Is "good enough" good enough? Ethical and responsible development of sign language technologies[1]

Maartje De Meulder

Abstract

This paper identifies some common and specific pitfalls in the development of sign language technologies targeted at deaf communities, with a specific focus on signing avatars. It makes the call to urgently interrogate some of the ideologies behind those technologies, including issues of ethical and responsible development. The paper addresses four separate and interlinked issues: ideologies about deaf people and mediated communication, bias in data sets and learning, user feedback, and applications of the technologies. The paper ends with several take away points for both technology developers and deaf NGOs. Technology developers should give more consideration to diversifying their team and working interdisciplinary, and be mindful of the biases that inevitably creep into data sets. There should also be a consideration of the technologies' end users. Sign language interpreters are not the end users nor should they be seen as the benchmark for language use. Technology developers and deaf NGOs can engage in a dialogue about how to prioritize application domains and prioritize *within* application domains. Finally, deaf NGOs policy statements will need to take a longer view, and use avatars to think of a significantly better system compared to what sign language interpreting services can provide.

Introduction

In our everyday lives, we increasingly (and often unconsciously) rely on technologies where the languages we see, hear and produce are mediated in real-time by technology. Indeed, we are well into the human-machine era (Sayers et al., 2021). We talk to our devices using Amazon's Alexa and Apple's Siri, we read tweets in different languages through Twitter's automatic translation feature, deaf people use Google Live Transcribe, Ava, and other apps for real-time speech-to-text access. All these features were built on years and years of human work, and years and years of training of machines. We know and accept that some of these features are far from perfect yet, but we use them anyway. Because these AI applications feed on data and our frequent use, technology is advancing quickly and improving all the time. Machines learn.

The last three decades have seen sign languages, and deaf people, who use these languages, as target groups of language technologies, being included in these efforts in various ways. This includes developments in automated translation from text-to-sign (e.g. Stoll et al., 2020), speech-to-sign (e.g. Cox et al., 2002; Glauert et al., 2006), or sign-to-text (currently still very limited, e.g. Camgöz et al., 2020a, b). The technology is being developed in the form of many existing technologies, for example wearable solutions like smart gloves and intelligent bots with sign language avatars (virtual humans). This makes it increasingly likely that we will sign to and through technology.

[1] The writing of this paper has benefited from ongoing discussions on this theme as part of the EU COST Action network 'Language in the Human Machine Era' https://lithme.eu/

Some of the technological solutions under development are campaigned and marketed as trailblazing, advanced developments, using cutting edge technologies. Inventors can be big tech companies but just as well hearing undergraduates who don't know any sign language. A good example is the robot-arm developed by hearing undergrad students who claimed this would advance deaf children's inclusion in education and could help "millions of deaf people". The only thing the robot arm could actually do, was rudimentary fingerspelling (Drewett, 2018). The rationales for developing these technologies 'for' deaf people often stem from a saviour complex ("we, abled people, need to help disabled people communicate"), and from "techno-solutionism" (Fleet, 2021) which leads to technology being developed uniformed by the lived experience of disabled people. Even so, these inventors often get significant press attention or even gain awards and grants from panels that do not include any signers (Lipomi, 2017). Most of these solutions are also one-way, placing the burden on deaf people (e.g. to wear smart gloves) so that 'hearing' people can understand them (Kouznetsova, 2016; Lu, 2016; Woodcock, 2012, 2020). There are also detectable ideologies behind sign-to-text technologies about the normativity of the spoken modality - in a sense, that deaf people's ideas make 'more sense' if converted to spoken form (Hill, 2013). Also, ideologies supporting the technologies reveal a lot about how deaf people are viewed and communication is normatively mediated.

Funded research projects in this field often claim that these technologies can assist with 'inclusion' of deaf people, 'social equality' and, finally, address the problem of the 'insufficient availability' and 'prohibitive costs' of sign language interpreting services. Recently funded and on-going projects in the EU attract quite a lot of funding, with the European Commission spending several millions of euro per project. Despite these claims, a lot of the technology is still notably limited in its development and usability. Also and importantly, much of the work has started and is on-going with minimal input from deaf communities (Erard, 2017).

This paper is not about the current technical limitations of language technologies for sign languages. These technologies have a lot of catching up to do compared to technologies for spoken languages. This will happen, one way or another. Instead, this paper addresses a more urgent issue: the ethical and responsible development of these technologies, specifically sign language avatars. This issue is currently virtually not discussed in the academic and practice community. Most publications on machine translation for sign languages are either technical accounts of how machine translation can work or uncritical technical evaluations of user experience (e.g. Kacorri et al., 2017, but see Quandt et al. 2021 for an exception). There are virtually no critical insights into ethical, societal, and ideological rationales for and consequences of technologies. Indeed, that discussion usually lags behind scientific innovation - ethical debates about new technologies often come after the fact of their use. But at least for sign languages, the lag time is becoming very long now, and there is a critical need to address some urgent questions.

Who invents the technologies, and what is their motivation for developing them? How are data being collected to make machines learn? Who evaluates the outcomes, and how? Is there an actual demand from the communities? Who are the end users and who decides that? Who benefits from these technologies, and who is at risk of being left behind? What are the current and potential future applications of those technologies? How will language rights keep pace with the development of language technologies? What are the ideologies behind these technologies?

This paper will mostly write from a Deaf Studies and sociolinguistics perspective, with a specific focus on sign language avatars (meaning communication towards sign language users). While I will semantically differentiate between the 'avatar' as the digital figure representing a (signing) person and the underlying 'translation engine' from or towards sign

languages, for reasons of simplicity I will occasionally use the term 'avatar' to include the underlying translation engine. These avatars can be created in different ways. The movement can be based on recorded motion capture from real human signers or can be based on computer-synthesized motions. More recent innovations use machine learning trained on existing video data to generate these avatars. Those are the innovations this paper will mainly focus on, discussing research projects that have recently ended or are on-going in the European Union and the UK, such as Content4All[2], EASIER[3], SignON[4], SignLab at University of Amsterdam (UvA)[5], and aID – Artificial Intelligence for the Deaf[6]). It will address four separate and interlinked issues: ideologies about deaf people and mediated communication, bias in data sets and learning, user feedback, and applications of the technology. The paper ends with a few take away points for both technology developers and deaf NGOs.

Ideologies about deaf people/mediated communication

Several of the above-mentioned projects start from specific ideologies about deaf people and mediated communication: the researchers perceive a problematic 'communication gap' or 'language barrier' between 'deaf' and 'hearing' people, and state that technologies can and should address this gap or barrier. The main aim of the SignON project, prominent on the home page and the funder page, is to "bridge the communication gap between Deaf, hard of hearing and hearing people". It will "cross the language barrier between Deaf sign language users, hard of hearing and hearing people. SignON will tear down this information barrier that currently exists." The EASIER project is "bridging the communication gap between the deaf and the hearing". The aiD project offers "AI solutions for communicating needs of deaf people". The SignLab at UvA highlights "breaking language barriers". Stoll et al. (2018, p. 891) see the facilitation of "easy and clear communication between the hearing and the Deaf" as the critical aim of text-to-sign technologies, stating "… there is no guarantee that someone whose first language is, for example, British Sign Language, is familiar with written English, as the two are completely separate languages" (p. 892).

Some projects then make the leap to stating that their research can mitigate the problem of the limited availability and prohibitive costs of sign language interpreting services. SignON sees sign language interpreters as "the main medium for signed-to-spoken, spoken-to-signed and signed-to-signed translation", and the availability and costs of these services are seen as "a limiting factor in communication between signers and non-signers".

Some of the projects and the literature explicitly state the aim of language technologies for sign languages is not to *substitute* human interpreters but aim to be there for when interpreters are not available. The Content4All project proclaims that "systems that can accurately translate and produce sign would be of use to the Deaf. Not to replace human interpreters, but to provide translation into native sign language when an interpreter is not available (Young, 2020). A Dutch newspaper reporting on the development of sign language avatars at UvA headlined their piece "Sign language interpreters are scarce. Therefore an Amsterdam

[2] September 2017-November 2020, H2020 funding €4.1 million, https://content4all-project.eu/ and https://cordis.europa.eu/project/id/762021

[3] January 2021-December 2023, H2020 funding €3.9 million, https://www.project-easier.eu/vision/ and https://cordis.europa.eu/project/id/101016982

[4] January 2021-December 2023, H2020 funding €5.6 million, https://signon-project.eu/ and https://cordis.europa.eu/project/id/101017255

[5] https://www.signlab-amsterdam.nl/index.html NWO VICI grant or a 5-year project, funding €1.5 million.

[6] December 2019-November 2023, H2020 funding €1.5 million, https://aideaf.eu/ and https://cordis.europa.eu/project/id/872139

institute is learning an avatar to sign" and reported the lead researcher saying that "a human interpreter is always to be preferred above an avatar, when such an interpreter is available" (Zijlmans, 2021). Sáfár and Glauert state that "… in the deaf community, there is often a fear that the hard won recognition of their sign language will result in moves to make machines take over the role of human interpreters" (Sáfár & Glauert, 2012). The World Federation of the Deaf and World Association of Sign Language Interpreters (2018) statement on the use of signing avatars is the only one that uses different words to describe the risks and challenges: "the difference in linguistic quality between humans and avatars is why WFD and WASLI caution against the use of signing avatars as a replacement for human signers (World Federation of the Deaf & World Association of Sign Language Interpreters, 2018). The availability issue stands against experiences of deaf people being offered video remote interpreting (VRI) instead of sign language interpreters on location in some contexts, and their objections to this (e.g. Collinson, 2018). So, in reality, the 'unavailability' argument is a red herring – live interpreters *will* be substituted for avatars.

Another, related, aspect, is that the benchmark used to evaluate the quality of sign language technologies are often, again, sign language interpreters. Sayers et al. (2021, p. 10) assert that "consensus among the Deaf community so far is that these [smart gloves, avatars] are a profoundly poor substitute for human interpreters". Content4All affirms that "generating translations of the same quality as a human interpreter is extremely challenging". Sign language interpreters, not deaf signers themselves, are thus seen as language models, and as the benchmark for accepted standards of language use. If anything, this shows low ambition and an inability to see who the technology is for. Sign language interpreters are not the end users, nor should they be the benchmark.

With spoken languages there is the recognition that machine translation is at the moment of inferior quality compared to human *translators*, but that aspect is much less foregrounded and emphasized compared to the sign language projects. This is because the situation is profoundly different for deaf people, who are made reliant (by policy, legislation, and normative views on the role of sign language interpreting services) on sign language interpreters in many aspects of their lives. For deaf people, language rights often are paramount to access to sign language interpreters *in the first place* (De Meulder, 2016). But sign language interpreting services are, in many cases, a Band-Aid solution. They are not scalable services, and not equally available to deaf people who use them. They mostly benefit those deaf people with certain interpreter-related privileges. Even so, the provision of sign language interpreting services has become the institutionally normative, often unquestioned, solution to grant deaf people access to education and public services (De Meulder & Haualand, 2021).

Data sets and bias

In the context of machine learning, and more specifically the subtopic of Natural Language Processing (NLP) most sign languages tend to belong to the category of 'low-resourced languages'. The 'low-resourced' aspect refers to a lack of available training data and the fragmentation of efforts in resource development (Sayers et al., 2021). Indeed, NLP applications require large datasets to be available on which to train new algorithms. As NLP falls into the category of 'supervised learning', the algorithms learn by example in the form of 'labels', which tell the algorithm what needs to be learned. To allow this, large datasets must be labelled — which is expensive, time consuming, and prone to error, which can introduce bias. If the dataset is not carefully curated, it is mainly through these labels that bias can sneak into the algorithm. For sign languages there is the additional issue of a different language modality, which makes data collection and machine training much more challenging than it is

between most spoken languages. Also, input in the form of text annotations in itself skews data sets because sign languages do not have a widely used written form (Jantunen et al., 2021).

Linked to this, there is the problem of generating data in a context of data sparsity and the risks for bias of language models. This is a common issue for NLP tools (Bender, 2019, 2021; Benjamin, 2020; Beukeboom and Burgers, 2017; Blodgett et al., 2020; Benjamin, 2020; Saleiro et al., 2020;). Dictionaries/lexicons for sign languages are developed for human use, not machine use (i.e. use by automated NLP systems). Digital sign language corpora have mostly focused on linguistic aspects (how signing is used) rather than computational processing (data from tracking movements, facial expressions, timing, etc.) (Sáfár & Glauert, 2012). These corpora are confronted with other problems that make them far less suitable for machine use linked to size and representativeness, and variety of discourses (Jantunen et al., 2021; Schembri & Cormier, 2022). In the absence of (semi-) automated annotation, manual annotation is demanding and time-consuming. This means that while for some sign languages there *is* a set of videorecorded data (although still small compared to most data sets for spoken languages), these are not suitable for machine learning because they are not, or only partially, annotated. Even if there is annotation, some of the larger sign language corpora currently only have basic annotation such as glosses and possibly translations but no other tags that can be provided by semi-automatic tagging tools (Hochesang, 2021; Schembri & Cormier, 2022).

On the other hand, we have to resign to, and therefore deal with, the realisation that machines are and will be trained on those corpora, which in themselves by design contain all the biases of the humans who design and assemble them (Saleiro et al., 2020). The largest sign language corpora now have participants numbering in the hundreds, but are often skewed by a native speaker bias, preferring focus on (often white) deaf native signers or early learners, who often went to residential schools (e.g. Schembri et al., 2013, for BSL). Jantunen et al. (2021, p. 4-5) go so far as to say that “the contribution of novices and non-native signers means decreased quality and accuracy” in corpora, and “to increase validity and recognition systems should be trained with real (native or near native) signers in realistic scenarios”. At the same time, they state that “datasets should include representative, generalizable samples from diverse age groups, gender, culture, various ethnicities, varying body types and physical traits, clothes, lighting conditions and more”. Work has been done on the development of more specialised corpora (see Schembri & Cormier, 2022). These focus on e.g. L2 learners (Mesch & Schönström, 2018), but not (yet) on, for example, signers from different racial and educational backgrounds, signers with immigration backgrounds, language deprivation, various disabilities, etc.

Due to this context of data scarcity, some of the on-going EU-funded projects start to collect their own data sets to train machines, using readily available internet data. In some cases, these are interpreted datasets. In these datasets both the signed input from the interpreter and the spoken source languages are available (or in the opposite case, the signed source language and the spoken output from the interpreter). This is also made possible by the COVID-19 pandemic, which led to an increase in recorded interpreted presentations, classes, press conferences etc. which are often available online. These datasets (with both deaf and hearing interpreters) are used as a training phase for machines to quickly enlarge the dataset.

In this stage already there is a significant risk for how bias can creep into the system when machines are trained on amalgams of data sets with input produced by either primarily white, native signers, or by interpreters. This is even more cause for concern combined with two related issues. The first is that sign language interpreters are often *already* language models for deaf learners in regular education (Caselli et al., 2020) and deaf people *already* often need

to adapt their signing to be understood by interpreters. The second is that there is a realistic future possibility that signing avatars are going to be used to train sign language interpreting students. This would happen in an already problematic context of sign language interpreting training programs in general not reflecting racial diversity and multicultural, multilingual deaf community needs (Robinson et al., 2020).

User feedback

Thirdly, there is the issue of user feedback. When prototypes of avatars are developed, the developers need feedback from end users. In two cases (SignLab UvA and Content4All) this is/was done by an online survey in which (self-selected) deaf people are given specific tasks based on a pre-determined set of responses.

In the SignLab project deaf people are asked how well they understand translations by an animated avatar (using SiGML for sign language synthesis) compared to video translation by a deaf signer, specifically in a medical setting. Similarly, in the Content4All survey respondents look at an animated avatar and a deaf signer translating the weather forecast and are given specific tasks to evaluate comprehension. For example, they have to indicate which Dutch words they understood from the signing, given multiple options, answer questions (e.g., where do the clouds go to, what did the moderator suggest doing tomorrow?) and then indicate how sure they are of their response. At the start of the survey, it is specifically asked to give opinions “on comprehension of the signs rather than the look of the avatar” – as if the two can be separated.

Although asking user feedback is important, there are also several issues that must be addressed, and it is here where interdisciplinary approaches and specifically input from Deaf Studies researchers is most critically needed. Most deaf people have a life-long experience understanding different signing styles, of widely varying quality (see also Green, 2014 and Kusters et al., 2020 work on ‘understanding’). The risk with asking this kind of user feedback is that deaf people will see avatars’ signing as another signing style they’ll have to put up with and learn to ‘understand’ (just as they need to learn to understand interpreters’ signing). This can lead to socially desirable responses. This is related to what Woodcock (2020) in this context calls a “mouse on the doormat design” and the savior complex of some inventors: respondents might say they understood just because they think they are expected to appreciate this technology that is made ‘for’ them. A third issue is the uncanny valley (Mori et al., 2012) which might make viewers uncomfortable when confronted with simulations that closely resemble humans but are not quite convincing enough. We regard a Toy Story character which is obviously not human, as cute, but an avatar which is meant to be human but is not, as creepy. This combined with most deaf respondents not having realistic (or just not having any) expectations about avatars (see also Sáfár & Glauert, 2012). While some deaf people are trained to manage expectations about sign language interpreters (knowing the limitations of interpretation) and may tolerate the limitations of Google Translate or existing speech-to-text technologies, most deaf people are not yet used to manage expectations about the robotic and unarticulated signing of most avatars. This means that either expectations can be too low so that ratings will be higher than reasonably justified, or that on the contrary expectations are too high. Add to this the lack of testing in real-world settings. Indeed, there is a big difference from watching an avatar from a screen in your own office for a short experiment, and having to watch it during a nerve-wrecking medical appointment. This is not just unintentionally creepy entertainment in the uncanny valley, when you can look away (Woodcock, 2020). If your health depends on it, you cannot afford to look away. Evaluations do not account for this.

Applications

A last issue are current and future applications of language technologies for sign languages, such as avatars. The use of those technologies in the context of highly constrained and predictable domains such as tourism or travelling could be reasonably justified. An avatar might be used to make an order in a coffee house, check in in a hotel, have access to announcements at airports or railway stations, or for interaction with specific customer services. Another potentially useful application of avatars could be in those situations where sensitive, confidential information must be shared (by a deaf person themselves or a human interpreter) and where this person prefers to remain anonymous, and in areas where deaf people feel uncomfortable about the presence of interpreters (see also Quandt et al. 2021).

However, some starter projects beginning with very fundamental limitations as outlined above, at this point already go directly into sensitive areas such as the medical domain (e.g. Roelofsen et al., 2021[7]). This happens despite explicit statements and warnings by authoritative deaf NGOs such as the WFD and WASLI that the medical domain is a no-go area. These NGOs expressed concerns on the use of avatars "when the information being delivered is live, complex or of significant importance to the lives of deaf citizens" (World Federation of the Deaf & World Association of Sign Language Interpreters, 2018). Applications in the medical domain have been spurred by the COVID-19 pandemic, which exposed communication problems between health care professionals and deaf people when everyone had to wear facemasks, interpreters were often not allowed in hospitals and interpreting via video relay was not always viable.

That this is happening in the first place, is again linked to how deaf people are viewed, and how communication is currently normatively mediated by sign language interpreters. Sign language interpreters, on location or remotely, are accepted for mediated healthcare communication, despite critical limitations (Kushalnagar et al., 2019). Because this practice is largely accepted and even normative, use of language technologies in healthcare situations is seen as the logical next step and as a justified application domain by technology developers.

Conclusion

This paper has identified some common and specific pitfalls in the development of sign language technologies targeted at deaf communities and has made the call to urgently interrogate the ideologies behind those technologies, including issues of ethical and responsible development. What has been done technologically so far is very promising, but if continued on the same path, there is a risk that technologies developed in the end will not be voluntarily adopted by end users. This uptake in use is important, because the more 'we' use AI, the better it will become. There must be a consideration though of who this 'we' is – who is the language technology for, and why? Sign language interpreters are not the end users here, nor should they be seen as the benchmark for language use. Placing interpreters on the centre of deaf peoples' lives (a constructed dependence) comes from a biased and hearing-centred view on communication.

For the technology developers, this paper makes the call to *diversify the team* and *work interdisciplinary.* Co-design or co-engineer (see also Jantunen et al., 2021) with the end users

[7] For demo see here https://www.signlab-amsterdam.nl/healthcare-demo.html

of the language technologies, a widely varied group of deaf people. Not just to ensure semantic value, but also to ensure that technologies are being developed incorporating the communities demands, values, and feedback, that there is given consideration to interaction design and user interfaces. Engage in interdisciplinary collaborations, combining Deaf Studies, language policy, sign linguistics, Sign Language Interpreting Studies, computer science, sociolinguistics, Human Computer Interaction, artificial intelligence, computational linguistics, etc., not only in the execution of research projects but also in writing and reviewing them.

Regarding data sets, this paper is not meant as a discouragement for researchers from advancing the state of the art. If there is a need to wait for the 'perfect' datasets to appear (which might never happen) instead of using what is readily available, the delay for practical working solutions might become much longer. This is a call though to at the very minimum *be mindful of the biases* that will inevitably creep into data sets, and to consider the long-term implications of this.

Regarding applications, there are some take-away points both for developers of technologies and for deaf NGOs who need to evaluate their use and application domains. At this point, many projects are using signing avatars to do many different things, in many different ways – some of which are probably less problematic, and some of which are more. A deaf tech developer and artist working on incorporating signed language into VR spaces is a very different development and application compared to a hearing non-signing engineer developing a signing avatar without any consultation or collaboration with deaf communities. For the developers of technologies and for deaf NGOs there are two takeaway points:
(1) *Prioritize the application domains*: there is a significant distinction to be made between for example an avatar presenting information on a government webpage, or an avatar used to mediate communication during a life-threatening healthcare situation. There is a lot of unproblematic low-hanging fruit: it is thus important to identify those research agendas which are problematic, while leaving space for those who are not. This is even more a case for deaf academics to (co-)lead these projects and to involve deaf people in various roles, in various stages of review of project applications, project development, project execution, and evaluation. This will make it possible to identify early which research agendas *are* problematic, and how this can be potentially mitigated, and which research agendas are worth pursuing. This will advance the state of the art in such a way that it is more likely technologies will be adopted by end users.
(2) *Prioritize within application domains*: building on the previous point, it will be critical to make much finer distinctions per different uses per domain. For example it is not helpful to tar all applications in the medical domain with the same brush. Some might be useful and necessary, while others might remain a no-go.

For the deaf NGOs there are two further points to keep in mind.
(a) *Look at the horizon*: statements by for example WFD and WASLI are based on the *current* state of the art, which appropriately advises against the use of avatars "when the information being delivered is live, complex or of significant importance to the lives of deaf citizens" (World Federation of the Deaf & World Association of Sign Language Interpreters, 2018). The technologies developed so far just do not warrant use of avatars in those situations. But this is not a status quo, and technology is advancing all the time. Policy statements will need to take a much longer view. Here again, the questions from the introduction of this article come to the foreground. What are potential *future* applications of those technologies? Not based on the technologies as they are now, but based on how they will inevitably develop? How will language rights keep pace with the development of language technologies? Who benefits from these technologies, and who is at risk of being left behind?

(b) *Use avatars to think of a significantly better system*: much of the current ideological framework by which the use of avatars is assessed is based on experiences with sign language interpreting services. Deaf NGOs appropriately advise against use of avatars in situations where it is not warranted, but at the same time, sign language interpreting services are often used in situations where they are not warranted either. The political institution of sign language interpreting services leaves a lot of questions to be considered regarding scalability and fairness. Let's not substitute one imperfect system with another. Let's use this moment in time, these technological possibilities, to try and design a better system.

Acknowledgments

Thanks to Jan Adriaenssens, Lorna Quandt, Joseph Murray, Tanguy Coenen, Kristin Snoddon, and the two anonymous reviewers for their generous feedback on previous drafts of this paper.

References

Bender, E. (2019). A Typology of Ethical Risks in Language: Technology with an Eye Towards Where Transparent Documentation Can Help. https://faculty.washington.edu/ebender/papers/Bender-A-Typology.pdf

Benjamin, R. (2020). 2020 Vision: Reimagining the Default Settings of Technology & Society. https://iclr.cc/virtual_2020/speaker_3.html

Beukeboom, C. J., & Burgers, C. (2017). Linguistic bias. In H. Giles & J. Harwood (Eds.), *Oxford Encyclopedia of Intergroup Communication*. Oxford University Press.

Blodgett, S. L., Barocas, S., Daumé, H., & Wallach, H. (2020). *Language (technology) is power: A critical survey of bias.* In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 454-476).

Camgöz, N.C., Koller, O., Hadfield, S., Bowden, R. (2020a). Multi-Channel Transformers for Multi-Articulatory Sign Language Translation. In *European Conference on Computer Vision,* 301-19.

Camgöz, N.C., Koller, O., Hadfield, S., Bowden, R. (2020b). Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation. In *Proceedings of the Ieee/Cvf Conference on Computer Vision and Pattern Recognition,* 10023-33.

Caselli, N. K., Hall, W. C., & Henner, J. (2020). American Sign Language Interpreters in Public Schools: An Illusion of Inclusion that Perpetuates Language Deprivation. *Maternal and Child Health Journal*, *24*, 1323-1329.

Collinson, A. (2018). The deaf patients 'left behind' by the NHS. https://www.bbc.com/news/health-44384503

Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., & Abbott, S. (2002). *TESSA, a system to aid communication with deaf people*. In *Proceedings of the 5th international ACM conference on assistive technologies* (pp. 205-212).

De Meulder, M. (2016). *The Power of Language Policy* [PhD]. University of Jyväskylä.

De Meulder, M., & Haualand, H. (2021). Sign language interpreting services: A quick fix for inclusion? *Translation and Interpreting Studies*, *16*(1), 19-40. https://doi.org/10.1075/tis.18008.dem

Erard, M. (2017). *Why sign-language gloves don't help deaf people*. Retrieved 2019-05-16 from https://www.theatlantic.com/technology/archive/2017/11/why-sign-language-gloves-dont-help-deaf-people/545441/

Fleet, C. [@ChanceyFleet] (2021, June 27). *What's true of other marginalized groups is true of us: techno-solutionism, uninformed by our wisdom and lived experience, can only harm us and waste investors' dollars and reputations. Stop it.*[Tweet] https://twitter.com/ChanceyFleet/status/1409221969166352393

Glauert, J. R. W., Elliott, R., Cox, S. J., Tryggvason, J., & Sheard, M. (2006). VANESSA – A System for Communication between Deaf and Hearing People. *Technology and Disability*, *18*, 207-216.

Green, E. M. (2014). Building the tower of Babel: International Sign, linguistic commensuration, and moral orientation. *Language in Society*, *43*(4), 445-465.

Hill, J. (2013). Language ideologies, policies, and attitudes towards signed languages. In R. Bayley, R. Cameron, & C. Lucas (Eds.), *The Oxford Handbook of Sociolinguistics* (pp. 680-697). Oxford University Press.

Hochesang, J. (2021). Documenting Language Use of the ASL Communities. https://www.youtube.com/watch?v=CJtfRQ0tsfM

Jantunen, T., Rousi, R., Rainò, P., Turunen, M., Valipoor, M. M., & García, N. (2021). Is There Any Hope for Developing Automated Translation Technology for Sign Languages? In M. Hämäläinen, N. Partanen, & K. Alnajjar (Eds.), *Multilingual Falicitation* (pp. 1-13). University of Helsinki.

Kacorri, H., Huenerfauth, M., Ebling, S., Patel, K., Menzies, K., & Willard, M. (2017). Regression Analysis of Demographic and Technology-Experience Factors Influencing Acceptance of Sign Language Animation. *ACM Transactions on Accessible Computing (TACCESS)*, *10*(1), 3-33. https://doi.org/10.1145/3046787

Kouznetsova, S. (2016). Why the Signing Gloves Hype Needs to Stop - Audio Accessibility. https://audio-accessibility.com/news/2016/05/signing-gloves-hype-needs-stop/

Kushalnagar, P., Paludneviciene, R., & Kushalnagar, R. (2019). Video Remote Interpreting Technology in Health Care: Cross-Sectional Study of Deaf Patients' Experiences. *JMIR Rehabil Assist Technol*, *6*(1), 1-8.

Kusters, A., Green, M., Moriarty Harrelson, E., & Snoddon, K. (2020). Sign language ideologies: Practices and politics. In A. Kusters, M. Green, E. Moriarty Harrelson, & K. Snoddon (Eds.), *Sign Language Ideologies in Practice*. De Gruyter Mouton.

Lipomi, D. (2017). *Lessons from working with the press: Cultural considerations and terminology surrounding American Sign Language in engineering research*. Retrieved 2019-05-16 from https://www.lipomigroup.org/blog/2017/9/12/cultural-considerations-and-terminology-surrounding-american-sign-language-in-materials-research

Lu, A. (2016). Deaf People Don't Need New Communication Tools — Everyone Else Does. *httpsmedium.com*. https://medium.com/@alexijie/deaf-people-dont-need-new-communication-tools-everyone-else-does-df83b5eb28e7

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138-163.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [from the field]. *IEEE Robotics & Automation Magazine*, *19*, 98-100.

Robinson, O. E., Sheneman, N., & Henner, J. (2020). Toxic Ableism Among Interpreters: Impeding deaf people's linguistic rights through pathological posturing. In C. McDermid, S. Ehrlich, & A. Gentry (Eds.), *Conference proceedings of the 2019 WASLI conference* (pp. 14-41). WASLI.

Roelofsen, F., Mende-Gillings, S. E., Esselink, L. D., & Smeijers, A. S. (2021). Supporting dialogue between healthcare professionals and Deaf patients through automated text-to-sign translation. *https://www.signlab-amsterdam.nl/publications/COVID-19.pdf*.

Sáfár, E., & Glauert, J. (2012). Computer modelling. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language. An International Handbook* (pp. 1075-1101). De Gruyter Mouton.

Saleiro, P., Rodolfa, K. T., & Ghani, R. (2020). *Dealing with Bias and Fairness in Data Science Systems: A Practical Hands-On Tutorial.* In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 3513-3514).

Sayers, D., Sousa-Silva, S., & Höhn. (2021). The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies. *Report for EU COST Action CA19102 'Language in the Human-Machine Era'*. https://doi.org/10.17011/jyx/reports/20210518/1

Schembri, A., & Cormier, K. (2022). Sign language corpora: future directions. In J. Hochesang & J. Fenlon (Eds.), *Sign language corpora*. Gallaudet University Press.

Stoll, S., Camgoz, N. C., Hadfield, S., & Bowden, R. (2020). Text2Sign: Towards Sign Language Production Using Neural Machine Translation and Generative Adversarial Networks. *International Journal of Computer Vision*, *128*, 891-908. https://doi.org/https://doi.org/10.1007/s11263-019-01281-2

Woodcock, K. (2012). Pick up a pencil. https://safeandsilent.wordpress.com/2012/07/11/pencil/

Woodcock, K. (2020). Paved with good intentions. https://safeandsilent.wordpress.com/category/technology/

World Federation of the Deaf & World Association of Sign Language Interpreters. (2018). WFD and WASLI Issue Statement on Signing Avatars - WFD. https://wfdeaf.org/news/wfd-wasli-issue-statement-signing-avatars/

Young, G. (2020). Automated Sign Language to produce signs helping the Deaf, not the hearing…. https://content4all-project.eu/automated-sign-language-to-produce-signs-helping-the-deaf-not-the-hearing/

Zijlmans, M. (2021). – *De Volkskrant*. Retrieved 2021-06-17 from https://www.volkskrant.nl/wetenschap/gebarentolken-zijn-schaars-daarom-leert-een-amsterdams-instituut-een-avatar-gebarentaal~b4449e2b/