LONG PAPER

Best practices for sign language technology research

Neil Fox[1] · Bencie Woll[2] · Kearsy Cormier[2]

Abstract

Research on sign language technology (SLT) has steadily increased in recent decades, and yet, common mistakes and pitfalls have significantly hindered progress in the field. The purpose of this paper is to examine some of the most prominent issues and suggest practical steps to overcome them, outlining the best practices to consider when conducting SLT research. These practices cluster around the five following issues: (1) knowledge of the specific sign language at the centre of the research and of sign languages more generally; (2) involving deaf people at the centre of research, including researchers who are themselves deaf; (3) motivations of the researcher and the relationship to the views of the sign language community; (4) what sign language data needs to be considered; (5) full recognition of the challenges posed by such research.

Keywords Sign language · Technology · Deaf technology · Research practices

1 Introduction

Sign language technology (SLT) has become a prominent research area for the computer vision and natural language processing (NLP) communities in the last 30 years [55, 86, 95]. Initial progress has been made in research into technologies that can aid communication between hearing and deaf communities. However, common mistakes have held the field back. As interest expands into this research area, we believe that best practices must be established to enable effective, continued, and long-lasting progress.

In this paper, we detail the most prominent issues that regularly arise in the current SLT research landscape. Often researchers do not fully appreciate the complexity of sign languages and the importance of the deaf community (Sect. 2). There has been a lack of deaf involvement in SLT projects (Sect. 3). SLT research has focused on 'problems' identified by hearing non-signers that are not actually problems at all, whilst some have proposed tools/advancements that have been enormously over-hyped by the media (Sect. 4). The data available for use in SLT have also

been limited (Sect. 5), with an as yet unmet requirement for continuous, diverse sign language datasets. Finally, the complexity of sign language translation has not been fully recognised, as multiple intermediary tasks must be tackled before this can be automated (Sect. 6).

To meet each of these challenges, this paper suggests practical steps, laying out best practice recommendations for SLT research. We hope this work can help establish effective guidelines for both new researchers and incumbents in the field, enabling meaningful progress. The main body of this paper describes the five points of consideration in more detail (Sects. 2–6) with conclusions in Sect. 7.

Before we begin, we wish to provide some context. We are a team of deaf (NF) and hearing (BW, KC) sign language researchers. We are part of and/or have worked closely with British deaf communities for many years, and we are all fluent signers. Because we work primarily on British Sign Language (BSL), most of our observations here relate to BSL, but most hold true for SLT relating to other sign languages as well.

2 Learn about sign languages and deaf people

Sign languages are the languages developed in and used by deaf communities [81]. There are many different sign languages in the world, each with their own grammar and

✉ Kearsy Cormier
  k.cormier@ucl.ac.uk

1   English Language and Linguistics, University
    of Birmingham, Birmingham, UK

2   Deafness, Cognition and Language Research Centre,
    University College London, London, UK

lexicon. The differences in the communicative channels used by spoken and sign languages result in differences in their linguistic structures. For example, spoken languages have access to only a single set of primary articulators (mouth, tongue, lips, teeth), while sign languages have two independent primary articulators (the two hands) and are thus able to make much greater use of simultaneous, rather than linear, grammar [91]. Additionally, in sign languages, communication is necessarily expressed both manually (hands) and non-manually (face and body poses) [84]. Fingerspellings (manual alphabets) are used within sign languages to represent the letters from the ambient written language for specific purposes, such as rendering proper names [65]. Fingerspelled words are distinct and different from the sign language lexicon, which is itself independent of the lexicon of the surrounding spoken/written language [85]. Understanding these complex linguistic features of sign languages is essential in conducting effective SLT research.

In sign languages studied to date, lexical signs are the most frequent form of sign—these are signs that have fairly conventional form and meaning, which can be expressed via one or more 'translation equivalent' words in another language [50] (although it should be noted that just as with translation between any two languages, there is often no one-to-one correspondence between signs and words.) But even lexical signs are produced less than 75% of the time in signed discourse [33]. Much of signed discourse involves pointing and/or depiction. Both pointing and depiction are context-dependent and involve some degree of improvisation. Pointings and depictions rarely look the same or mean the same thing more than once in any signed discourse, which makes them difficult to deal with in a machine learning context [48]. Their unconventionality means that in SLT they are treated as single sign tokens (see Sect. 5 on single tokens).

In addition to learning about how sign languages work, gaining basic deaf awareness is a minimal requirement for researchers in the field [4]. Some assume wrongly that deaf people have the same challenges as people with various disabilities, while others assume that deaf people have the same cultural norms as hearing people. Learning about deaf communities and the different ways in which deaf people view the world is fundamental to producing valid sign language research. Researchers also need to learn how deaf people *do* and *do not* refer to themselves, in order to avoid offensive terminology [12]. For example, terms such as 'deaf and dumb' and 'deaf-mute' are completely unacceptable and their use in sign language technology research has led to retractions by publishers (see, e.g. [39, 58]). In addition, referring to sign languages as 'gestures', 'mimicry' or 'communication tools', or being 'specifically developed for [deaf] people' (as in, e.g. [6, 8, 45, 87]; and many others) are inaccurate and offensive ways to talk about natural human languages. Börstell [9] has shown that this problem of ableist language use when referring to sign languages and deaf communities is far more prevalent in the field of technology than other fields like linguistics, education, and health—reflecting low levels of deaf awareness and deaf involvement in SLT research.

3 Involve deaf people in research

The ultimate aim of SLT research is to develop technology for the deaf community, to aid communication and accessibility. It is only logical that deaf people must then be involved in the research itself [10]. Deaf perspectives bring the engagement with the community that a successful project should seek to include. Ideally, deaf people should be involved at every level including in the planning stages before any work begins [25, 64], yet few projects and publications reflect this level of deaf involvement. Exceptions include work by Vogler and Metaxas [92, 93], Padden and Gunsauls [65], Cormier, Fox, et al. [23], Glasser et al. [40], and EU projects such as EASIER (https://www.project-easier.eu/) and SignOn (https://signon-project.eu/) which have involved deaf organisations at every stage and deaf lay audiences in user testing. One weakness with many projects to date is that engagement has happened too late, after the main development work has taken place, and the perspective has become one of reporting back to the community, rather than ascertaining whether the community thinks the project is worthy in the first place [34].

One danger of involving deaf people in SLT research minimally is tokenism, and this should be avoided. Tokenism is not an issue as long as one aims for allyship instead [46]. To be an ally is to work towards improving deaf representation in the research in various ways—not just as participants, but also as researchers, advisors, investigators. In areas where deaf people are underrepresented in these roles, hearing allies should recruit and train them so that they can be leaders in the future. It should be an aim of the SLT research community to provide not only equal training opportunities for deaf researchers, but additional training, and fast track possibilities where funding allows, to enable professional development, including via non-traditional routes. Such opportunities apply not just in the day-to-day running of research projects but also in presenting the research, e.g. in publications and conference participation. In these contexts, visibility is key, and hearing allies can play a role in shaping this.

For example, hearing researchers who are invited to contribute to a publication or conference or keynote where the topic is focused on SLT should encourage the inclusion of deaf colleagues by the default provision of interpreting at SL conferences (see, e.g. [38]), and by giving space and time to

deaf researchers to showcase their work. Additionally, any workshops and conferences covering sign languages that do not have deaf invited speakers or deaf authors in their proceedings, should be viewed as not deaf-inclusive.

4 Consider the reasons for carrying out the research

When conducting SLT research, ask yourself this: 'What problem am I trying to solve? Is it actually a problem?'. Technology is never going to solve the problem of deaf signers and hearing non-signers understanding each other [41], but it can be used to develop tools to help towards this end [10]. If you are developing a tool, who exactly would use it and for what purpose [59]? By engaging early with deaf people and deaf communities [60, 80], research can better meet their needs and preferences. Some topics, which could actually benefit deaf people, have received insufficient attention from the research community, while some technologies, such as 'data gloves for deaf people', at best have no practical purpose at all [30, 47] and at worst 'perpetuate cultural appropriation and audism' [35].

Another problem to address is that many school and college projects, which are touted as technology which will help deaf people to communicate with the hearing community, are initiated almost exclusively by hearing people. The tools that are developed from these projects are clearly only prototypes, often dealing with limited aspects of communication among deaf people (e.g. recognition of fingerspelled handshapes [51, 77] or with signs in isolation [57]) and receive no further development. More importantly, they often serve no useful purpose to deaf people at all. Despite this, because they appear innovative to hearing non-signers, such projects attract publicity and funding.

In addition to attracting funding, technological projects of this type are often picked up by the media and presented as technology that will remove barriers to communication between deaf and hearing people [11, 20, 21]. Media hype nearly always ends up alienating the deaf community as it comes from a mainly hearing perspective. Just as researchers need deaf perspectives, so do the media. This too would be improved with more deaf people involved in the research from the beginning. These responsibilities should also be shared with the funding bodies and their vetting process. If there was an obligation for funding bodies to ensure that their resources are appropriately allocated, it follows that deaf participation would increase and deaf perspectives would be more realistically reflected. The focus would thus shift from research as a self-perpetuating enterprise to one that aims to provide benefit to the community.

Despite the criticisms outlined above, there are some good candidates for useful SLT: for example, the deaf community might well welcome increased access to smart assistants/home control systems such as Siri and Alexa [31], or the ability to search signed videos [29], or a signed wiki [40]. Unfortunately, without awareness in private sector R&D departments that the needs of deaf people may be fundamentally different to those of hearing people, progress is unlikely.

5 Consider the type of source data needed

Sign language corpora exist for a growing number of sign languages around the world [26, 27, 44, 75, 100]. The sources and uses of these corpora are varied: continuous natural studio-recorded datasets originally designed for linguistic use [44, 63, 75], project-specific isolated studio-recorded datasets [16, 26] or sign interpreted broadcast footage [1, 3, 13, 15, 18, 22, 36], to name a few.

The suitability of each type of dataset for SLT research on recognition and output must be considered before use. When considering which sign language dataset to use for SLT research, there are many important factors to be aware of [10]. These include the diversity of signers present in the data, variability across signers of different ages and different proficiency and age of acquisition, the size of the vocabulary, whether the data are isolated or continuous, whether the data are from a laboratory recording, the internet or broadcast footage, and what types of annotation have been undertaken. Yin et al. [98] provide a detailed breakdown of further properties to consider when selecting an appropriate sign language dataset.

In addition, datasets can vary between a spoken language source (interpreted into a signed language, e.g. with picture-in-picture interpreter) and a sign language source (interpreted via voice-over into a spoken language). The most widely used sign machine learning datasets have consisted of broadcast interpretations, most notably television weather reports [36, 53]. Although these have proved useful, there are considerations and concerns about whether it is appropriate to use datasets from such restricted domains of discourse with limited vocabulary size [16, 26], rather than the large domains found in spontaneous, natural signing [18, 44]. One disadvantage, however, of very large domains such as spontaneous conversation is that many signs are represented with only a handful of instances, which poses difficulties for data-hungry machine learning algorithms.

A critical issue in research of this sort is ensuring that the data to be used as the source material represent the actual target of the analysis: in the case where material interpreted from English into BSL is used as the source data, the question arises of whether material comprised of BSL produced by hearing and deaf interpreters is appropriate as source material for developing automated translation from BSL to

English. Additional questions arise in relation to automated translation from English to BSL. As with spoken languages, deaf fluent signers can and do make grammaticality and acceptability judgements, assessing whether other signers are fluent or not, whether they are native signers or not, and whether they use the language in everyday contexts. In this respect, there are three important questions to be addressed: (1) to what extent does scripted language (whether produced by hearing or deaf people) differ from the spontaneously produced BSL of deaf people?; (2) are there any differences between interpreted and spontaneously produced BSL?; (3) are there any differences between the interpreted BSL produced by hearing interpreters and that produced by deaf interpreters? This final question is of particular relevance in relation to automatic translation of sign language facilitated by recognition of mouthing patterns used by signers, since there is some evidence that hearing and deaf signers differ in their use of mouthing, both in terms of amount and of form [66].

In the general interpreting/translation literature, there is recognition that translated or interpreted language differs from the source (whether spontaneous or scripted) not only in terms of target language but also on a number of dimensions (so-called translation, or interpreting, 'universals' [5, 24, 37, 78, 79]). These differences include a number of features, such as a general tendency towards simplification, and because of their similarity across different source and output texts in different languages, they have been termed 'interpretese'.

Shlesinger and Ordan [79] compared three types of text: interpreted texts, manually transcribed from the spoken outputs of four professional interpreters working in conference settings; translated written texts in (approximately) the same domains, rendered by professional translators; and original semi-scripted speech in (approximately) the same domains by conference presenters. They found that interpreted texts exhibited far more similarities to original speech than to written translation, reflecting that *interpretese* is more *spoken* than *translated*. On the other hand, they found that features such as simplification and lower type-token ratio (which are characteristic of translation) are found to be more salient in interpreted output, as compared to spontaneous language.

There is little literature in the field addressing these questions in relation to sign language interpreting, although there is evidence that there are differences between hearing and deaf interpreters [83, 94]. Stone [83] addresses differences between hearing and deaf interpreters in the process of preparing a sign language interpretation from an English script, by examining prosodic features of these interpretations, for example in the use of non-manual features such as mouthing (with hearing interpreters more likely to produce multisyllabic mouthings). Additionally, signed translations

can either be done from written scripts via autocue in real time or interpreted from a spoken language in real time. For broadcast television material, there are differences between deaf and hearing interpreters, although both produce their final version 'live'. Hearing interpreters, although they have access to a written script to prepare, undertake limited preparation from these written forms and rely to a greater extent on hearing the spoken version to interpret in real time. In contrast, deaf interpreters prepare extensively from the written text, enabling them to create a translation rather than an interpretation, using the autocue to support the final version in real time [83].

In relation to the question of possible differences between deaf and hearing interpreters, silent mouthing of words from spoken language is also one possible source of difference. It is known that mouthing differs along sociolinguistic parameters such as region, gender, age, nativeness, and level of education, even among deaf signers [7, 68]. No studies have explicitly explored mouthing differences between interpreters on the basis of hearing status. However, this is a topic worthy of research.

Another issue in considering choice of data source regardless of whether it is interpreted or not is annotation. In order to computationally process sign datasets, time-aligned machine-readable annotation is necessary. For machine learning SLT research, these annotations should be accurate and exhaustive, with detailed segmentation and ideally gloss labelling of each sign. However, this process is significantly labour-intensive and requires fluent signers.

It is also important to consider the extraction of data for computational processing, most commonly pose keypoints [19, 99]. Computer models are able to accurately estimate 2D body pose [19], but hand pose estimation, especially with two hands, is still very challenging [43]. Recent work of Moryossef et al. [62] has shown that human body pose estimation quality is potentially a limiting factor when used for SLT and requires further research. To optimise pose estimation results, datasets of higher quality and resolution must be adopted [18].

6 Recognise the challenges of automatic sign language analysis

Automatic translation between signed and spoken languages is the ultimate aim of many SLT projects [18, 23, 28], yet this task is incredibly complex [89]. The computer science community often underestimates the linguistic complexity of sign languages and treats automatic translation as a standard video-to-text/text-to-video problem or as similar to a simple gesture recognition/production problem [56, 90]. This oversimplifies translation

models, leading to inaccurate end results and ultimately poor access for deaf people [52, 97].

There are substantial differences between an automated sign-to-spoken language translation process compared to automated translation between two spoken languages. Spoken language translation can involve speech-to-text as a first stage, followed by translation from source language text to target language text, followed by text-to-speech. Sign languages lack a written form [84], and they must be represented in a continuous format for computation [71], in contrast to the discrete representation of written language. Therefore, there is a requirement for bespoke architectures specific to sign language.

In addition, when tackling sign language translation, the natural variability in human translation must be captured, as there is more than one way to translate an utterance between a spoken and a signed language, just as between spoken languages. Many translations are equally valid, but some may be judged as better or more accurate than others. It is important that computational models use judgements of accuracy to take natural variability into account. Currently, the most common SLT research areas are sign recognition: the recognition of isolated lexical signs from a video [42, 57]; sign language translation: the translation of sign language videos to continuous spoken language [17, 18]; and sign language production: the generation of sign language content from spoken language [73, 82, 96].

Although recognition is a logical first step when tackling a full automated translation task, any application of isolated sign recognition has limited use to deaf people. Isolated sign recognition is useful for some tasks such as searching for individual signs in videos and dictionaries but not for recognising sign language discourse. This continued focus on isolated recognition is indicative of a lack of progress in the field [10, 53] and a lack of understanding of sign language and deaf needs. Although continuous sign translation [15] and production [71] are much harder tasks, they are considerably more helpful as tools. SLT research must turn towards continuous translation and production to progress.

However, before unconstrained sign language translation can be achieved, there are multiple additional intermediary tasks in sign language processing that must be tackled. Current intermediate problems include, but are not limited to, active signer detection [2], subtitle alignment [14], sign segmentation [69], visual anonymisation [70, 72], visual representation learning [3], continuous recognition [54], sign animation [74], sign spotting [61, 89], fingerspelling detection [67, 76], detailed 3D human shape estimation [32, 49], facial expressions, head pose and body movements [88], and multi-signer scenarios [64].

7 Conclusions

In this paper, we have outlined the current state of sign language technology (SLT) research, arguing that progress has been hindered by five prominent issues. To tackle this, we have proposed best practices every researcher should consider when conducting SLT research. We hope the insights provided here will enhance progress and value in the field for both hearing and deaf people.

Declarations

References

1. Albanie, S., Varol, G., Momeni, L., Bull, H., Afouras, T., Chowdhury, H., Fox, N., Woll, B., Cooper, R., McParland, A., Zisserman, A.: BOBSL: BBC-Oxford British Sign Language Dataset. https://arxiv.org/abs/2111.03635 (2021a)

2. Albanie, S., Varol, G., Momeni, L., Afouras, T., Brown, A., Zhang, C., Coto, E., Camgöz, NC., Saunders, B., Dutta, A., Fox, N., Bowden, R., Woll, B., Zisserman, A.: Signer diarisation in the wild. https://www.robots.ox.ac.uk/~vgg/publications/2021/Albanie21a/albanie21a.pdf (2021b)

3. Albanie, S., Varol, G., Momeni, L., Afouras, T., Chung, J.S., Fox, N., Zisserman, A.: BSL-1K: scaling up co-articulated sign language recognition using mouthing cues. In: Comp Vis–ECCV

2020: 16th Europ Conf Proc, Part XI 16. Springer International, New York, pp 35–53, (2020) doi:https://doi.org/10.48550/arXiv.2007.12131

4. Atherton, M.: A feeling as much as a place: leisure, deaf clubs and the British deaf community. Leis Stud **28**(4), 443–454 (2009). https://doi.org/10.1080/02614360902951690

5. Baker, M.: Corpus Linguistics and Translation Studies: Implications and Applications. In: Baker, M., Francis, G., Tognini-Bonelli, E. (eds.) Text and Technology In Honour of John Sinclair, pp. 233–250. John Benjamins, Amsterdam (1993). https://doi.org/10.1075/z.64

6. Batnasan, G., Gochoo, M., Otgonbold, M.E., Alnajjar, F., Shih, T.K.: ArSL21L: Arabic sign language letter dataset benchmarking and an educational avatar for metaverse applications. In: 2022 IEEE Glob Eng Ed Conf (EDCON). IEEE, New York, pp. 1814–1821 (2022). doi:https://doi.org/10.1109/EDUCON52537.2022.9766497

7. Bauer, A.: How words meet signs: a corpus-based study on variation of mouthing in Russian Sign Language. Linguistische Beiträge zur Slavistik **24**, 9–35 (2019)

8. Bilgin, M., Mutludoğan, K.: American sign language character recognition with capsule networks. 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE, New York, pp 1–6, (2019). doi:https://doi.org/10.1109/ISMSIT.2019.8932829

9. Börstell, C.: Ableist language teching over sign language research. In: Proc 2nd Workshop on Resources and Representations for Under-Resourced Languages and Domains (RESOURCEFUL-2023), pp. 1–10, (2023). https://aclanthology.org/2023.resourceful-1.0

10. Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A. et al.: Sign language recognition, generation, and translation: An interdisciplinary perspective. Proc 21st International ACM SIGACCESS Confernece on Computers and Accessibility, pp. 16–31. (2019). doi:https://doi.org/10.1145/3308561.3353774

11. Bridge, M.: Google's Next Translation: Sign Language. The Times, London, 26 August 2019. (2019). https://www.thetimes.co.uk/article/googles-next-translation-sign-language-gvnmldjw3

12. British Deaf Association (BDA): British Deaf Association—Definitions of Hearing Impairments. (2017). https://www.derbyshire.gov.uk/site-elements/documents/pdf/social-health/adult-care-and-wellbeing/disability-support/hearing-impaired/british-deaf-association-definitions-of-hearing-impairments.pdf

13. Buehler, P., Zisserman, A., Everingham, M.: Learning sign language by watching TV (using weakly aligned subtitles). In: 2012 IEEE Conference Computer and Vision and Pattern Recognition. pp. 2961–2968. (2009). doi:https://doi.org/10.1109/CVPRW.2009.5206523

14. Bull, H., Afouras, T., Varol, G., Albanie, S., Momeni, L., Zisserman, A.: Aligning subtitles in sign language videos. (2021). arXiv Preprint. https://arxiv.org/abs/2105.02877

15. Camgöz, N.C., Hadfield, S., Koller, O., Ney, H., Bowden, R.: Neural Sign Language Translation. In: Proceedings of IEEE Confernce on Computer Vision Pattern Recognition (CVPR) (2018). doi:https://doi.org/10.1109/CVPR.2018.00812

16. Camgöz, N.C., Kındıroğlu, A.A., Karabüklü, S., Kelepir, M., Ozsoy, A.S., Akarun, L.: BosphorusSign: a Turkish Sign Language recognition corpus in health and finance domains. In: Proc 10th Intl Conf on Lang Resources and Eval (LREC'16), pp. 1383–1388. (2016). https://aclanthology.org/L16-1220

17. Camgöz, N.C., Koller, O., Hadfield, S., Bowden, R.: Sign language transformers: Joint end-to-end sign language recognition and translation. In: Proc IEEE Conf on Comp Vis and Pattern Recognit. (CVPR), pp. 10023–10033. (2020). doi:https://doi.org/10.1109/CVPR42600.2020.01004

18. Camgöz, N.C., Saunders, B., Rochette, G., Giovanelli, M., Inches, G., Nachtrab-Ribback R, et al.: Content4All Open Research Sign Language Translation Datasets. IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), pp. 1–5. (2021). doi:https://doi.org/10.1109/FG52635.2021.9667087

19. Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., Sheikh, Y.: OpenPose: realtime multi-person 2D pose estimation using part affinity fields. Proc IEEE Conf. Comput. Vis. Pattern Recognit (CVPR) **43**(1), 172–186 (2017). https://doi.org/10.1109/TPAMI.2019.2929257

20. Coldewey, D.: SignAll is slowly but surely building a sign language translation platform. (2018). https://techcrunch.com/2018/02/14/signall-is-slowly-but-surely-building-a-sign-language-translation-platform

21. Coldewey, D.: SLAIT's real-time sign language translation promises more accessible online communication. (2021). https://techcrunch.com/2021/04/26/slaits-real-time-sign-language-translation-promises-more-accessible-online

22. Cooper, H., Bowden, R.: Learning signs from subtitles: a weakly supervised approach to sign language recognition. In: IEEE Conf on Comp Vis and Pattern Recognition, pp. 2568–2574. (2009). doi:https://doi.org/10.1109/CVPR.2009.5206647

23. Cormier, K., Fox, N., Woll, B., Zisserman, A., Camgöz, N.C., Bowden, R.: ExTOL: automatic recognition of british sign language using the BSL corpus. In: Proc 6th Workshop on Sign Language Translation and Avatar Technology (SLTAT) (2019). https://openresearch.surrey.ac.uk/esploro/outputs/conferencePresentation/ExTOL-Automatic-recognition-of-British-Sign-Language-using-the-BSL-Corpus/99514750802346

24. Dayter, D.: Collocations in Non-Interpreted and Simultaneously Interpreted English. In: Vandevoorde, L., Daems, J., Defrancq, B. (eds.) New Empirical Perspectives on Translation and Interpreting, pp. 67–91. Routledge, Abingdon (2019). https://doi.org/10.4324/9780429030376-4

25. De Meulder, M.: Is "Good Enough" Good Enough? Ethical and Responsible Development of Sign Language Technologies. Proc 18th Biennial Machine Translation Summit, 1st Intl Workshop on Automatic Translation for Signed and Spoken Languages, Vol 1. (2021). https://www.semanticscholar.org/paper/Is-%E2%80%9Cgood-enough%E2%80%9D-good-enough-Ethical-and-of-sign-Meulder/590d4da2864b57f05e249b02dc1c1778d39b192e

26. Ebling, S., Camgöz, N.C., Braem, P.B., Tissi, K., Sidler-Miserez, S., Stoll, S., et al.: SMILE Swiss German Sign Language Dataset. Proc Intl Conf on Language Resources and Evaluation (LREC). (2018). http://www.lrec-conf.org/proceedings/lrec2018/pdf/25.pdf

27. Efthimiou, E., Fotinea, S.E.: GSLC: creation and annotation of a Greek sign language corpus for HCI. In Proc Univ Access in Hum Comp Interaction: Coping with Diversity: 4th Intl Conf on Universal Access in Hum-Comp Interact I:4. Springer, Abingdon pp. 657–666/ (2007). https://link.springer.com/chapter/https://doi.org/10.1007/978-3-540-73279-2_73

28. Efthimiou, E., Fotinea, S-E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., et al.: Sign language technologies and resources of the dicta-sign project. In: Proc 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon (LREC), pp. 35–44. (2012). http://www.sign-lang.uni-hamburg.de/lrec/pub/12025.html

29. Elliott, R., Cooper, H., Ong, E-J., Glauert, J., Bowden, R., Lefebvre-Albaret, F.: Search-by-Example in Multilingual Sign Language Databases. In: 2nd Intl. Workshop on Sign Language Translation and Avatar Technology (SLTAT) (2011). http://personal.ee.surrey.ac.uk/Personal/H.Cooper/research/papers/SBE_SLTAT.pdf

30. Erard, M.: Why Sign-Language Gloves Don't Help Deaf People. The Atlantic 9 November 2017. (2017). https://www.theatlantic.com/technology/archive/2017/11/why-sign-language-gloves-dont-help-deaf-people/545441/

31. Evans, J.: Apple's accessibility tools are changing the world. Apple Must 25 June 2020. (2020). https://www.applemust.com/apples-accessibility-tools-are-changing-the-world/

32. Feng, Y., Choutas, V., Bolkart, T., Tzionas, D., Black, M.J.: Collaborative regression of expressive bodies using moderation. (2021). arXiv Preprint. https://arxiv.org/abs/2105.05301

33. Fenlon, J., Schembri, A., Rentelis, R., Vinson, D., Cormier, K.: Using conversational data to determine lexical frequency in British Sign Language: the influence of text type. Lingua **143**, 187–202 (2014). https://doi.org/10.1016/j.lingua.2014.02.003

34. Ferndale, D.: "Nothing About Us Without Us": navigating engagement as hearing researcher in the Deaf Community. Qual. Res. Psychol. **15**(4), 437–455 (2018). https://doi.org/10.1080/14780887.2017.1416802

35. Forshay, L., Winter, K., Bender, E., et al.: University of Washington Letter in Response to SignAloud. (2016). http://depts.washington.edu/asluw/SignAloud-openletter.pdf

36. Forster, J., Schmidt, C., Hoyoux, T., Koller, O., Zelle, U., Piater, J., Ney, H.: RWTH-PHOENIX-Weather: A large vocabulary sign language recognition and translation corpus. In: Proc. Intl. Conf. Lang. Resour. Eval. 2012 (LREC). (2012). http://www.lrec-conf.org/proceedings/lrec2012/pdf/844_Paper.pdf

37. Fu, R., Wang, K.: Hedging in interpreted and spontaneous speeches: a comparative study of Chinese and American political press briefings. Text and Talk **42**(2), 153–175 (2022). https://doi.org/10.1515/text-2019-0290

38. Gawne, L., Hodge, G.: Planning communication access for online conferences. (2021). https://researchwhisperer.org/2021/12/21/planning-accessible-online-conferences/

39. Ghule, S., Chavaan, M.: (2021 - retracted). Implementation of hand gesture recognition system to aid deaf-dumb people. Advances in Signal and Data Processing. Retracted version doi:https://doi.org/10.1007/978-981-15-8391-9_14. Retraction Note at https://link.springer.com/chapter/https://doi.org/10.1007/978-981-15-8391-9_49

40. Glasser, A., Minakov, F., Bragg, D.: ASL Wiki: an Exploratory Interface for Crowdsourcing ASL Translations. In: Proc 24th Intll ACM SIGACCESS Conf. Comput. Accessibility (ASSETS '22). Assoc for Computing Machinery Article 16, 1–13. (2022). doi:https://doi.org/10.1145/3517428.3544827

41. Grieve-Smith, A.: 10 Reasons why sign-to-speech technology won't be practical anytime soon. (2016). https://limpingchicken.com/2016/05/04/angus-grieve-smith-10-reasons-why-sign-to-speech-technology-wont-be-practical-anytime-soon/

42. Grobel, K., Assan, M.: Isolated sign language recognition using hidden Markov models. In: Proc 1997 IEEE Intl Conf. Syst., Man, Cybern. 1: 162-167. (1997). doi:https://doi.org/10.1109/ICSMC.1997.625742

43. Hampali, S., Sarkar, S.D., Rad, M., Lepetit, V.: Solving joint identification in challenging hands and object interactions for accurate 3D pose estimation. (2021). arXiv Preprint. https://arxiv.org/abs/2104.14639

44. Hanke, T., König, L., Wagner, S., Matthes, S.: DGS Corpus and Dicta-Sign: the Hamburg Studio Setup. 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010) (2010). https://www.sign-lang.uni-hamburg.de/lrec2010/lrec_cslt_01.pdf

45. He, S.: Research of a sign language translation system based on deep learning. In: Int. Conf. Artif. Intell. Adv. Manuf. (AIAM), pp. 392–396. (2019). doi:https://doi.org/10.1109/AIAM48774.2019.00083

46. Hearing Allyship (2021) Guiding Principles for Hearing Allyship. https://www.hearingallyship.org/

47. Hill, J.: Do deaf communities actually want sign language gloves? Nat. Electron. **3**(9), 512–513 (2020). https://doi.org/10.1038/s41928-020-0451-7

48. Jantunen, T., Rousi, R., Rainò, P., Turunen, M., Moeen Valipoor, M., García, N.: Is There Any Hope for Developing Automated Translation Technology for Sign Languages? In: Hämäläinen, M., Partanen, N., Alnajjar, K. (eds.) Multilingual Facilitation, pp. 61–73. University of Helsinki, Rootroo (2021). https://doi.org/10.31885/9789515150257.7

49. Jiang, T., Camgöz, N.C., Bowden, R.: Skeletor: Skeletal Transformers for Robust Body-Pose Estimation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition: pp. 3394–3402. (2021). https://ieeexplore.ieee.org/document/9522847

50. Johnston, T., Schembri, A.C.: On defining lexeme in a signed language. Sign Lang & Ling **2**(2), 115–185 (1999). https://doi.org/10.1075/sll.2.2.03joh

51. Kim, T., Shakhnarovich, G., Livescu, K.: Fingerspelling recognition with semi-Markov conditional random fields. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). (2013). doi:https://doi.org/10.1109/ICCV.2013.192

52. Kipp, M., Nguyen, Q., Heloir, A., Matthes, S.: Assessing the deaf user perspective on sign language avatars. In: Proc 13th Int. ACM SIGACCESS Conf. Comput. Access. (ASSETS), pp.107–114. (2011). doi:https://doi.org/10.1145/2049536.2049557

53. Koller, O.: Quantitative survey of the state of the art in sign language recognition. (2020). *arXiv Preprint*. doi:https://doi.org/10.48550/arXiv.2008.09918

54. Koller, O., Forster, J., Ney, H.: Continuous sign language recognition: towards large vocabulary statistical recognition systems handling multiple signers. Comput. Vis Image Underst. (CVIU) **141**, 108–125 (2015). https://doi.org/10.1016/j.cviu.2015.09.013

55. Liang, R.-H. & Ouhyoung, M.: A sign language recognition system using hidden Markov model and context sensitive search. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology (pp. 59–66). (1996). doi:https://doi.org/10.1145/3304181.3304194

56. Liang, R-H., Ouhyoung, M.: A real-time continuous gesture recognition system for sign language. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 558–567. (1998). doi:https://doi.org/10.1109/AFGR.1998.671007

57. Lim, K.M., Tan, A.W.C., Lee, C.P., Tan, S.C.: Isolated sign language recognition using Convolutional Neural Network hand modelling and Hand Energy Image. Multimed. Tools Appl. **78**, 19917–19944 (2019). https://doi.org/10.1007/s11042-019-7263-7

58. Marcus, A.: Springer Nature to retract chapter on sign language critics call "unbelievably insulting". Retraction Watch, February 2021. (2021). https://retractionwatch.com/2021/02/01/springer-nature-to-retract-chapter-on-sign-language-critics-call-unbelievably-insulting/

59. Matchar, E.: Sign language translating devices are cool. But Are They Useful? Smithsonian Magazine. February 2019. (2019). https://www.smithsonianmag.com/innovation/sign-language-translators-are-cool-but-are-they-useful-180971535/

60. McKee, M., Schlehofer, D., Thew, D.: Ethical issues in conducting research with deaf populations. Am. J. Public Health **103**(12), 2174–2178 (2013). https://doi.org/10.2105/AJPH.2013.301343

61. Momeni, L., Varol, G., Albanie, S., Afouras, T., Zisserman, A.: Watch, read and lookup: learning to spot signs from multiple supervisors. Proc. Asian Conf. Comput. Vis. (2020). https://doi.org/10.1007/978-3-030-69544-6_18

62. Moryossef, A.: Tsochantaridis, I., Dinn, J., Camgöz, N.C., Bowden, R. et al.: Evaluating the immediate applicability of pose estimation for sign language recognition. In: Proceedings of IEEE/CVF Confernce on Computer Vision and Pattern Recognition, pp. 3434–3440. (2021). doi:https://doi.org/10.1109/CVPRW53098.2021.00382

63. Neidle, C., Thangali, A., Sclaroff, S.: Challenges in the development of the American Sign Language Lexicon Video Dataset (ASLLVD) Corpus. In: Proceedings of 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon. International Conference on Language Resources and Evaluation (LREC), pp.143–150. (2012). https://www.sign-lang.uni-hamburg.de/lrec/pub/12027.html

64. Núñez-Marcos, A., de Viñaspre, O.P., Labaka, G.: A survey on sign language machine translation. Expert Syst. Appl. **213**, 118993 (2023). https://doi.org/10.1016/j.eswa.2022.118993

65. Padden, C.A., Gunsauls, D.C.: How the alphabet came to be used in a sign language. Sign Lang. Stud. **4**(1), 10–33 (2003). https://doi.org/10.1353/sls.2003.0026

66. Perniss, P., Vinson, D., Vigliocco, G.: Making sense of the hands and mouth: the role of "secondary" cues to meaning in British Sign Language and English. Cognit. Sci. **44**(7), e12868 (2020). https://doi.org/10.1111/cogs.12868

67. Prajwal, K.R., Bull, H., Momeni, L., Albanie, S., Varol, G., Zisserman, A.: Weakly-supervised Fingerspelling Recognition in British Sign Language Videos British Machine Vision Conference. (2022). https://bmvc2022.mpi-inf.mpg.de/609/

68. Proctor, H., Cormier, K.: Sociolinguistic variation in mouthings in British Sign Language (BSL): a corpus-based study. Lang. Speech **66**(2), 1–30 (2022). https://doi.org/10.1177/00238309221107002

69. Renz, K., Stache, N.C., Albanie, S., Varol, G.: Sign language segmentation with temporal convolutional networks. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) pp. 2135–2139. (2021). doi:https://doi.org/10.1109/ICASSP39728.2021.9413817

70. Saunders, B., Camgöz, N.C., Bowden, R.: Everybody sign now: translating spoken language to photo realistic sign language video. (2020a). arXiv Preprint. https://arxiv.org/2011.09846

71. Saunders, B.,Camgöz, N.C., Bowden, R.: Progressive transformers for end-to-end sign language production. In: Proceedings of European Confernce on Computer Vision (ECCV) (2020b). doi:https://doi.org/10.1007/978-3-030-58621-8_40

72. Saunders, B., Camgöz, N.C., Bowden, R.: AnonySign: novel human appearance synthesis for sign language video anonymisation. In: 16th IEEE Intl Conf on Automatic Face and Gesture Recognition (FG 2021), pp. 1–8. (2021a). doi:https://doi.org/10.1109/FG52635.2021.9666984

73. Saunders, B., Camgöz, N.C., Bowden, R.: Continuous 3D multichannel sign language production via progressive transformers and mixture density networks. Int. J. Comput. Vis. **129**, 1–23 (2021). https://doi.org/10.1007/s11263-021-01457-9

74. Saunders, B., Camgöz, N.C., Bowden, R.: Mixed SIGNals: sign language production via a mixture of motion primitives. In: Proceedings of International Conferencen Computer Vision (ICCV), pp. 1899–1909. (2021c). doi:https://doi.org/10.1109/ICCV48922.2021.00193

75. Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S., Cormier, K.: Building the British Sign Language Corpus. Lang Documentation & Conservation 7: 136–154. (2013). http://hdl.handle.net/10125/4592

76. Shi, B., Brentari, D., Shakhnarovich, G., Livescu, K.: Fingerspelling detection in American Sign Language. In: Proceedings of 60th Annual Mtg of the Associate for Computational Linguistics (Volume 1: Long Papers), pp. 1699–1712. (2022). https://aclanthology.org/2022.acl-long.119/

77. Shi, B., Rio, A.M.D., Keane, J., Brentari, D., Shakhnarovich, G., Livescu, K.: Fingerspelling Recognition in the Wild with Iterative Visual Attention. In: Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV) 5399–5408. (2019). doi:https://doi.org/10.1109/ICCV.2019.00550

78. Shlesinger, M.: Towards a Definition of Interpretese: An Intermodal, Corpus-Based Study. John Benjamins Publishing Company, Amsterdam (2009). https://doi.org/10.1075/btl.80.18shl

79. Shlesinger, M., Ordan, N.: More spoken or more translated?: Exploring a known unknown of simultaneous interpreting. Target. Int. J. Transl. Stud. **24**(1), 43–60 (2012). https://doi.org/10.1075/target.24.1.04shl

80. Singleton, J.L., Jones, G., Hanumantha, S.: Toward ethical research practice with deaf participants. J. Empir. Res. Hum. Res. Ethics **9**(3), 59–66 (2014). https://doi.org/10.1177/1556264614540589

81. Stokoe, W.C.: Sign Language Structure. Ann. Rev. Anthropol. **9**(1), 365–390 (1980). https://doi.org/10.1146/annurev.an.09.100180.002053

82. Stoll, S., Camgöz, N.C., Hadfield, S., Bowden, R.: Sign language production using neural machine translation and generative adversarial networks. In Proceedings of British Machine Vision Conference (BMVC). (2018). http://www.bmva.org/bmvc/2018/contents/papers/0906.pdf

83. Stone, C.: Toward a Deaf Translation Norm. Gallaudet University Press, Washington (2009). https://doi.org/10.2307/j.ctv2rcng24

84. Sutton-Spence, R., Woll, B.: The Linguistics of British Sign Language: An Introduction. Cambridge University Press, Cambridge (1999). https://doi.org/10.1017/CBO9781139167048

85. Sutton-Spence, R., Woll, B., Allsop, L.: Variation and recent change in fingerspelling in British sign language. Lang. Var. Change **2**(3), 313–330 (1990). https://doi.org/10.1017/S0954394500000399

86. Tamura, S., Kawasaki, S.: Recognition of sign language motion images. Pattern Recognit. **21**(4), 343–353 (1988). https://doi.org/10.1016/0031-3203(88)90048-9

87. Tyagi, A., Bansal, S.: Feature extraction technique for vision-based indian sign language recognition system: a review. Comput. Methods Data Eng. (2021). https://doi.org/10.1007/978-981-15-6876-3_4

88. Tze, C., Filntisis, P., Dimou, A., Roussos, A., Maragos, P.: Neural sign reenactor: deep photorealistic sign language retargeting. (2022). Archiv Preprint. https://arxiv.org/abs/2209.01470

89. Varol, G., Momeni, L., Albanie, S., Afouras, T., Zisserman, A.: Read and attend: Temporal Localisation in Sign Language Videos. In: Proceedings of IEEE Confernce on Computer Vision and Pattern Recognition (CVPR) 16852–16861. (2021). doi:https://doi.org/10.1109/CVPR46437.2021.01658

90. Verma, H.V., Aggarwal, E., Chandra, S.: (2013) Gesture recognition using kinect for sign language translation. In: IEEE 2nd International Conference on Image Information Processing (ICIIP), pp. 96–100. doi:https://doi.org/10.1109/ICIIP.2013.6707563

91. Vermeerbergen, M., Leeson, L., Crasborn, O.A.: Simultaneity in Signed Languages: Form and Function. John Benjamins Publishing, Amsterdam (2007). https://doi.org/10.1075/cilt.281

92. Vogler, C., Metaxas, D.: A framework for recognizing the simultaneous aspects of American Sign Language. Comput. Vis. Image Underst. **81**(3), 358–384 (2001). https://doi.org/10.1006/cviu.2000.0895

93. Vogler, C., Metaxas, D.: Handshapes and movements: multiple-channel american sign language recognition. In: Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop (GW 2003) Selected Revised Papers

5, pp. 247–258. (2003). doi:https://doi.org/10.1007/978-3-540-24598-8_23

94. Wehrmeyer, E.: Linguistic Interference in Interpreting from English to South African Sign Language. In: Hickey, R. (ed.) English in Multilingual South Africa: The Linguistics of Contact and Change, pp. 371–393. Cambridge University Press, Cambridge (2019). https://doi.org/10.1017/9781108340892.018

95. Wilson, B.J., Anspach, G.: Neural networks for sign language translation. Applications of Artificial Neural Networks IV, Vol 1965, pp. 589–599.In: International Society for Optics and Photonics (SPIE). (1993). doi:https://doi.org/10.1117/12.152560

96. Wolfe, R., McDonald, J.C., Hanke, T., Ebling, S., Van Landuyt, D., et al.: Sign language avatars: a question of representation. Information **13**(4), 206 (2022). https://doi.org/10.3390/info13040206

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

98. Yin, K., Moryossef, A., Hochgesang, J., Goldberg, Y., Alikhani, M.: Including signed languages in natural language processing. In: Proceedings of 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Vol 1: Long Papers. Assoc for Computational Linguistics). (2021). https://aclanthology.org/2021.acl-long.570.pdf

99. Zelinka, J., Kanis, J.: Neural sign language synthesis: words are our glosses. In: IEEE Winter Confernce on Applications of Computer Vision (WACV), pp.3384–3392. (2020). doi:https://doi.org/10.1109/WACV45572.2020.9093516

100. Zhang, J., Zhou, W., Xie, C., Pu, J., Li, H.: Chinese Sign Language Recognition with Adaptive HMM. In: 2016 IEEE International Conference on Multimedia and Expo (ICME). (2016). doi:https://doi.org/10.1109/ICME.2016.7552950