Additionally, the group of participants described in this paper comprises speakers of only one low-resourced language; speakers of other low-resourced languages would likely have very different needs and concerns. This case study is meant only to provide examples of the concerns of speakers of a particular low-resourced language. It is important to avoid generalizing low-resourced languages and their speakers.

This paper does not cover all of the potential harms of machine translation; further efforts are needed to uncover other concerns for individual language communities. If only the harms I described in this paper were taken into consideration in the development of a machine translation system, it is certain that other important concerns would be missed, which could cause substantial harms to speaker populations.

9 Conclusion

Using Value Scenarios, this paper illustrates some potential harms that a general-purpose machine translation system could have for speakers of a low-resourced language. Avoiding these harms requires direct collaboration with stakeholders before the creation of a machine translation system intended for low-resourced languages. To do so, machine translation for low-resourced languages should be undertaken as a language-specific task.

Acknowledgments

References

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2016. Neural Machine Translation by Jointly Learning to Align and Translate. In *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, California. ArXiv: 1409.0473.

Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623, Virtual Event Canada. ACM.

Isaac Caswell and Bowen Liang. 2020. Recent Advances in Google Translate.

David M. Eberhard, Gary F. Simons, and Charles D. Fennig, editors. 2021. *Ethnologue: Languages of the World*, 24th edition. SIL International, Dallas, Texas.

Marzieh Fadaee, Arianna Bisazza, and Christof Monz. 2017. Data augmentation for low-resource neural machine translation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 567–573, Vancouver, Canada. Association for Computational Linguistics.

Batya Friedman. 1996. Value-sensitive design. *Interactions*, 3(6):16–23.

Batya Friedman, David G. Hendry, and Alan Borning. 2017. A Survey of Value Sensitive Design Methods. *Foundations and Trends® in Human–Computer Interaction*, 11(2):63–125.

Jiatao Gu, Hany Hassan, Jacob Devlin, and Victor O.K. Li. 2018. Universal neural machine translation for extremely low resource languages. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 344–354, New Orleans, Louisiana. Association for Computational Linguistics.

Gayané Hagopian. 2007. *Armenian for Everyone: Western And Eastern Armenian in Parallel Lessons.*, 2nd edition. Yerevan Printing, Glendale, California, USA. OCLC: 150335569.

Alex Hanna and Tina M. Park. 2020. Against Scale: Provocations and Resistances to Scale Thinking. *Proceedings of the CSCW 2020 Workshop: Reconsidering Scale and Scaling in CSCW Research*. ArXiv: 2010.08850.

Pratik Joshi, Christain Barnes, Sebastin Santy, Simran Khanuja, Sanket Shah, Anirudh Srinivasan, Satwik Bhattamishra, Sunayana Sitaram, Monojit Choudhury, and Kalika Bali. 2019. Unsung challenges of building and deploying language technologies for low resource language communities. In *Proceedings of the 16th International Conference on Natural Language Processing*, pages 211–219, International Institute of Information Technology, Hyderabad, India. NLP Association of India.

Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online. Association for Computational Linguistics.

Wei-Jen Ko, Ahmed El-Kishky, Adithya Renduchintala, Vishrav Chaudhary, Naman Goyal, Francisco Guzmán, Pascale Fung, Philipp Koehn, and Mona Diab. 2021. Adapting high-resource NMT models to translate low-resource related languages without parallel data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 802–812, Online. Association for Computational Linguistics.

Sachin Kumar, Antonios Anastasopoulos, Shuly Wintner, and Yulia Tsvetkov. 2021. Machine translation into low-resource language varieties. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 110–121, Online. Association for Computational Linguistics.

Guillaume Lample, Alexis Conneau, Ludovic Denoyer, and Marc'Aurelio Ranzato. 2018. Unsupervised Machine Translation Using Monolingual Corpora Only. In *Proceedings of the 6th International Conference on Learning Representations*, page 14.

Lisa P. Nathan, Predrag V. Klasnja, and Batya Friedman. 2007. Value scenarios: a technique for envisioning systemic effects of new technologies. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, pages 2585–2590, San Jose CA USA. ACM.

Wilhelmina Nekoto, Vukosi Marivate, Tshinondiwa Matsila, Timi Fasubaa, Taiwo Fagbohungbe, Solomon Oluwole Akinola, Shamsuddeen Muhammad, Salomon Kabongo Kabenamualu, Salomey Osei, Freshia Sackey, Rubungo Andre Niyongabo, Ricky Macharm, Perez Ogayo, Orevaoghene Ahia, Musie Meressa Berhe, Mofetoluwa Adeyemi, Masabata Mokgesi-Selinga, Lawrence Okegbemi, Laura Martinus, Kolawole Tajudeen, Kevin Degila, Kelechi Ogueji, Kathleen Siminyu, Julia Kreutzer, Jason Webster, Jamiil Toure Ali, Jade Abbott, Iroro Orife, Ignatius Ezeani, Idris Abdulkadir Dangana, Herman Kamper, Hady Elsahar, Goodness Duru, Ghollah Kioko, Murhabazi Espoir, Elan van Biljon, Daniel Whitenack, Christopher Onyefuluchi, Chris Chinenye Emezue, Bonaventure F. P. Dossou, Blessing Sibanda, Blessing Bassey, Ayodele Olabiyi, Arshath Ramkilowan, Alp Öktem, Adewale Akinfaderin, and Abdallah Bashir. 2020. Participatory research for low-resourced machine translation: A case study in African languages. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2144–2160, Online. Association for Computational Linguistics.

Amandalynne Paullada. 2020. How does Machine Translation Shift Power? *Resistance AI Workshop at 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*.

Amandalynne Paullada, Inioluwa Deborah Raji, Emily M. Bender, Emily Denton, and Alex Hanna. 2021. Data and its (dis)contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11):100336.

Barret Zoph, Deniz Yuret, Jonathan May, and Kevin Knight. 2016. Transfer learning for low-resource neural machine translation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1568–1575, Austin, Texas. Association for Computational Linguistics.

A   Appendix: Interview Topics and questions

Below are the questions that I used to guide each interview, separated by topics. Because the interviews were conducted in a semi-structured format, this list does not include every question I asked participants.

Topic 1: Use of Armenian Language

- Can you describe how you speak Armenian?

- What variety do you speak?

- How often do you speak it?

- With whom do you speak Armenian?

- How do you use Armenian online?

Topic 2: Experience of Being Armenian Online

- What is your experience as an Armenian speaker online?

- How do you engage with other Armenians or Armenian communities online?

- How do you engage with non-Armenians online?

- How difficult is it for you to communicate in Armenian online?

- Have you ever been the subject of harassment? If so, can you tell me more about that?

Topic 3: Use of Machine Translation Tools

- What is your experience using machine translation tools?

- What tools do you use?

- How well does it work for your variety?

- How usable is it for you?

- When you translate from English to Armenian, does it give you something that sounds like the way you would speak?

- When you translate from Armenian to English, do you run into any problems that relate to the way you speak Armenian?

- What is your understanding of how it works?

Topic 4: Desired Improvements and Potential Uses

- What concerns do you have about how machine translation currently works for Armenian?

- What would have to change about machine translation for Armenian to make it more useful for you?

- If machine translation for Armenian (to and from English) improved, how do you think it would affect you?

- How would it affect people you know?

Topic 5: Anticipating Improvements to Machine Translation

- How do you imagine other people (both Armenians and non-Armenians) would use an improved machine translation system?

- What benefits do you anticipate?

- What harms do you anticipate?

- How would it affect you if your data (speech or text) was used to improve it?

- What if machine translation was substantially improved for Standard Eastern Armenian, but not for other varieties? What impact would this have on you? What are the potential benefits you would expect in this scenario? What are the potential harms?

- How would it affect you if non-Armenians were able to understand you when you speak Armenian? Specifically, how would it affect you if you were understood by a) your friends, b) strangers on the internet, or c) trolls?

- Let's imagine a best-case scenario for improved machine translation. What would that look like? How do you think people would use it? How would you use it?

- Let's imagine a worst-case scenario. What would that look like? How would that affect you and people you know?

Topic 6: Miscellaneous

- What other concerns do you have about improvements to machine translation for Armenian?

- Is there anything else you'd like to add?