

Ableist Language *Teching* over Sign Language Research

Carl Börstell

Dept. of Linguistic, Literary and Aesthetic Studies (LLE)

University of Bergen (UiB)

carl.borstell@uib.no

Abstract

The progress made in computer-assisted linguistics has led to huge advances in natural language processing (NLP) research. This research often benefits linguistics in a broader sense, e.g., by digitizing pre-existing data and analyzing ever larger quantities of linguistic data in audio or visual form, such as sign language video data using computer vision methods. A large portion of research conducted on sign languages today is based in computer science and engineering, but much of this research is unfortunately conducted without any input from experts on the linguistics of sign languages or deaf communities. This is obvious from some of the language used in the published research, which regularly contains ableist labels. In this paper, I illustrate this by demonstrating the distribution of words in titles of research papers indexed by Google Scholar. By doing so, we see that the number of tech papers is increasing, while the number of linguistics papers is (relatively) decreasing, and that ableist language is more frequent in tech papers. By extension, this suggests that much of the tech-related work on sign languages – heavily under-researched and under-resourced languages – is conducted without collaboration and consultation with deaf communities and experts, against ethical recommendations.

1 Introduction

Sign language linguistics is a young field, with its inception in the 1960s (McBurney, 2012). Due to the modality of sign languages, using the visual channel for linguistic signals, storing data of signing in its true form has only been possible for as long as video recording has been possible. The analysis of large-scale sign language datasets has consequently relied on an increase in digital storage capacity, and even more recent advances in

computer vision and AI technology have led to a growing interest in analyzing sign languages using methods from computer science and natural language processing (NLP) (Bragg et al., 2019, 2021; Yin et al., 2021; Shterionov et al., 2022).

However, while bias and ableism are discussed in NLP and AI research more broadly (Bender et al., 2020; Shew, 2020; Hassan et al., 2021; Kamikubo et al., 2022), a substantial part of the research on sign languages in these fields is conducted without including or consulting deaf communities and experts. Such inclusive collaboration is crucial for ethical research on sign languages and their communities (Harris et al., 2009; Kusters et al., 2017; Hill, 2020; De Meulder, 2021; Hochgesang, 2021b; Hochgesang and Palfreyman, 2022; SLLS), and research without it can easily lead to useless, exploitative and even damaging outcomes, despite good intentions. Shew (2020, 43) introduces the term *technoableism*, defined as the “rhetoric of disability that at once talks about empowering disabled people through technologies while at the same time reinforcing ableist tropes about what body-minds are good to have and who counts as worthy”, which in the context of deaf people can be the development of technical tools – e.g., hearing aids, “sign language gloves” and signing avatars – without consulting the intended users about whether such tools are even wanted (see Hill, 2020).

One key identifier of sign language research being conducted without consideration for, and input from, deaf communities and experts is the use of *ableist* language in the research itself, which sign language linguists observe frequently in published papers. The term *ableist language* is used here to mean language referring to disabled people – in this paper particularly deaf and hard-of-hearing people – from the perspective of abled people, equating *disabled* with ‘deficient’. In this paper, ableist language refers specifically to words

or phrases used as labels or descriptors of deaf and hard-of-hearing people – NB: these offensive labels will be mentioned explicitly in this paper. In the context of deaf and hard-of-hearing communities, particularly those in North America and Europe, labels such as “deaf-mute”, “deaf and dumb” and “hearing impaired” are considered ableist and offensive. The terms listed here are generally considered ableist and offensive in English, although the direct translations into other languages may be judged differently by the local deaf community.

One specific paper on sign language technology garnered some attention after heavy criticism due to the offensive, ableist language used in it. The criticism was outlined in an open letter to the publisher – signed by over 100 academics, professionals and community members – questioning the decision to publish work with highly offensive, ableist language in its title, and resulted in the publisher officially retracting the paper (Hochgesang, 2021a). Nonetheless, sign language research papers containing ableist language continue to be published, which is evident from notifications of new publications in the field from services such as Google Scholar Alerts. Thus, in this paper, I¹ look at the impact of tech-related fields on current sign language research, and how it correlates with ableist language, using data from Google Scholar and Google Scholar Alerts.

2 Method

The data for this paper come from two overlapping sources: 1) Google Scholar’s online search interface and 2) Google Scholar Alerts email notification service. I, the author, have had a Google Scholar Alerts notification for the term “sign language” since October 2017, and the data spans all Google Scholar Alerts from October 6, 2017 to February 27, 2023. The email data is also put in an overall context compared to data on sign language papers searched through the online Google Scholar interface, searched systematically by annual intervals from 2012 to 2022.

2.1 Searching Google Scholar

To obtain rough figures of sign language publications indexed in Google Scholar, the online interface was searched on March 10, 2023 for "sign

¹I, the author, am a hearing, signing linguist from – and currently based in – Scandinavia, working mainly on the linguistic structure and use of sign languages using quantitative methods.

language" for each year as the defined search interval, from 2012 to 2022 – see Table 1.

Year	Papers
2012	10 700
2013	12 300
2014	12 600
2015	12 800
2016	13 600
2017	14 600
2018	15 100
2019	16 400
2020	17 000
2021	19 000
2022	18 700
Total	162 800

Table 1: Number of papers on Google Scholar matching “sign language”, from 2012 to 2022.

These numbers thus constitute the total number of papers on sign language per year, which were then complemented by a delimited search per year using the operator “+” and one of the keywords "ai", "computer" and "glove" (for tech-related papers) and "grammar", "linguistics" and "morphology" (for linguistics-related papers). A broad search for the entire span 2012 to 2022 was also conducted using the search strings ["sign language" + "computer" -"linguistics"] and ["sign language" + "linguistics" -"computer"], each with one of the additional words "deaf-mute" and "dumb", prefixed with “+” or “-”, to see the proportion of papers including (or not) these two ableist terms. I deliberately chose these two ableist terms, as they are specifically targeting deaf and hard-of-hearing people and are the most overtly offensive. Additional terms that can be considered ableist, such as “(ab)normal” or “impair(ed|ment)”, are harder to know if they are in fact referring to deaf or hard-of-hearing people, or at all referencing disabilities (e.g., “normal distribution” in statistics). However, such additional terms are included in the Google Scholar Alerts analysis (see Section 2.2), where it is possible to also conduct a manual check of their use.

2.2 Processing Google Scholar Alerts

The Google Scholar Alerts emails were downloaded, processed and analyzed using R 4.2.2

(R Core Team, 2022) and the packages `rvest` (Wickham, 2022), `scales` (Wickham and Seidel, 2022), `tidytext` (Silge and Robinson, 2016) and `tidyverse` (Wickham et al., 2019). In total, 832 notification emails were processed, extracting the titles of each listed paper, resulting in a total of 8368 papers – see Table 2.

Year	Papers
2017	360
2018	1488
2019	1500
2020	1710
2021	1509
2022	1556
2023	245
Total	8368

Table 2: Number of papers from Google Scholar Alerts for “sign language”, from October 6, 2017 to February 27, 2023.

Each paper title was scanned for keywords to annotate them as **tech** if the title string matched any substring in

```
(^| )ai($| )|
(^| )app($| |s|lication)|
android|arduino|artificial|
automat|cnn|comput|controller|
convert|convolutional|deep|
device|glove|
(^| |\\()ml($|\\-| |\\))|
machine|neural|nlp|python|
raspberry|recognition|
real-time|sensor|software|
system|tech|tensorflow|
to( |\\-)text|to( |\\-)speech|
transformer|tool|virtual|
vocaliz|wearable
```

and **linguistics** if the title matched any substring in

```
claus(al|e)|communica|
conversation|corpus|
discourse|gramma(r|tic)|
iconic|interaction|
linguist|morpholog|
neuroling|object|
phon(etic|olog)|psycholing|
semantic|socioling|
subject|synta(x|ct)
```

and each paper was labeled as **ableist** if the title matched any of the following substrings:

```
abnormal| normal($| )|
(and |and|deaf|\\-| )mute|
dumb($| )|( |-)impair
```

It should be noted here that this search method is crude and it is possible that some titles categorized as ableist are in fact using the terms metalinguistically, problematizing the terms rather than using them unquestioned. However, it might be expected that the metalinguistic use of ableist language would be higher in papers relating to discourse analysis, social anthropology and disability studies, which are arguably closer to the humanities side of linguistics than computational approaches. Any matches for ableist terms were thus manually checked after the automated identification.

3 Results

3.1 Google Scholar

Looking first at the search results from the Google Scholar online interface, we can see from Figure 1 (as well as Table 1) that the absolute number of papers found using the search the term "sign language" is increasing steadily over the years from 2012 to 2022.² Using the total number of sign language papers as the baseline, we can compare the detailed search terms including also terms for concepts associated with **tech** or **linguistics** in Figure 2. Though not mutually exclusive, it is worth noting that in 2022, about half of the papers with "sign language" also contained "computer", whereas around a quarter contained "linguistics". From Figure 2, we can observe that whereas a term such as "ai" (i.e. *artificial intelligence*) is clearly increasing in the past five years with regard to the proportion of indexed sign language papers overall that contain the term, words associated with linguistics such as "grammar" and "morphology" appear to be decreasing relative to the total number of papers.

Combining the search term "sign language" with an additional search term "computer" or "linguistics" (mutually exclusive), with one of two well-known ableist terms either included or excluded, we can see from Figure 3 that a much larger proportion of papers including "computer" will simultaneously use an ableist term, compared to papers including

²Since the search was done in the first quarter of 2023, it is possible that some papers for 2022 had not yet been indexed and that the slight drop for 2022 may not be accurate.

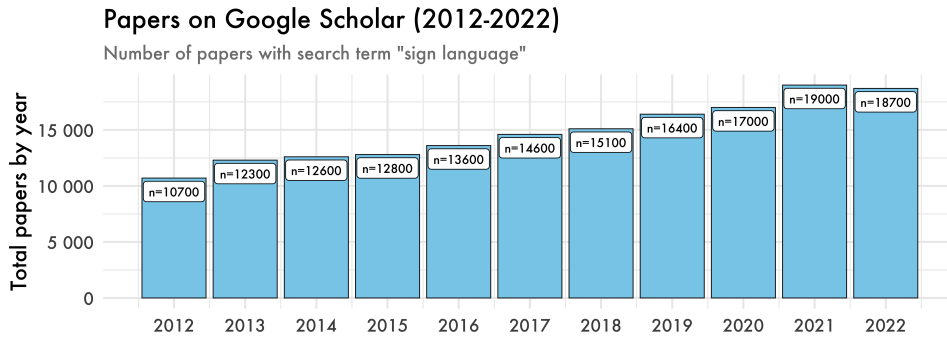


Figure 1: Number of papers indexed by Google Scholar in the years 2012 to 2022 found with the search term “sign language”.

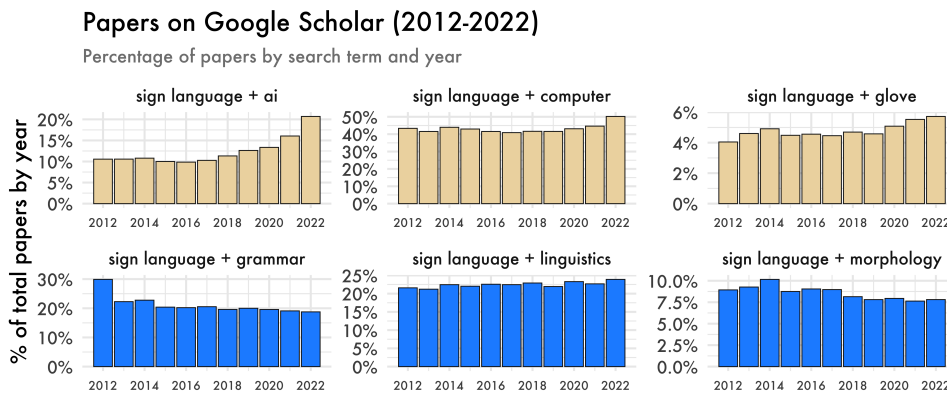


Figure 2: Proportion of papers indexed by Google Scholar in the years 2012 to 2022 found with the search term “sign language” and additional terms relating to tech or linguistics.

"linguistics". As many as 12% of the papers found searching "sign language" + "computer" simultaneously include the highly ableist term "dumb" (as in “deaf and dumb”), while papers that involve "linguistics" (but not "computer") only have around half of that proportion of papers with this ableist term. While it should be remembered that these terms in some cases may be used metalinguistically rather than referentially/generically, this does not explain why "computer" papers are more likely to include ableist terms. Rather, it is likely that the skewed distribution reflects a difference in whether or not the authors are aware of the fact that these terms are considered ableist and offensive.

3.2 Google Scholar Alerts

We now turn to the Google Scholar Alerts data, which involves 8368 papers listed in 832 notification emails (in digest form) from October 6,

2017 to February 27, 2023. Figure 4 shows the total number of papers per year from 2018 to 2022 (excluding 2017 and 2023 as incomplete years). From this, we can see that there has been a fairly even number of papers listed for the Google Scholar Alerts notification for “sign language”, with around 1500 papers annually. Comparing this to the proportion of papers with either **tech-** or **linguistics-**related terms in the titles, we can see that whereas tech-related papers appear to be on the rise in both absolute and relative numbers (Figure 5), linguistics-related papers show the opposite trend, generally decreasing over time (Figure 6). These numbers can be directly compared to Figure 7, which shows the proportion and number of titles containing ableist words, indicating a slow increase over time from 2017 to 2023. These results, also shown in Table 3, in themselves suggest that there is a correlation between the increase of tech-related papers and the number of ableist titles, and simultaneously an inverse corre-

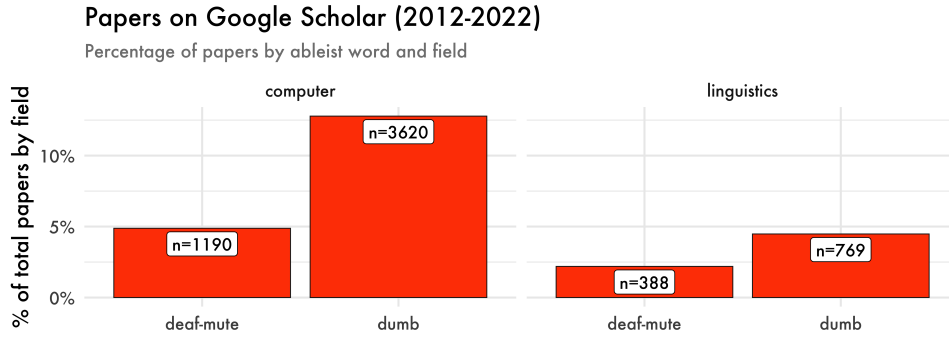


Figure 3: Number and proportion of papers indexed by Google Scholar in the years 2012 to 2022 found with the search term “sign language” with either “computer” or “linguistics” as added term (mutually exclusive) and including one of two well-known ableist terms.

lation between the decrease in linguistics-related papers and the number of ableist titles. While correlation does not imply causation, we can look at the direct overlap of titles containing ableist words and their categorization as either **tech**, **linguistics**, **both** (in case a title is simultaneously categorized as tech and linguistics) and **no label** (categorized as neither tech nor linguistics). Here, Figure 8 shows that tech-related papers are responsible for almost half of all observed titles using ableist language – the rest mostly being papers categorized as **no label**, meaning the titles do not contain any overtly tech- nor linguistics-associated words, but could still be either or both. Among these, Table 4 shows the frequency distribution of ableist words identified. The most commonly used word that can be considered ableist is *impair** (including word forms like *impaired*, *impairment*, etc.). While this word is still readily used in many contexts (e.g., medical literature), it is generally considered ableist due to its direct reference to “deficiency” in relation to a default “normal”.

Year	Total	Tech	Linguistics	Ableist
2017	360	102	53	14
2018	1488	458	223	54
2019	1500	464	195	66
2020	1710	548	259	75
2021	1509	576	184	74
2022	1556	666	139	59
2023	245	124	20	21

Table 3: Number of papers from Google Scholar Alerts for “sign language” from October 6, 2017 to February 27, 2023 that can be labeled as **tech** and/or **linguistics** and/or contain ableist words.

Word	Tokens
<i>impair*</i>	270
<i>dumb</i>	48
<i>mute</i>	42
<i>(ab)normal</i>	11
Total	371

Table 4: Frequency of ableist words in paper titles listed in Google Scholar Alerts emails from October 6, 2017 to February 27, 2023.

However, it is possible that this method of identifying ableist words and categorizing topics is not entirely accurate. For example, as discussed earlier, there could be cases of ableist terms being used metalinguistically, or that there are incorrect classifications of topics due to the keywords not covering all cases. The latter case is undoubtedly true, since Figure 8 showed a substantial portion of ableist titles categorized as **no label** (i.e. no topic keywords identified). Thus, I manually annotated the 363 titles that were identified as containing ableist titles with respect to three features: 1) whether the title contained an ableist term in an offensive way; 2) whether the title was correctly identified as **tech**; 3) whether the title was correctly identified as **linguistics**. An attempt was also made to categorize **no label**-titles into a named category. In this manual annotation, eight titles were removed due to formatting inconsistencies, for example if journal names or partial abstracts had been parsed incorrectly as part of the title, and the ableist term had only been identified *outside* the actual title. In the remaining 355 titles, 97.7% (n=347) had been correctly identified as directly ableist, the other eight being used ei-

ther metalinguistically/critically (n=1) or in borderline cases involving other medical terminology (n=7). 89% of the titles automatically categorized as **tech**, and 83.6% of the titles automatically categorized as *linguistics*, were confirmed as such by the manual annotation. Thus, the manual annotation reveals that the automated process is fairly accurate, and that the non-ableist uses of the selected terms are very marginal. In the manual annotation, all uncategorized (i.e. **no label**) cases were assigned a category. This manual re-categorization is shown in Figure 9, which illustrates that **tech**-related papers constitute the majority of titles with ableist terms, followed by **education**, **health** (e.g., public health research as well as psychology and medical studies) and **other** (e.g., acoustics, law, sociology). Figure 10 further corroborates this point, illustrating that 6–7% of papers categorized as **tech**-related contain ableist language in their titles, compared to around 3% of papers categorized as **no tech**. This clearly demonstrates a pattern of tech-related sign language papers being more likely to contain ableist language, based solely on the titles of the paper themselves, without even looking at the text content of the papers.

4 Discussion

Sign language linguistics is a young field (McBurney, 2012) and sign languages are both under-researched and under-resourced. Part of the under-resourced issue can be attributed to the visual modality and that technical advances were needed before the recording and analysis of sign language data became possible, but is also likely a result of their minoritized and marginalized status. In the past decade, the number of lexical databases and corpora for various sign languages has exploded (Fenlon and Hochgesang, 2022; Kopf et al., 2022). Additionally, there are ongoing efforts to include sign languages in the building of NLP and other language technology resources (Bragg et al., 2021; Yin et al., 2021; Morgan et al., 2022; Shterionov et al., 2022), although there are still many obstacles to overcome before achieving, e.g., machine translation of sign languages (Jantunen et al., 2021). While there is current NLP and language technology work that is being conducted ethically, by and in direct collaboration with deaf experts and stakeholders, in accordance with the ethical standards for research on deaf communities and sign languages (cf. Harris et al., 2009; De Meul-

der, 2021; Leeson et al., 2021; Hochgesang and Palfreyman, 2022), this paper has illustrated that much of the recent and ongoing work involving technical approaches to sign language research is ignorant of basic nomenclature in the context of deaf communities and their languages. By extension, the use of ableist language in scientific reporting does not instill confidence that the research itself, and resources and tools stemming from it, will be of any higher quality. This is because without engaging with and involving deaf scholars in the process of developing language resources, the input of deaf experiences and expertise is left out (cf. Kusters et al., 2017), which often leads to tools and solutions that are not acceptable or even wanted by the deaf community (see Hill, 2020; De Meulder, 2021), often summarized by a collective outcry of *not another sign language glove* (cf. Hochgesang, 2021a).

Unfortunately, many technical applications that target deaf and hard-of-hearing people – as well as other disabled people – often constitute a type of ableism labeled *technoableism*, in which so-called “solutions” that, while they may be well-intentioned, are rooted in ideas of “fixing” or “mending” any differences from the norm, regardless of whether those solutions are practical or requested by the target group (Shew, 2020). Similar questioning of normativity in language use and the study of language has been raised within linguistics in a broader sense (Cheng et al., 2021; Henner and Robinson, 2021; Namboodiripad and Henner, 2022), and feeds into the more narrow question of language technology and resources, both in terms of biases present in the underlying datasets (Kamikubo et al., 2022) and in terms of who is developing the technology and for whom (Hill, 2020; De Meulder, 2021; Leeson et al., 2021).

The results of this paper illustrate that many of the language resources and tools that are being developed for sign languages – which are all still very much under-studied and under-resourced languages – are likely developed without direct input from the stakeholders, and without sufficient background in the history and ethics of sign language research, as is evidenced by ableist language use. Besides being harmful, this also leads to a distrust in the usability of such tools and resources, and potentially also difficulties in finding and funding the quality work due to the overwhelming increase in publications and resources being developed.

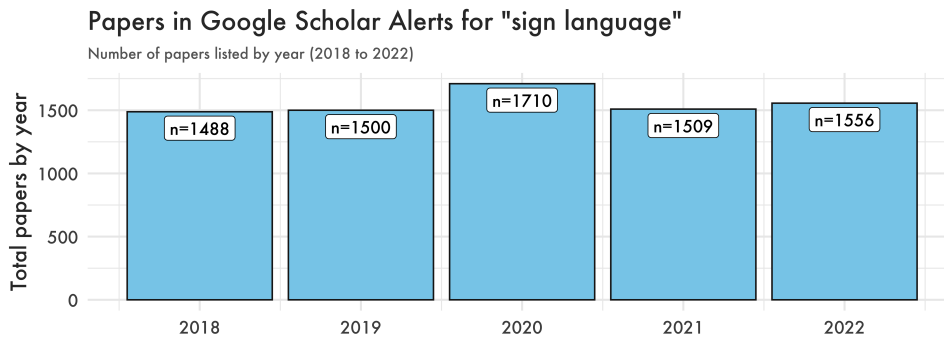


Figure 4: Number of papers listed in Google Scholar Alerts emails for the term “sign language” from 2018 to 2022.

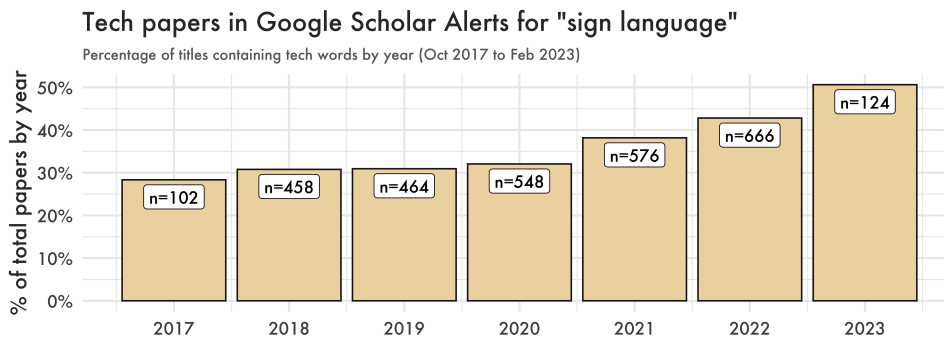


Figure 5: Number and proportion of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words identified as tech.

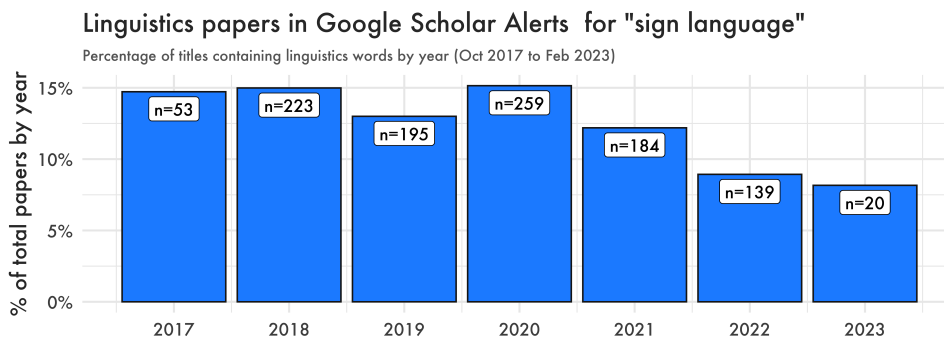


Figure 6: Number and proportion of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words identified as linguistics.

5 Conclusion

In this paper, I have shown how ableist language in publications on sign languages correlates with tech-related research. Besides being offensive in itself, thus leading to further oppression of deaf and disabled people, it suggests a low level of awareness of the actual wants and needs of deaf communities when it comes to technology, and

reinforces biases of both researchers (in terms of who is involved) and their output (in terms of resources and applications). Researchers working on developing language resources for any group, but particularly marginalized ones, should at the very least be expected to have enough knowledge and awareness about the context and history of the group to not reinforce offensive and oppressive

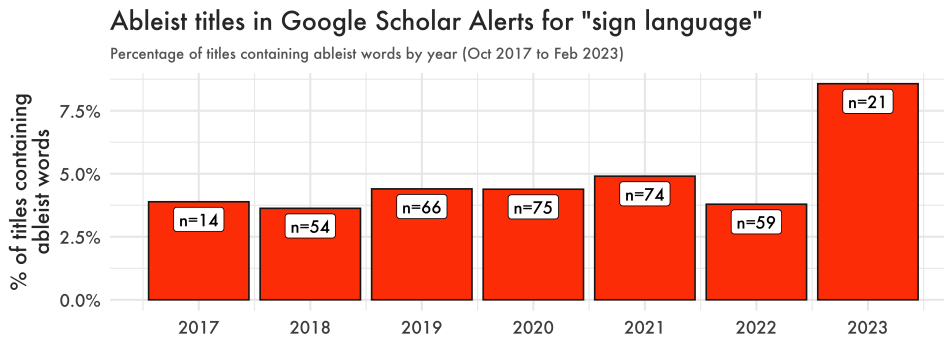


Figure 7: Number and proportion of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words identified as ableist.

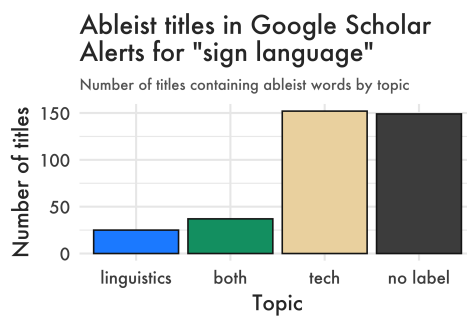


Figure 8: Number of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words identified as ableist by topic.

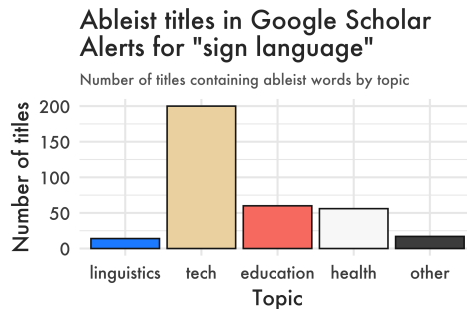


Figure 9: Number of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words *manually* identified as ableist by topic.

language use, but ideally also work in direct consultation and collaboration with the community in question. This has been outlined by many scholars advocating for inclusive and ethical research (e.g. Harris et al., 2009; Hill, 2020; De Meulder, 2021; Hochgesang, 2021b; Hochgesang and Pal-

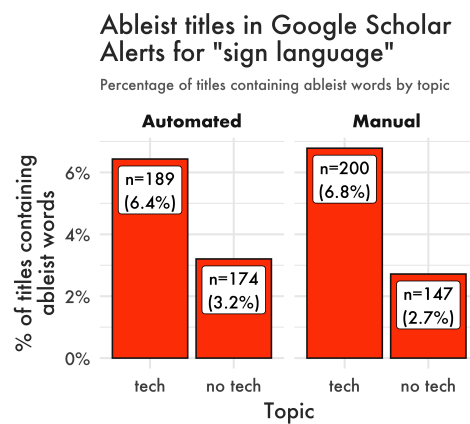


Figure 10: Proportion of papers listed in Google Scholar Alerts emails for the term “sign language” from October 6, 2017 to February 27, 2023 with titles containing words identified as ableist by topic divided into **tech** vs. **no tech** based on keywords in the titles, using both automated and manual methods of categorizing paper titles.

freyman, 2022), and among these, De Meulder (2021, 18) directly addresses researchers developing sign language technology, encouraging them to always ask themselves the question “who is the language technology for, and why?”

Acknowledgments

Thanks to Julie Hochgesang for providing feedback and comments on an earlier draft of this paper, and for sharing presentation slides on the topic of ethics in sign language research. I also thank three anonymous reviewers for their very valuable feedback.

References

- Emily M. Bender, Dirk Hovy, and Alexandra Schofield. 2020. Integrating Ethics into the NLP Curriculum. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, pages 6–9, Online. Association for Computational Linguistics.
- Danielle Bragg, Naomi Caselli, Julie A. Hochgesang, Matt Huenerfauth, Leah Katz-Hernandez, Oscar Koller, Raja Kushalnagar, Christian Vogler, and Richard E. Ladner. 2021. The FATE Landscape of Sign Language AI Datasets: An Interdisciplinary Perspective. *ACM Transactions on Accessible Computing*, 14(2):1–45.
- Danielle Bragg, Oscar Koller, Mary Bellard, Laran Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. 2019. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pages 16–31, Pittsburgh PA USA. ACM.
- Lauretta S. P. Cheng, Danielle Burgess, Natasha Vernooij, Cecilia Solís-Barroso, Ashley McDermott, and Savithry Nambodiripad. 2021. The Problematic Concept of Native Speaker in Psycholinguistics: Replacing Vague and Harmful Terminology With Inclusive and Accurate Measures. *Frontiers in Psychology*, 12:715843.
- Maartje De Meulder. 2021. Is “good enough” good enough? Ethical and responsible development of sign language technologies. In *Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, pages 12–22, Virtual. Association for Machine Translation in the Americas.
- Jordan Fenlon and Julie A. Hochgesang, editors. 2022. *Signed Language Corpora*. Number 25 in Sociolinguistics in deaf communities. Gallaudet University Press, Washington, D. C.
- Raychelle Harris, Heidi M. Holmes, and Donna M. Mertens. 2009. Research Ethics in Sign Language Communities. *Sign Language Studies*, 9(2):104–131. Publisher: Gallaudet University Press.
- Saad Hassan, Matt Huenerfauth, and Cecilia Ovesdotter Alm. 2021. Unpacking the Interdependent Systems of Discrimination: Ableist Bias in NLP Systems through an Intersectional Lens. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3116–3123, Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Jon Henner and Octavian Robinson. 2021. Unsettling Languages, Unruly Bodyminds: Imaging a Crip Linguistics. preprint, PsyArXiv.
- Joseph Hill. 2020. Do deaf communities actually want sign language gloves? *Nature Electronics*, 3(9):512–513.
- Julie Hochgesang. 2021a. Open Letter to Springer Editors and Their Response.
- Julie A. Hochgesang. 2021b. Ethics of working with signed language communities. In *AI & Sign Language Convention 2021*, Washington, DC.
- Julie A. Hochgesang and Nick Palfreyman. 2022. Signed Language Corpora and the Ethics of Working with Signed Language Communities. In Jordan Fenlon and Julie A. Hochgesang, editors, *Signed Language Corpora*, number 25 in Sociolinguistics in deaf communities, pages 158–195. Gallaudet University Press, Washington, D. C.
- Tommi Jantunen, Rebekah Rousi, Päivi Rainò, Markku Turunen, Mohammad Moen Valipoor, and Narciso García. 2021. Is There Any Hope for Developing Automated Translation Technology for Sign Languages? In Mika Hämäläinen, Niko Partanen, and Khalid Alnajjar, editors, *Multilingual Facilitation*, pages 61–73. University of Helsinki.
- Rie Kamikubo, Lining Wang, Crystal Marte, Amnah Mahmood, and Hernisa Kacorri. 2022. Data Representativeness in Accessibility Datasets: A Meta-Analysis. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–15, Athens Greece. ACM.
- Maria Kopf, Marc Schulder, and Thomas Hanke. 2022. The Sign Language Dataset Compendium: Creating an Overview of Digital Linguistic Resources. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 102–109, Marseille, France. European Language Resources Association.
- Annelies Kusters, Maartje De Meulder, and Dai O’Brien, editors. 2017. *Innovations in deaf studies: the role of deaf scholars*. Perspectives on deafness. Oxford University Press, New York.
- Lorraine Leeson, Dimitar Shterionov, Vincent Vandeghinste, Henk van den Heuvel, Josep Blat, Jorn Rijkaert, Myriam Vermeerbergen, Catia Cucchiari, and Aoife Brady. 2021. D9.2 Annual Ethical Report on Practices Towards Stakeholder Communities. Technical report, Trinity College, Dublin.
- Susan McBurney. 2012. History of sign languages and sign language linguistics. In Roland Pfau, Markus Steinbach, and Bencie Woll, editors, *Sign language: An international handbook*, pages 909–948. Walter de Gruyter, Berlin/Boston, MA.
- Hope Morgan, Onno Crasborn, Maria Kopf, Marc Schulder, and Thomas Hanke. 2022. Facilitating the Spread of New Sign Language Technologies across Europe. In *Proceedings of the LREC2022*

- 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 144–147, Marseille, France. European Language Resources Association.
- Savithry Nambodiripad and Jonathan Henner. 2022. Rejecting Competence – Essentialist Constructs Reproduce Ableism and White Supremacy in Linguistic Theory: A Commentary on “Undoing Competence: Coloniality, Homogeneity, and the Overrepresentation of Whiteness in Applied Linguistics”. *Language Learning*.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ashley Shew. 2020. Ableism, Technoableism, and Future AI. *IEEE Technology and Society Magazine*, 39(1):40–85.
- Dimitar Shterionov, Mirella De Sisto, Vincent Vandeghinste, Aoife Brady, Mathieu De Coster, Lorraine Leeson, Josep Blat, Frankie Picron, Marcello Paolo Scipioni, Aditya Parikh, Louis ten Bosh, John O’Flaherty, Joni Dambre, and Jorn Rijckaert. 2022. Sign Language Translation: Ongoing Development, Challenges and Innovations in the SignON Project. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, pages 325–326, Ghent, Belgium. European Association for Machine Translation.
- Julia Silge and David Robinson. 2016. tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *Journal of Open Source Software*, 1(3):37.
- SLLS. SLLS Ethics Statement for Sign Language Research.
- Hadley Wickham. 2022. *rvest: Easily Harvest (Scrape) Web Pages*.
- Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemond, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Pedersen, Evan Miller, Stephan Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. 2019. Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43):1686.
- Hadley Wickham and Dana Seidel. 2022. *scales: Scale Functions for Visualization*.
- Kayo Yin, Amit Moryossef, Julie Hochgesang, Yoav Goldberg, and Malihe Alikhani. 2021. Including Signed Languages in Natural Language Processing. 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 7347–7360, Online. Association for Computational Linguistics.